iCIIT 2023

The 4th International Conference on Innovations in Info-business & Technology (ICIIT 2023)

04th August 2023 at Galadari Hotel, Colombo

CONFERENCE PROCEEDINGS



Organized by

INFORMATICS INSTITUTE OF TECHNOLOGY

Platinum Sponsor

SLTMOBITEL

Gold Sponsor

AXIATA DIGITAL LABS

Silver Sponsor

ZONE24x7

Bronze Sponsors

NCINGA

imark

Technical Sponsors

Event Partners

CSSL COMPUTER SOCIETY OF SRI LANKA

BCS Sri Lanka Section

FITIS

ISACA Sri Lanka Chapter

SLASSCOM

Electronic Media Partner

Hi!TV

Print Media Sponsors

Daily FT

The Sunday Times

Education Times

Hospitality Partner

GALADARI COLOMBO

PROCEEDINGS OF THE 4TH INTERNATIONAL CONFERENCE ON INNOVATIONS IN INFO-BUSINESS & TECHNOLOGY (ICIIT 2023)

Unlocking digital potential for economic growth in emerging regions

04th August 2023 at Galadari Hotel, Colombo



Informatics Institute of Technology
57, Ramakrishna Road,
Colombo 06,
Sri Lanka.
www.iciit.iit.ac.lk

DISCLAIMER

The proceedings comprise the papers presented at the 4th International Conference on Innovations in Info-business & Technology (ICIIT 2023), held on the 04th August 2023 at Galadari Hotel, Colombo.

The responsibility for opinions expressed in articles, in studies and other contributions in this publication rests solely with the respective authors. The 4th International Conference on Innovations in Info-business & Technology (ICIIT 2023) of Informatics Institute of Technology shall have no liability or responsibility to any person or entity regarding any loss or damage incurred, or alleged to have incurred, directly or indirectly, by the information contained in this book.

Edited by the Editorial Board – ICIIT 2023

- Dr. Kathiravelu Thabotharan (Editor-in-Chief)
- Dr. Rohan Samarasinghe (Editor)
- Prof. D.M.G. Preethichandra (Editor)
- Dr. Chitraka Wickramarachchi (Editor)
- Prof. Nimal Wickramasinghe (Editor)
- Dr. Frederic Kleinermann (Editor)
- Dr. Dharshana Kasthurirathna (Editor)
- Dr. Chamli Pushpakumara (Editor)

Published by
Informatics Institute of Technology
57, Ramakrishna Road,
Colombo 06,
Sri Lanka.
www.iciit.iit.ac.lk

Published date
04th August 2023

ISSN 3021-6036
© Informatics Institute of Technology, Sri Lanka

4TH INTERNATIONAL CONFERENCE ON INNOVATIONS IN INFO-BUSINESS & TECHNOLOGY (ICIIT 2023)

The 4th annual "International Conference on Innovations in Info-business & Technology (ICIIT 2023)" will be held on 04th August 2023 at Galadari Hotel, Colombo under the theme "Unlocking digital potential for economic growth in emerging regions". ICIIT will bring together academia, industry, technologists, and research scholars. The conference will have keynote addresses by leading international research scholars, workshops by domain experts, research paper presentations, a panel discussion and a poster presentation.

The conference this year is chaired by Prof. Prasad Wimalaratne, who obtained his PhD from University of Salford, UK, who is a professor of Computing at University of Colombo School of Computing and also a Professor, teaching at Informatics Institute of Technology.

There are 27 papers presented this year under the three tracks of Data Science & AI, Computing and Technology of Business. The keynote will be delivered by Professor Umashanger Thayasivam who is a professor in Statistics and Data Science at Dept. of Mathematics, Rowan University, New Jersey. USA. There will be four other invited speeches by research scholars and the paper presentations.

The time is right for more research work to happen in bringing out ideas for our research community and this would be an ideal opportunity to meet up to discuss on research topics that are currently worked upon for researchers and academics.



MESSAGE FROM THE FOUNDER AND CHAIRMAN,

INFORMATICS GROUP OF COMPANIES, SRI LANKA

It is a pleasure to extend a warm welcome to all participants of ICIIT 2023, focusing on the theme of "Unlocking digital potential for economic growth in emerging regions." This conference provides an ideal platform for academics, industry professionals, and students to come together in one forum and engage in discussions about potential challenges and solutions.

ICIIT 2023 offers a valuable opportunity for participants to connect with one another, fostering collaboration in areas of mutual interest. Given the current circumstances in Sri Lanka, the conference emphasizes innovation and technology as key focus areas for developing new product ideas and propelling Sri Lanka towards a manufacturing economy. The conference aims to inspire participants to think innovatively and take action in the future based on the ideas generated during the event.

Informatics Group has been at the forefront of providing quality IT solutions and leading in IT education, as well as making significant contributions to the agriculture sector. Informatics Institute of Technology (IIT) stands as the first private higher education institute awarding high quality British degrees. Having produced over five thousand graduates and awarded more than nine hundred scholarships, IIT has played a crucial role in shaping skilled professionals for various industries.

I extend my gratitude to the organizing committee and all individuals involved in making this conference possible. I wish everyone a successful conference full of ideas and meaningful discussions that will lead to timely actions, benefiting the scientific community and our country. I hope all overseas participants have a memorable stay in Sri Lanka. Let us seize this opportunity to collaborate, share knowledge, and explore innovative solutions that will contribute to the growth and progress of emerging regions. Together, we can make a positive impact and drive our collective vision for a brighter future in the digital era.



MESSAGE FROM THE CHIEF EXECUTIVE OFFICER,

INFORMATICS INSTITUTE OF TECHNOLOGY, SRI LANKA

ICIIT 2023 is another very important day in the IIT calendar. Since its inception in 1990, it has been in the forefront, transforming Sri Lankan students to be innovative and tech savvy bringing the state-of-the-art technology to Sri Lanka. This year's theme "Unlocking digital potential for economic growth in emerging regions" is very timely to push our nation ahead on its way to economic recovery.

The current generation of students have faced many issues due to pandemic and subsequent challenges due to economic crisis. Education has been confined to digital media which has led to social issues for some students. We are now at a juncture where a right mix of digital and physical learning are clubbed together to enable our students to reach the heights of higher education to push our country forward. The ICIIT conference will provide an opportunity and a platform for research, learning and industry to meet in one stage to look at problems and practical solutions.

At IIT we encourage research, more so applied research. We strive to take the research one step ahead to make it into a reality to create value for organizations. We are also looking at helping this through data driven decisions which can improve the quality of decisions taken, being more informative and forward looking.

We would like to see the ideas that stem from ICIIT to be taken forward into a reality of innovations in products and services which will help pave the way for the country's future. So a keen welcome to all participants, to be a part of this invaluable journey for yourselves and for our country.

Mohan Fernando
CEO, Informatics Institute of Technology



MESSAGE FROM THE DEAN – SCHOOL OF COMPUTING,

SCHOOL OF COMPUTING, INFORMATICS INSTITUTE OF TECHNOLOGY, SRI LANKA

Welcome to the 4th International Conference on Innovations in Info-business and Technology (ICIIT-2023) organized by the Informatics Institute of Technology (IIT). IIT takes pride in producing high-caliber Computing, Data Science and Business graduates who have played a pivotal role in empowering industries and organizations globally.

ICIIT 2023 in its 4th year demonstrates its maturity by featuring numerous academic presentations from both local and international universities in the fields of Computer science, AI & Data Science and Business Managment. The conference aims to highlight the work conducted by researchers worldwide, showcasing their findings and ongoing projects to academia and industry. By fostering innovation and advancement, ICIIT enables researchers from all over the world to present novel perspectives on diverse issues related to the conference theme.

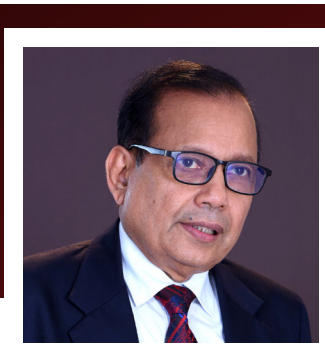
This year's theme, "Unlocking digital potential for economic growth in emerging regions," provides a platform for researchers to explore and harness the power and capabilities of digital technologies, aiming to stimulate and drive economic growth in emerging regions. By leveraging digital technologies, such as information technology, artificial intelligence, data analytics, cloud computing, Internet of Things (IoT), and other cutting-edge innovations, these regions can overcome the various challenges they face and unlock new opportunities for economic advancement.

A special welcome is extended to our keynote and invited speakers, and we warmly welcome all distinguished guests and those presenting papers and posters. We are confident that the diverse topics, expertise, and insights presented in these talks will contribute significantly to the conference theme and enrich the experience for all participants. Your valuable contributions are highly anticipated, and they will undoubtedly add significance and scope to the conference. We also extend our gratitude to our panel of judges, whose experience and knowledge bring immense value to our conference sessions.

We are genuinely grateful to the local and international reviewers and the conference organizing committee for their tireless efforts in ensuring the success of this conference. Their dedication and hard work have made this event possible.

Wishing you all a very fruitful and rewarding conference. Let us come together to explore innovative ideas and pave the way for a brighter future in overcoming a myriad of challenges and discovering novel avenues for economic progress.

Ms. Naomi Krishnarajah
Dean, School of Computing, Informatics Institute of Technology



MESSAGE FROM THE DEAN – FACULTY OF BUSINESS,

FACULTY OF BUSINESS, INFORMATICS INSTITUTE OF TECHNOLOGY, SRI LANKA

I am delighted to pen this message to proceedings of the 4th International Conference on Innovations in Info-business & Technology (ICIIT 2023) held on 04th August 2023 at Galadari Hotel, Colombo under the theme "Unlocking the digital potential for economic growth in emerging regions". First, I would like to appreciate the IIT management for the continuous support they extended with the hope of creating vibrant research culture within the institution covering both academic disciplines and teaching pedagogies. The conference series on info-business technology was initiated by the IIT four years ago within this broader vision creating a platform opened not only to its own staff and students but also to those of other higher learning institutions in the country and overseas. As the Dean of the IIT Business School, it is encouraging to notice that this conference series is receiving growing attention and recognition every year from academia and the industry.

I was informed by the conference committee headed by Professor Prasad Wimalaratne that this year's conference creates a platform to discuss and deliberate new ideas and research findings among academia, industry, technologists, and research scholars. Further, it features a series of academically rewarding activities such as plenary speeches, keynote addresses by leading international research scholars, tutorials and workshops by domain experts, research paper presentations, industry exhibits, and booths. I am confident that the research papers presented at the ICIIT 2023 under all three tracks will set the stage for thought-provoking debates and deliberations due to the rigorous reviewing process adopted by the conference committee. With this backdrop, I am sure that ICIIT 2013 will give you, the participants, some fruits to reimagine the future: your future, the future of Computing, Business, and Data Science. I wish to express my sincere appreciation to the Conference Chair, the members of the main conference committee, and all subcommittees for making this event a reality. I would also like to give my appreciation to the authors who have submitted their excellent works to this conference and to reviewers for their critical review of the submitted papers. Finally, I wish all the authors and attendees a highly stimulating, informative, and enjoyable conference.

Professor P. S. M. Gunaratne

Dean, School of Business Informatics Institute of Technology



MESSAGE FROM THE CONFERENCE CHAIR

ICIIT 2023

It is my pleasure to chair the International Conference on Innovations in Info-business & Technology (ICIIT 2023), which is held on the timely theme "Unlocking digital potential for economic growth in emerging regions". The goal of ICIIT is to bring together diverse groups to share knowledge with peers, explore innovative work, update with new trends & network with industry and academia.

This year's conference features with five key note speeches by the scholars from USA , Australia, UK and Germany. I am sure those keynotes will enlighten delegates with new knowledge of the latest developments. In addition, the programme features three tech talks by the industry experts and a panel discussion by experts. This year's conference has three tracks on AI and Data Science, Computing and Tech of Business with over 25 research paper presentations and over 15 poster presentations. Two post conference workshops on AI and Cyber security will be conducted by industry experts.

We are particularly indebted to keynote and invited speakers, our generous sponsors and event partners, management of IIT, organizing committee members, review panel and everyone who contributed to making the event a success.

I welcome delegates to the conference, and I am sure the participants will have a fruitful and exciting experience.

Prasad Wimalaratne PhD,
Conference Chair

ORGANIZING COMMITTEE

CONFERENCE CHAIR

General Chair

- Prof. Prasad Wimalaratne - PhD(Salford), BSc (Col), SMIEEE, MCSSL

STEERING COMMITTEE

- Ms. Naomi Krishnarajah- Dean, School of Computing, Informatics Institute of Technology
- Mr. Mohan Fernando - CEO, Informatics Institute of Technology
- Prof. P.S.M Gunaratne - Dean, School of Business, Informatics Institute of Technology

CONFERENCE COORDINATOR

- Dr Dinesh Arunathilaka

CONFERENCE SECRETARY

- Ms. Imasha Silva

TECHNICAL PROGRAM COMMITTEE

CHAIR

- Professor Nihal D. Kodikara

Professor of Computer Science - B.Sc. (Col), Dip in App. Stat. (Col), M.Sc. (Man.), Ph.D.(Man), MBCS, MCSSL, Informatics Institute of Technology

ORGANIZING COMMITTEE

CO-CHAIR

• Professor S Vasanthapriyan
Doctor of Engineering (Wuhan University of Technology), MSc (Eindhoven University of Technology, Netherlands), B.Sc(Hons) University of Peradeniya
Head of Department, Sabaragamuwa University

COMMITTEE MEMBER

• Dr. Ruvan Weerasinghe - PhD(Cardiff), MSc(Cardiff), Bsc (Col), Senior Lecturer, University of Colombo School of Computing

• Dr. Ajith Pasqual - B.Sc. Eng. (Moratuwa), M.Eng. (Tokyo), Ph.D. (Tokyo), MIEEE, MACM, Senior Lecturer, University of Moratuwa

• Dr. Manjula Sangirigama - PhD, CS (Ehime), Japan, Senior Lecturer, Faculty of Engineering, University of Peradeniya

• Prof Anuja Dharmarathne - PhD Hiroshima University(Japan), Computer Science (Honours) (Colombo) Associate Head (Education)School of IT, Monash University, Malaysia

• Dr. Chaman Wijesiriwardhana - PhD(Col), BSc(Peradeniya), Senior Lecturer, University of Moratuwa

• Prof. Damitha Karunaratna - BSc (Col), MSc (Swansea), PhD (Cardiff), MCS (SL), Professor, University of Colombo School of Computing

• Mr. Austen Mascranghe

PUBLICATION COMMITTEE

CHAIR

• Dr. Kathiravelu Thabotharan - Senior Lecturer,Department of Computer Science, Faculty of Science, University of Jaffna, PhD in CS (Colombo), Ph.L. in CS (Uppsala, Sweden), M.Sc. in CS (Akron, US) B.Sc. Special in CS (Jaffna)

CO-CHAIR

• Prof. Nimal Wickramasinghe - PhD (Kansas State University), Former Head, Department of Statistics University of Colombo

• Dr. Rohan Samarasinghe - PhD,(Iizuka, Japan) 2003, BSc (Kelaniya), Head of ICT Department, Faculty of Technology, University of Colombo

COMMITTEE MEMBER

• Dr. Frederic Kleinermann - PhD, Research Scientist in Computer Science & Applied Mathematics, Vrije Universiteit, Brussels Belgium

• Prof. D.M.G. Preethichandra - (BSc.Eng., M.Eng., MEBM, PhD, C.Eng., SMIEEE), Deputy Dean (Research) , School of Engineering and Technology, Central Queensland University, Australia

• Dr. Dharshana Kasthurirathna -Assistant Professor, SLIIT, PhD iSydney), BSc in CSE, University of Moratuwa

• Dr. Chitraka Wickramarachchi - PhD, University of Canterbury, Senior Lecturer, University of Sri Jayewardenepura

• Dr. Chamli Pushpakumara - PhD (Sheffield Hallam), Dean, Faculty of Computing and Technology, University of Kelaniya

TRACK CHAIRS

DATA SCIENCE & AI TRACK

CHAIR
Mr. Nipuna Senanayake

CO-CHAIR
Ms. Janani Harishchandra

COMPUTING

CHAIR
Mr. Iresh Bandara

CO-CHAIR
Ms. Lakna Gammedda

TECHNOLOGY OF BUSINESS

CHAIR
Ms. Abarana Kirupananda

CO-CHAIR
Ms. Gayashini Rathnayake

WORKSHOP & TUTORIALS COMMITTEE

CHAIR
Mr. Guhanathan Poravi

CO-CHAIR
Mr. Torin Wirasinghe

FINANCE COMMITTEE

CHAIR

- Mr. Sudarshan Welihinda

CO-CHAIR

- Ms. Theja Perera
- Ms. Ganesha Thondilege

MEMBER

- Mr. Leon Charles

MARKETING, PUBLIC & INTERNATIONAL LIAISON COMMITTEE

- Mr. Ashwaq Ahamed
- Ms. Valarine De Zilva
- Mr. Pradeep Kariyawasam
- Mr. Chamith Maduranga
- Mr. Kushantha Toussaint

REGISTRATION & CERTIFICATION COMMITTEE

CHAIR

- Dr. Jayantha Rajapakse

CO-CHAIR

- Mr. Cassim Farook
- Ms. Janice Abeykoon

LOGISTICS COMMITTEE

- Mr. Prashantha Rodrigo
- Mr. Sarath Ananda
- Ms. Wayomi Mahanama
- Mr. Thanuja Piyaratne

TECHNICAL PROGRAM COMMITTEE

CHAIR

- Professor Nihal D. Kodikara

CO-CHAIR

- Professor S Vasanthapriyan

COMMITTEE MEMBER

- Dr. Ruvan Weerasinghe
- Dr. Ajith Pasqual
- Dr. Manjula Sangirigama
- Prof Anuja Dharmarathne
- Dr. Thilak Chaminda Hapugahage
- Dr. Chaman Wijesiriwardhana

REVIEW COMMITTEE

Prof. Tharusha N Gooneratne, University of Colombo

Prof. Sampath Kehelwalatenna, University of Colombo

Prof. Pavithra Kailasapathy, University of Colombo

Prof. P.D. Nimal, University of Sri Jayewardenepura

Prof. N. N. J Navaratne, University of Colombo

Prof. M.P.P. Dharmadasa, University of Kelaniya

Prof. M.J.M. Razi, University of Colombo

Prof. Lalith Edirisinghe, CINEC

Prof. Klaus Schmidt, Illinois State University

Prof. K.A.S.P Kaluarachchi, University of Colombo

Prof. D.M Semasinghe, University of Kelaniya

Prof. Chalinda Beneragama, University of Peradeniya

Prof D.D. Karunaratna, University of Colombo

Prof. B. Nishantha, University of Colombo

Prof. Anuja Dharmaratne, Monash University Malaysia

Dr. Windhya Rankothge, IEEE Sri Lanka Section

Dr. Thilina Halloluwa, University of Colombo

Dr. Thilak Chaminda, Lanka Nippon BizTech Institute

Dr. Thareendhra Wijayasiriwardhane, University of Kelaniya

Dr. Thabotharan Kathiravelu, University of Jaffna

Dr. Suzanne Embury, University of Manchester

Dr. Sudath Heiyanthuduwage, Charles Sturt University. Sydney

Dr. Sidath Liyanage, University of Kelaniya

Dr. Sameera Viswakula, University of Colombo

Dr. Sachintha Pitigala, University of Kelaniya

Dr. S.R Manorathne, University of Colombo

Dr. Ruwan Navarathna, University of Peradeniya

Dr. Rohan Samarasinghe, University of Colombo

Dr. Rashmi Bomiriya, Metrics Asia Holdings

Dr. Pushpike Thilakarathne, Janssen Pharmaceuticals, Beers, Belgium

Dr. Prabhani Kuruppumullage, Eberly College of Science, PennState

Dr. Nishani Ekanayake, University of Colombo

Dr. Nihal Hennayake, University of Colombo

Dr. Lakshika Navarathna, University of Peradeniya

Dr. Jianhua Shao, Cardiff University

Dr. Isuru Hewapathirana, University of Kelaniya

Dr. Indika W.A., University of Ruhuna

Dr. Hakim Usoof, University of Peradeniya

Dr. H.A.K.N.S. Surangi, University of Kelaniya

Dr. Dharshana Kasthurirathna, Institute of Information Technology

Dr. Chinthaka Kuruwita, Hamilton College, Clinton

Dr. C.N. Wickramasinghe, University of Kelaniya

Ms. Mandu Jayasundera, Tik Tok USA

Ms. Deshanee Wickramarachchi, National University of Singapore

MS Ganesha Thondilege, Informatics Institute of Technology

Mr Deshan Sumanathilaka, Informatics Institute of Technology

Mr Prasanna Haddela, Sri Lanka Institute of Information Technology

Dr. B. M. Thosini, University of Kelaniya

Dr. Rohan Samarasinghe, University of Colombo

Mr Cassim Farook, Informatics Institute of Technology

Dr Dharshana Kasthurirathna, Institute of Information Technology

Dr. Sameera Viswakula, University of Colombo

Dr Lakshika Nawarathna, University of Peradeniya

Mr Torin Wirasingha, Informatics Institute of Technology

MS Nideshika Ellepola, SLIIT Academy

Mr Pumudu Fernando, Informatics Institute of Technology

KEYNOTE SPEAKER

PROF. UMASHANGER THAYASIVAM

PROFESSOR AT ROWAN UNIVERSITY, USA



Prof. Umashanger obtained his BSc(Hons) degree in Statistics from University of Colombo, Sri Lanka in 2001 and proceeded to University of Georgia in USA to obtain his MS in Statistics in 2005 and the PhD from the same University in 2009.

At present, he is a Professor in Statistics and Data Science at Department of Mathematics, in Rowan University, New Jersey, USA.

His Research Interests include; Statistical Data Mining, Supervised Learning Approaches for Biomarker Discovery, Robust Model Selection and Estimation, Spoof detection & Speaker Identification, Mixture Distribution Parameter/Component Estimation with L2E, Significance of Renyi Divergence Professor Uma Shankar is a member of American Statistical Organization, Institute of Mathematical Statistics and Institute of Applied Statistics, Sri Lanka. He has contributed to many significant projects and is the author of many research related publications.

TITLE: FUTURE OF DATA SCIENCE- OPPORTUNITIES AND CHALLENGES OF LARGE LANGUAGE MODELS (LLM)

Prof. Umashanger Thayasivam
Professor at Rowan University, USA
thayasivam@rowan.edu

Abstract:

As we enter a new era of technological advancements, the field of data science is undergoing a significant transformation. In particular, the emergence of large language models revolutionized the way we process and understand vast amounts of textual data. This talk aims to explore the opportunities and challenges presented by these powerful models and their impact on the future of data science. The opportunities presented by large language models demonstrated remarkable proficiency in tasks such as machine translation, sentiment analysis, content generation, and question answering.

Industries ranging from healthcare to finance or marketing to education all stand to benefit from the potential applications of large language models. Along with these immense opportunities, large language models also bring forth a set of challenges that need to be addressed. Ethical considerations surrounding bias, fairness, and privacy in data collection, model training, and deployment require careful attention. The potential for the amplification of misinformation and fake news also demands proactive measures. The future of data science lies at the intersection of human intelligence and machine learning, and large language models represent a significant milestone in this journey. We can shape a future where data-driven decision- making is more accurate, efficient, and beneficial to society.

MAIN SPEAKER

PROF ANURA JAYASUMANA

PROFESSOR AT COLORADO STATE UNIVERSITY, USA



Prof. Jayasumana obtained his BSc (Engineering) from University of Moratuwa, Sri Lanka before proceeding to USA for further studies.

He is currently a Professor in Electrical and Computer Engineering Department (with joint appointments in Computer Science Department, Systems Engineering and Biomedical Engineering). He is also the Associate Director, Information Sciences and Technology Center (ISTeC) at Colorado State University. He founded the Computer Networking Research Laboratory at Colorado State University, and is a member of NSF Engineering Research Center for Collaborative Adaptive Sensing of the Atmosphere.

His research interests include; Computer & Communication Networking, Internet of Things, Social Networks, Sensor/ Optical/ Wireless/ Wired Networks, Performance Modeling, Network Protocols, VLSI: Testing and Testable Design of Integrated Circuits.

He has published over 300 research papers and is a consultant to many companies.

TITLE: MACHINE LEARNING AND SYNTHETIC DATA – POTENTIAL AND PITFALLS

Prof Anura Jayasumana
Professor at Colorado State University, USA

Abstract:

Machine Learning (ML) models have become indispensable for solving complex problems. However, they require sufficient volumes of representative training data to be effective. There are many problem domains where the collection of sufficient data is expensive or even infeasible. Unconventional Data Sets (UDS) that are inherently sparse and incomplete are quite common in social and behavioral domains. Within a UDS, it is possible for various data fields to be absent from different records, owing to multiple factors such as lack of measurement, non-measurability, lack of knowledge of its existence, or simply inapplicability to the specific record. Furthermore, collection challenges, cost, lack of participation, small affected populations, etc., also result in very small and sparse sets of data. Synthetic data provides a potential solution to overcome such limitations for ML.

Our recent research resulted in techniques to model and synthesize large sets of complex data objects from UDS while maintaining the same statistical and topological characteristics. Synthetic data sets we generated have shown marked improvements in the performance of ML algorithms in problems such as detection of radicalization potential, video traffic classification, and phishing website detection. We will also overview the state of the art in the use of synthetic data for training ML models as well as ML techniques for generating such data.

The resulting symbiotic relationship, between synthetic data generated by ML models and ML models trained using synthetic data, promises a bold new field with creative solutions and perilous pitfalls.

GUEST SPEAKER

DR BARBARA VILLARINI

UNIVERSITY OF WESTMINSTER, UNITED KINGDOM



Dr. Villarini graduated in Information Engineering at the University of Perugia, Italy, in 2005. She received her Master's degree in informatics and Telecommunications Engineering from the University of Perugia in 2008, and her Doctorate degree in Information Engineering also from the University of Perugia in 2012. Her doctoral work was focused on multimedia and image processing, 2D/3D image and video quality assessment and error protection.

She is a Lecturer in Software Engineering in the Department of Computer Science and an active researcher in medical imaging processing and 3D data registration and reconstruction at University of Westminster, United Kingdom.

Her current research interests include; 3D Registration and Reconstruction, Image and Video Quality Assessment, Medical Image Processing

She has many publications to her credit in these research areas.

TITLE: ARTIFICIAL INTELLIGENCE FOR AUTOMATIC ORGAN RECONSTRUCTION AND MORPHOLOGICAL FEATURES EXTRACTION FROM MEDICAL IMAGES.

Dr Barbara Villarini
University of Westminster, United Kingdom

Abstract:

The Impact on Healthcare. Medical imaging technology has rapidly evolved in the last decades providing detailed anatomical images and enabling non-invasive techniques. Abdominal radiological scans, such as Magnetic Resonance Imaging (MRI) and Computer Tomography (CT), are extensively performed, holding potential for assessing medical conditions and predicting risks. However, manually inspecting large- scale datasets poses challenges for radiologists, and expert-led segmentation as it is time-consuming and subject to variability. The talk will explore the role of Artificial Intelligence (AI) in computer- aided diagnosis (CADx) systems, specifically focusing on an AI-driven open-source framework for automatic organ segmentation and reconstruction from medical images. The segmentation approach extends to the classification and extraction of prognostic biomarkers for medical conditions. Developing robust automated segmentation techniques faces challenges due to variations in organ structure, size, spatial resolutions, and image quality across different modalities. Convolutional neural networks (CNNs) have significantly enhanced imaging tasks, including abdominal organ segmentation, particularly for deformable organs with indistinct boundaries. Furthermore, integrating automatic organ segmentation into digital twins enables healthcare professionals to gain accurate and detailed information, facilitating precise visualization, analysis, and simulation for personalized interventions, surgical planning, and treatment evaluation. This symbiotic relationship between digital twins and automatic organ segmentation revolutionizes healthcare practices, ultimately enhancing patient care.

GUEST SPEAKER

PROF. DR. WOLFGANG ZIEGLER

PROFESSOR AT HOCHSCHULE KARLSRUHE UNIVERSITY

OF APPLIED SCIENCES, GERMANY



Prof. Ziegler obtained his PhD in theoretical physics from University of Wurzberg in Germany. He is currently a Professor at Hochschule Karlsruhe University of Applied Sciences, Germany, In the faculty of Information Management and Media.

His Research interests are in Global information management (focus on Asia), Configuration specific content management, Content Delivery and Digital Information Services and Semantic models and knowledge management.

He has many publications to his credit in English as well as in German language.

His other duties are including a Consultant for information and content management in technical communication, Professor for Content Development & Publishing, Karlsruhe University of Applied Sciences and Conference Advisory Board tekom Beirat ETLTC (ACM Conference, Univ. of Aizu, Japan).

TITLE: DIGITAL INFORMATION SERVICES IN TECHNICAL COMMUNICATION BASED ON CONTENT MANAGEMENT, DELIVERY AND SEMANTICS

Prof. Dr. Wolfgang Ziegler
Professor at Hochschule Karlsruhe University of Applied Sciences, Germany

Abstract:

Information Management in the domain of Technical Communication has gained recent interest due to its contributions to digitalization scenarios and developments of digital services in B2B and B2C industries. We will show recent market data in central Europe and focus methodologies used in underlying Component Content Management and Delivery Systems. Semantic technologies like knowledge graphs are used increasingly in this domain for enhancing user support and (meta)data quality. Of course, also in this area, the use of AI technologies are heavily discussed and (will) have an interplay with the above introduced technologies. We will also report on some recent industrial projects using graph-based content correlations in search interfaces.

GUEST SPEAKER

PROF. DAMMINDA ALAHAKOON

PROFESSOR AT LA TROBE UNIVERSITY, AUSTRALIA



Prof. Alahakoon obtained his BSc (Honours) from University of Colombo, Sri Lanka and his PhD from Monash University, Australia.

He is a Professor and Discipline Leader in Business Analytics, LaTrobe University Business School (since 2015). He is also the Founding Director of the Research Centre for Data Analytics and Cognition (CDAC) at La Trobe.

His research is focused on the areas of Machine Learning, Artificial Intelligence, Advanced Data Analytics, Text and Social Media Analytics, Business Intelligence, and the harnessing of such theories for practical tools and innovative technology for the industry.

His work on unsupervised self-structuring AI and the Growing Self Organizing Map (GSOM) algorithm has received international recognition and adopted in diverse fields by industry, government, and academia. He has advanced and extended the GSOM as a new paradigm in artificial learning with theoretical innovations and practical outcomes published in over 150 peer-reviewed research articles.

TITLE: ARTIFICIAL GENERAL INTELLIGENCE (AGI) : IS IT TIME FOR MACHINES THAT LEARN AND THINK FOR THEMSELVES

Prof. Damminda Alahakoon

Professor at La Trobe University, Australia

Abstract:

Although the term artificial intelligence (AI) was coined by John McCarthy in 1956 at the famous Dartmouth workshop, the idea that human thought processes could be mechanized has existed for many centuries. Early Greek, Indian and Chinese philosophers contemplated how human thinking could be artificially mechanized and manipulated by intelligent non-human machines. These ideas focussed mostly on humanoid robots with mechanical versions of human thought processes. The advent of computing in the mid 1900s revolutionised such thought with computers becoming the 'machines' which enabled and 'housed' the AI. Norbert Wiener's cybernetics, Claude Shannon's information theory and Alan Turing's theory of computation led the way towards this computing based AI which went through several phases of varying popularity in the next half century. Symbolic AI was the dominant paradigm at the early stages which was also called Good Old Fashioned AI (GOFAI).A more practical variation of AI called machine learning became widely used since the latter part of the 20th century where computer programs learnt how to carry out tasks from data, instead of having to be specifically programmed by humans which were known as sub-symbolic AI. Sub-symbolic AI, especially artificial neural networks and more recently deep learning have achieved huge successes but a major limitation is their black box nature which restricts explainability.

The recent years has seen a major shift in not only the thinking and application of AI, but the role of AI in society and the relationship with humans. The wide acceptance and popularity of the internet and technologies such as cloud computing and mobile devices has resulted in a connected global community where many tasks and processes are automated. The social media phenomenon within this environment enabled people to create, disseminate and share information as well as socialise and converse, thus bringing in the human element into a fast-growing profusion of digital data. The outcome of all this has been heralded as the information revolution, creating a virtual digital world and the emergence of a digital ecosystem; where human interactions via social media, internet of things (IOT) represent objects, smart cities and homes create the connected environment and big data fabricates the digital space.

The new digital environment and ecosystem has completely transformed the traditional utilization of AI and instead created a role for AI co-existing with humans. In the past, digital information representing natural environments, people and behaviours had to be generated and fed into AI programs which would then output results – as such AI functioned mostly as smarter tools used by humans. The AI learned what humans required them to learn and were mostly built upon a supervised learning paradigm. The new digital environment and ecosystem in contrast generates a digitized version of most of the natural world. Social media and other communication mediums capture traces of our interests, opinions and emotions which smart programs could develop into complete profiles or avatars by the fusion of multiple data sources over time. This environment results in big data – massive volumes of digital content from a variety of sources (sensors, IOT, text, social media, images, video etc) which are also highly volatile. As such the traditional practise of training AI by humans for specific activities using well defined data has become obsolete, and in many cases impossible. The new breed of AI for such environments need to be much more autonomous and capable of self-building and have the ability to self-learn using the unsupervised learning paradigm.

To cater to such environments, and to be able to harness the value of big data and power of technological advancement, there is a call for AI with more autonomy, ability to interact, communicate, socialise with humans, act and re-act quickly and efficiently to new situations without requiring exact pre-training. Such AI has been called Artificial General Intelligence (AGI) or Strong AI as opposed to Narrow AI, which have to be mostly pre-trained for clearly defined tasks. There are several and varying definitions of AGI which range from more advanced deep learning while others argue that deep learning can never achieve AGI. The recent advances in NLP technologies including ChatGPT has made advances in AI a talking point and a subject of debate.

This talk will discuss the evolution of AI to it's current position and the possibilities of AGI, machines that can learn and think for themsleves. Several key ideas, theories and concepts such as latent learning, modality independent data fusion, theory of mind will be presented as potential components of achieving AGI.

WORKSHOPS AND TUTORIALS

APPLICATIONS AND USE CASES OF AI & MACHINE LEARNING FOR ENTERPRISE BUSINESSES

Presented by Analytics & AI Team at Axiata Digital Labs



SPEAKERS:

- Buddhika S. Gamage, Head of Competency Excellence - Data Engineering, Senior Architect at ADL, Bachelor in Electronics & Communication Engineering, MANIT, Master in Business Administration, Master in Data Science

- Dr. Lahiru Ariyasinghe, Associate Architect - Data Science at ADL, PhD in Computer Science, BSc (Hons) in IT with First Class

- Isuru Alagiyawanna, Senior Data Scientist at ADL & Assistant Lecturer at SLIIT, BSc in Electronic and Telecommunication Engineering

- Himasha Jayawardene, Senior Data Engineer at ADL, BSc in Electronic and Telecommunication Engineering

- Sajani Mayadunna, Data Engineer at ADL, BSc (Hons) in Information Technology Specializing in Data Science

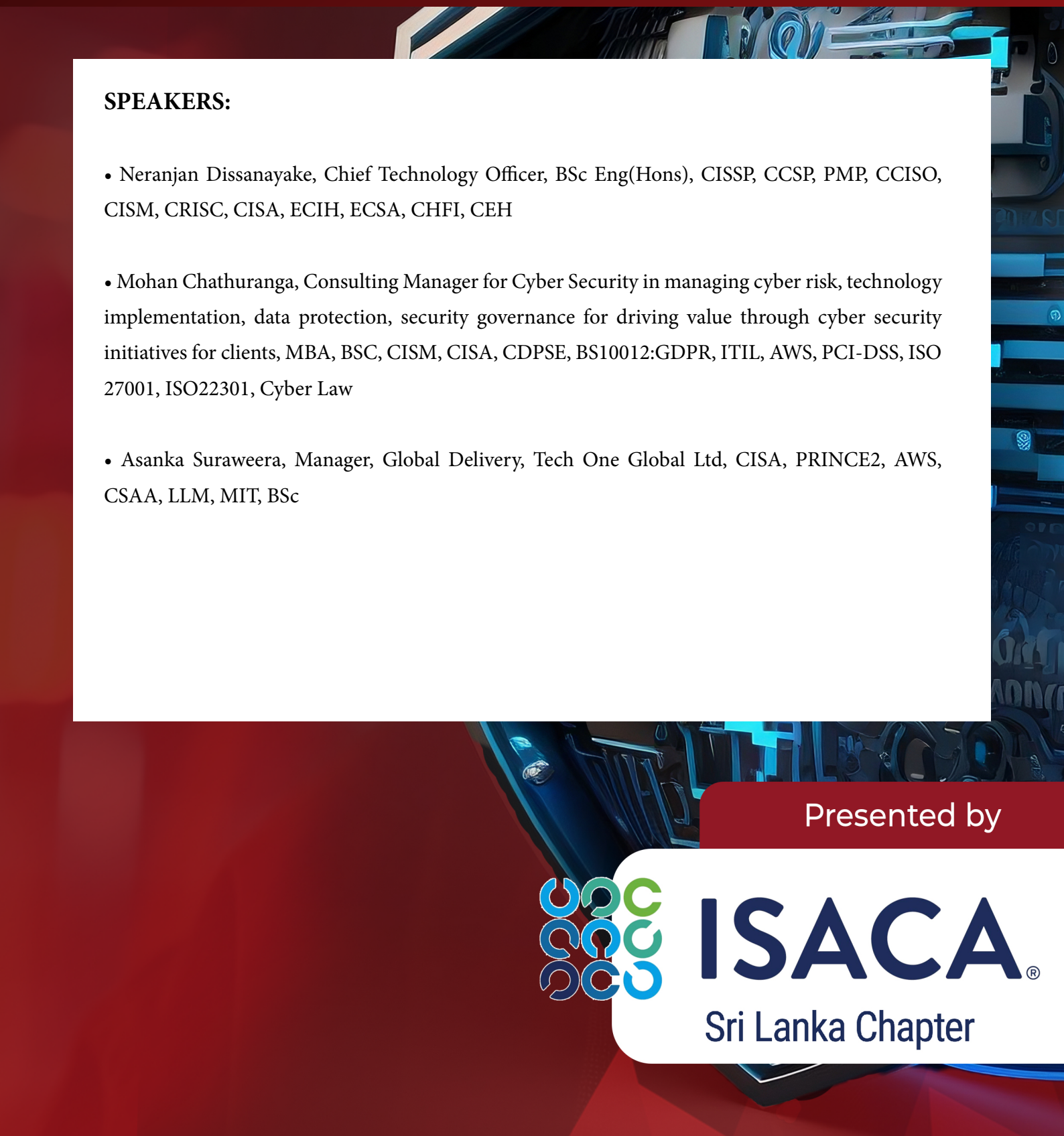
Presented by

AXIATA DIGITAL LABS

WORKSHOPS AND TUTORIALS

DEMYSTIFYING THE DIGITAL TRUST, CONSUMERS' ROLE IN DIGITAL TRUST, LEGAL SAFEGUARDS AGAINST THE VIOLATIONS OF DATA PROTECTION & PRIVACY : ENSURING DIGITAL TRUST: SAFEGUARDING PRIVACY AND ENHANCING DATA PROTECTION

Presented by ISACA (Sri Lanka Chapter)



SPEAKERS:

- Neranjan Dissanayake, Chief Technology Officer, BSc Eng(Hons), CISSP, CCSP, PMP, CCISO, CISM, CRISC, CISA, ECIH, ECSA, CHFI, CEH

- Mohan Chathuranga, Consulting Manager for Cyber Security in managing cyber risk, technology implementation, data protection, security governance for driving value through cyber security initiatives for clients, MBA, BSC, CISM, CISA, CDPSE, BS10012:GDPR, ITIL, AWS, PCI-DSS, ISO 27001, ISO22301, Cyber Law

- Asanka Suraweera, Manager, Global Delivery, Tech One Global Ltd, CISA, PRINCE2, AWS, CSAA, LLM, MIT, BSc

Presented by

ISACA®

Sri Lanka Chapter

iCIIT 2023

THE 4TH INTERNATIONAL CONFERENCE ON INNOVATIONS IN INFO-BUSINESS & TECHNOLOGY

AGENDA



TIME	DESCRIPTION		
08.00 - 09.00	Registration of Participants		
09.00 - 09.05	Inauguration		
09.00 - 09.10	The lighting of the Oil lamp		
09.10 - 09.13	National Anthem		
09.13 - 09.15	Opening of the conference- Opening Video		
09.15 - 09.25	Welcome address by Chief Executive Officer of IIT, Mr. Mohan Fernando		
09.25 - 09.27	RE-CAP ICIIT video		
09.27 - 09.33	Address by the ICIIT Conference Chair - Prof. Prasad Wimalaratne		
09.33 - 09.38	Introduction of Sponsors / Event partners		
09.38 - 09.40	Launching of the eProceedings by the Chairman of IIT		
09.40 - 10.15	Keynote Speech by Prof. Umashangar Thayasivam, Rowan University, USA		
10.15 - 10.35	Tea Break		
10.35 - 11.10	Invited Speech by Prof. Anura Jayasumana, Colorado State University, USA "Machine Learning and Synthetic Data – Potential and Pitfalls"		
11.10 - 11.25	**Tech Talk 1** by SLTMobitel, Platinum Sponsor: Anuradha Udunuwara, Senior Engineer, Sri Lanka Telecom **"Making Sri Lanka an Innovation Hub"**		
Hall Name	**Track 1 - Main Hall**	**Track 2 - Orchid Room**	**Track 3 - Jasmine Room**
Conference Tracks	**AI & Data Science**	**Computing**	**Technology of Business**
11.35 – 12.10	Invited Speech by Prof. Damminda Alahakoon, La Trobe University, Australia *"Is it time for Artificial General Intelligence (AGI)?: Machines that Learn and Think for Themselves*	Invited Speech by Dr. Barabara Villarni, University of Westminster, United Kingdom *"Artificial Intelligence for Automatic Organ Reconstruction and Morphological Features Extraction from Medical Images. The Impact on Healthcare"*	Invited Speech by Prof. Wolfgang Ziegler, Hochschule Karlsruhe University of Applied Sciences, Germany *"Digital Information Services in Technical Communication based on Content Management, Delivery and Semantics"*
Technical Sessions	**Technical Session 1** Session Chair - Mr Niranjan Silva, Former Director/ CEO of Metropolitan Computers	**Technical Session 1** Session Chair - Prof Nihal Kodikara, Professor IIT	**Technical Session 1** Session Chair - Prof Nalaka Wickramasinghe, Associate Dean, Business School, IIT

12.10 - 12.30	**Exploring Emotion Classification in a Labeled Twitter Dataset: A Comparative Study of Machine Learning Approaches** Sugeeshwa S P Galhena, Ajantha S Atukorale *(University of Colombo School of Computing, University of Colombo, Sri Lanka)*	**Developing An Automated System to Analyzing the Software Developers Interest Using Twitter Data** [1]K. N. M. S. Nandasiri, [1]K. G. L. S. Ananda, 1D. V. S. Maduranga, [1]K. D. H. S. Bandara, [1]C. P. Wijesiriwardena, [2] I. T. S. Piyatilake ([1]Department of Information Technology, University of Moratuwa, Sri Lanka, [2]Department of Computational Mathematics, University of Moratuwa, Sri Lanka)	**AlzhCare: A cognitive therapy application for Alzheimer's Patients with smart results, and a safety tool to support their Caregivers** Jesica Rayen, Namal Malalasena *(Department of Business School Informatics Institute of Technology, Sri Lanka)*
12.30 - 12.50	**Health Care Data Privacy Preserving using Federated Learning: A case of Heart Disease Prediction** A.S Aathika, L. Rupasinghe *(Computer Systems & Engineering, Sri Lanka Institute of Information Technology, Sri Lanka)*	**Predicting the Possibility of Chronic Pancreatitis** Bawantha Wickramarachchi, Iresh Bandara (Department of Computing, Informatics Institution of Technology, Sri Lanka)	**The Impact of Electronic Word-of-Mouth on Customer Purchase Intention** Vishaleni Sureshan, Shonali Aponso *(School of Business, Informatics Institute of Technology, Sri Lanka)*
12.50 - 13.50	Lunch Break		
13.50 - 17.00	Poster Presentations Session in Main Hall		
Technical Sessions	**Technical Session 2** Session Chair - Dr. Ruvan Weerasinghe, Head of Data Science, IIT	**Technical Session 2** Session Chair Ms. Naomi Krishnarajah, Dean Computing School, IIT	**Technical Session 2** Session Chair - Mr. Austen Mascranghe, Head of Department, Business School, IIT
14.20 - 14.40	**Tech Talk 2: Muti server Scalable Systems**. RoshanSembacuttiaratchy, Ex-CEO, AppMeta Solutions (Pvt) L td	**Optimization of Climatic Parameters of the Greenhouse to Maximize Chilli Production** C.S. Silva, A.P. Wijayawantha, M.T.U.D. Mallikarathne, A.M.H.M. Abeysinghe *(University of Colombo)*	**Remaddict: A Solution Put Forward To Predict Withdrawal Symptoms And Act As A Maintenance Tool To Assist Adolescents Challenged With Substance Use Disorder** Avisha de Mel, Janice Abeykoon *(School of Business, Informatics*
14.40 - 15.00	**Detecting Dyslexia and Dysgraphia Risks in Sinhala-Speaking Children Using Neural Networks** [1] Budhvin T. Withana, [2]Sulochana Rupasinghe *([1] Computer Science and Engineering, University of Westminster, United Kingdom, [2] Department of Computing, Informatics Institute of Technology, Sri Lanka)*	**Machine Learning Approach to Predict Next Transaction Day with Customer Lifetime Value** Ms. Punsisi Kaludewa, Ms. Rukshala Weerasingha *(Department of Computing, Informatics Institute of Technology, Sri Lanka)*	**The Influence of Online Shopping Experience and Attitudinal Loyalty on Online Impulsive Buying in Sri Lankan Online Fashion Retail Industry** Chethana Gallage, Gayashini Ratnayake *(School of Business, Informatics Institute of Technology, Sri Lanka)*

15.00 - 15.20	**A Comparative Analysis of Extensible Approaches for Document Layout Segmentation** Chanaka M. Wickramasinghe, Anushka D. Vithanage, Dharshana Kasthurirathna, Ajantha S. Atukorale *(Department of Computation and Intelligent Systems, University of Colombo School of Computing, Sri Lanka)*	**Empowering Early Childhood Education Students: A User-Centered Approach to Enhance Spelling Skills** Shidn De Silva, MTA Wickramasinghe, D Gunasekara *(Department of Information Technology, Faculty of Computing, General Sir John Kotelawala Defence University, Sri Lanka)*	**Impact of Social Media Communication and Brand Equity on Consumer Perceptions of Brands in the Fast-Food Industry** Ashweka Vasanthakumar, Gayashini Ratnayake *(School of Business, Informatics Institute of Technology, Sri Lanka)*
15.20 - 15.40	**Styllizea: Name Recommending System for Fashion Products Using Image Captioning on Single-Object Images and Multiple-Object Images** Mahidi Wijayantha, Niwarthana Kariyabaduge *(School of Computing, Informatics Institute of Technology, Sri Lanka)*	**Developing a System to Transliterate Singlish Twitter Posts to Sinhala** [1] M. D. C. Amarasekara, [1] R. A. D. P. Rajapaksha, 1 H. M. D. T. Jayarathna, 1 H. M. G. K. Karunarathna, [1] C. P. Wijesiriwardana, [2] I. T. S. Piyatilake *(1Department of Information Technology, University of Moratuwa, Sri Lanka, [2] Department of Computational Mathematics, University of Moratuwa, Sri Lanka)*	**Influence of social media marketing using Instagram and consumer trust on consumer purchase intentions in the skincare industry in Colombo District in Sri Lanka** Hansini Navodya Somaratna, Gayashini Ratnayake *(School of Business, Informatics Institute of Technology, Sri Lanka)*
15.40 - 16.00	Tea Break		
Technical Sessions	**Technical Session 3** Session Chair - Nipuna Senanayake, Senior Lecturer, IIT	**Technical Session 3** Session Chair - Pumudu Fernando, Senior Lecturer, IIT	**Technical Session 3** Session Chair - Ms Kumudini Sarathchandra, Senior Lecturer, IIT
16.00 - 16.20	**Deepfake Low Resource Image Detection with Explainable Reporting** [1] Abdul Baasith Shiyam, [2] Guhanathan Poravi *([1] Computer Science and Engineering, University of Westminster, United Kingdom, [2] Department of Computing, Informatics Institute of Technology, Sri Lanka)*	**Helaa: A Sinhala Language-Based Programming Language** Rajitha Yasasri, Damitha Karunarathna *(University of Colombo School of Computing, University of Colombo, Sri Lanka)*	**Mama Aid - An application to identify postpartum complications and assist in re-establishing mother's health** Erandi Kavindya Liyanagamage, Abarnah Kirupananda *(School of Business, Informatics Institute of Technology, Sri Lanka)*
16.20 - 16.40	**Sinhala Active Voice into Passive Voice Converter using Rule Based Approach with Grammar Error Correction** Oshini Ilukkumbura, Sulochana Rupasinghe *(School of Computing, Informatics Institute of Technology, Sri Lanka)*	**Alternative Transport Layer Protocol** [1] Harshi P. S. Ratnayake, [2] Geethapriya Liyanage *([1] Informatics Institute of Technology affiliated to University of Westminster, Sri Lanka, [2] Department of Computer Science and Engineering, University of Moratuwa, Moratuwa)*	**Influence of Internal Employer Branding and Organizational Trust on Employee Engagement In The IT Industry in Sri Lanka** Sashini Vasanthan, Gayashini Ratnayake *(School of Business, Informatics Institute of Technology, Sri Lanka)*

16.40 - 17.00	**HealthySmyl: Analysis and Recommendations Based Lifestyle Enhancing Application for HIV Patients** Selshia Sbanumi Regis Sagayaratnam, Kumudini Sarathchandra *(Business School, Informatics Institute of Technology, Sri Lanka)*		**Ambula - A Mobile Application Solution to Promote Sri Lankan Food Culture** [1] Sarah Hausotter, [2] Lakshika Paiva *([1] School of Business, Informatics Institute of Technology, Sri Lanka, [2] WSO2 Lanka Pvt. Ltd., Colombo 04, Sri Lanka)*
17.00 - 17.20	**Tech Talk 3: by Zone24x7 on AI-Powered Analytics: Transforming Real-World Challenges into Solutions in the Physical Realm.** Manjulal Waas, Head of Client Services, Zone24x7 Pvt Ltd		**Influence of Supervisor Behavior & Social Support on Employee Wellbeing in the Apparel Industry of Sri Lanka** Osandi Udalamatta, Gayashini Ratnayake *(School of Business, Informatics Institute of Technology, Sri Lanka)*
17.25 - 17.55	**Panel Discussion : "Learning and Assessment in the age of Generative AI"** Panelists :, Dr Ruvan Weerasinghe, Prof. Umashangar Thayasiwam, Prof. Nalaka Wickramasinghe, Mr Jeevan Gananam, CEO St Anthony's Group. Ms. Naomi Krishnarajah; Moderator- Mr. Havindra Gunawardene		
17.55 – 18:00	Raffle Draw		
18:00 – 18:30	Awards		
18:30 – 18:35	Vote of Thanks, Dr. Dinesh Arunathilaka, Conference Coordinator		

iCIIT 2023

THE 4TH INTERNATIONAL CONFERENCE ON INNOVATIONS IN INFO-BUSINESS & TECHNOLOGY

PAPER PRESENTATIONS

Online proceedings available at www.iciit.iit.ac.lk



TABLE OF CONTENT

DATA SCIENCE & AI

Paper ID	Title and Author Page	Page
27	**Exploring Emotion Classification in a Labeled Twitter Dataset: A Comparative Study of Machine Learning Approaches** Sugeeshwa S P Galhena, Ajantha S Atukorale *(University of Colombo School of Computing, University of Colombo, Sri Lanka)*	1
73	**Health Care Data Privacy Preserving using Federated Learning: A case of Heart Disease Prediction** A.S Aathika, L. Rupasinghe *(Computer Systems & Engineering, Sri Lanka Institute of Information Technology, Sri Lanka)*	15
14	**Detecting Dyslexia and Dysgraphia Risks in Sinhala-Speaking Children Using Neural Networks** [1] Budhvin T. Withana, [2] Sulochana Rupasinghe *([1] Computer Science and Engineering, University of Westminster, United Kingdom, [2] Department of Computing, Informatics Institute of Technology, Sri Lanka)*	26
28	**A Comparative Analysis of Extensible Approaches for Document Layout Segmentation** Chanaka M. Wickramasinghe, Anushka D. Vithanage, Dharshana Kasthurirathna, Ajantha S. Atukorale *(Department of Computation and Intelligent Systems, University of Colombo School of Computing, Sri Lanka)*	37
56	**Styllizea: Name Recommending System for Fashion Products Using Image Captioning on Single-Object Images and Multiple-Object Images** Mahidi Wijayantha, Niwarthana Kariyabaduge *(School of Computing, Informatics Institute of Technology, Sri Lanka,*	49
03	**Deepfake Low Resource Image Detection with Explainable Reporting** [1] Abdul Baasith Shiyam, [2] Guhanathan Poravi *([1] Computer Science and Engineering, University of Westminster, United Kingdom, [2] Department of Computing, Informatics Institute of Technology, Sri Lanka)*	60
15	**Sinhala Active Voice into Passive Voice Converter using Rule Based Approach with Grammar Error Correction** Oshini Ilukkumbura, Sulochana Rupasinghe *(School of Computing, Informatics Institute of Technology, Sri Lanka)*	70
39	**HealthySmyl: Analysis and Recommendations Based Lifestyle Enhancing Application for HIV Patients** Selshia Sbanumi Regis Sagayaratnam, Kumudini Sarathchandra *(Business School, Informatics Institute of Technology, Sri Lanka)*	80

TABLE OF CONTENT

COMPUTING

Paper ID	Title and Author Page	Page
25	**Developing An Automated System to Analyzing the Software Developers Interest Using Twitter Data** [1] K. N. M. S. Nandasiri, [1] K. G. L. S. Ananda, [1] D. V. S. Maduranga, [1] K. D. H. S. Bandara, [1] C. P. Wijesiriwardena, [2] I. T. S. Piyatilake *([1] Department of Information Technology, University of Moratuwa, Sri Lanka, [2] Department of Computational Mathematics, University of Moratuwa, Sri Lanka)*	94
38	**Predicting the Possibility of Chronic Pancreatitis** Bawantha Wickramarachchi, Iresh Bandara *(Department of Computing, Informatics Institution of Technology, Sri Lanka)*	102
5	**Optimization of Climatic Parameters of the Greenhouse to Maximize Chilli Production** A.P. Wijayawantha, M.T.U.D. Mallikarathne, A.M.H.M. Abeysinghe, C.S.Silva *(Department of Instrumentation and Automation Technology, Faculty of Technology, University of Colombo, Sri Lanka)*	113
60	**Machine Learning Approach to Predict Next Transaction Day with Customer Lifetime Value** Ms. Punsisi Kaludewa, Ms. Rukshala Weerasingha *(Department of Computing, Informatics Institute of Technology, Sri Lanka)*	122
7	**Empowering Early Childhood Education Students: A User-Centered Approach to Enhance Spelling Skills** Shidn De Silva, MTA Wickramasinghe, D Gunasekara *(Department of Information Technology, Faculty of Computing, General Sir John Kotelawala Defence University, Sri Lanka)*	133
24	**Alternative Transport Layer Protocol** [1] Harshi P. S. Ratnayake, [2] Geethapriya Liyanage *([1] Informatics Institute of Technology affiliated to University of Westminster, Sri Lanka, [2] Department of Computer Science and Engineering, University of Moratuwa, Moratuwa)*	162
20	**Developing a System to Transliterate Singlish Twitter Posts to Sinhala** [1] M. D. C. Amarasekara, [1] R. A. D. P. Rajapaksha, [1] H. M. D. T. Jayarathna, [1] H. M. G. K. Karunarathna, [1] C. P. Wijesiriwardana, [2] I. T. S. Piyatilake *([1] Department of Information Technology, University of Moratuwa, Sri Lanka, [2] Department of Computational Mathematics, University of Moratuwa, Sri Lanka)*	144
42	**Helaa: A Sinhala Language-Based Programming Language** Rajitha Yasasri, Damitha Karunarathna *(University of Colombo School of Computing, University of Colombo, Sri Lanka)*	151

TABLE OF CONTENT

TECHNOLOGY OF BUSINESS

Paper ID	Title and Author Page	Page
32	**AlzhCare: A cognitive therapy application for Alzheimer's Patients with smart results, and a safety tool to support their Caregivers** Jesica Rayen, Namal Malalasena *(School of Business, Informatics Institute of Technology, Sri Lanka)*	179
21	**The Impact of Electronic Word-of-Mouth on Customer Purchase Intention** Vishaleni Sureshan, Shonali Aponso *(School of Business, Informatics Institute of Technology, Sri Lanka)*	191
52	**Remaddict: A Solution Put Forward To Predict Withdrawal Symptoms And Act As A Maintenance Tool To Assist Adolescents Challenged With Substance Use Disorder** Avisha de Mel, Janice Abeykoon *(School of Business, Informatics Institute of Technology, Sri Lanka)*	203
26	**The Influence of Online Shopping Experience and Attitudinal Loyalty on Online Impulsive Buying in Sri Lankan Online Fashion Retail Industry** Chethana Gallage, Gayashini Ratnayake *(School of Business, Informatics Institute of Technology, Sri Lanka)*	213
22	**Impact of Social Media Communication and Brand Equity on Consumer Perceptions of Brands in the Fast-Food Industry** Ashweka Vasanthakumar, Gayashini Ratnayake *(School of Business, Informatics Institute of Technology, Sri Lanka)*	225
54	**Influence of social media marketing using Instagram and consumer trust on consumer purchase intentions in the skincare industry in Colombo District in Sri Lanka** Hansini Navodya Somaratna, Gayashini Ratnayake *(School of Business, Informatics Institute of Technology, Sri Lanka)*	237
65	**Mama Aid - An application to identify postpartum complications and assist in re-establishing mother's health** Erandi Kavindya Liyanagamage, Abarnah Kirupananda *(School of Business, Informatics Institute of Technology, Sri Lanka)*	249
19	**Influence of Internal Employer Branding and Organizational Trust on Employee Engagement In The IT Industry in Sri Lanka** Sashini Vasanthan, Gayashini Ratnayake *(School of Business, Informatics Institute of Technology, Sri Lanka)*	258

TABLE OF CONTENT

TECHNOLOGY OF BUSINESS

Paper ID	Title and Author Page	Page
67	**Ambula - A Mobile Application Solution to Promote Sri Lankan Food Culture** [1] Sarah Hausotter, [2] Lakshika Paiva *([1] School of Business, Informatics Institute of Technology, Sri Lanka, [2] WSO2 Lanka Pvt. Ltd., Colombo 04, Sri Lanka)*	270
31	**Influence of Supervisor Behavior & Social Support on Employee Wellbeing in the Apparel Industry of Sri Lanka** Osandi Udalamatta, Gayashini Ratnayake *(School of Business, Informatics Institute of Technology, Sri Lanka)*	283

TABLE OF CONTENT

EXTENDED ABSTRACTS - DATA SCIENCE & AI

Paper ID	Title and Author Page	Page
16	**Voice-bot for book an appointment** N. M. Pathirage, D. W. W. Gunasekara, A. P. Kumara, N. Madanayake, C. Bandara *(SL T digital lab, Sri Lanka Telecom PL C, Sri Lanka)*	295
17	**Dialogue System for Lifelong Learning of Verified Knowledge** T. A. Tennakoon, D. S. Alwis *(School of Computing, Informatics Institute of Technology, Sri Lanka)*	299
40	**DiabetCare – An Artificial Intelligence based Mobile Application for Predicting the Risk of Prediabetes Mellitus, Gestational Diabetes Mellitus and Type 2 Diabetes Mellitus** N. J. Pathiranage, S. M. D. A. S. Suraweera *(School of Computing, Informatics Institute of Technology, Sri Lanka)*	302
43	**SyntaxSwap: Supervised Java to JavaScript Code-To-Code Translator With Relative Positional Encoding Attention Mechanism** D. D. Jayamanne, I. Bandara *(School of Computing, Informatics Institute of Technology, Sri Lanka)*	305
49	**Development of a Smart Shoe for Health Monitoring with Energy Harvesting** B. Sathiyakeerthy, L. Lokeswaran, B.F.Rajon, C.S.Silva *(Faculty of Technology, University of Colombo, Sri Lanka)*	308
50	**Sri Oryzaugur: A Multivariate Approach to Forecast Rice Prices in Sri Lanka** [1] L. Rumaiz, [2] G.Liyanage *([1] School of Computing, Informatics Institute of Technology, Sri Lanka, Sri Lanka, [2] Ascentic Pvt Ltd, Colombo, Sri Lanka)*	311
63	**Depression Detection in Sinhala-English Code-Mixed Language using Social Media Data** F. Fazal, C. Farook *(School of Computing, Informatics Institute of Technology, Sri Lanka)*	314

TABLE OF CONTENT

EXTENDED ABSTRACTS - COMPUTING

Paper ID	Title and Author Page	Page
18	**Transfer Learning-based Approach to Detect Onychomycosis Fungal State & Fungal Growth on the Nail** P. Bandara, S. Perera *(School of Computing, Informatics Institute of Technology, Sri Lanka)*	317
62	**Automated Deception Detection System Using Human Facial Cues and Facial Chirality** C. G. Vidanagam, A. Kirupananda *(School of Computing, Informatics Institute of Technology, Sri Lanka)*	320



TABLE OF CONTENT

EXTENDED ABSTRACTS - TECHNOLOGY OF BUSINESS

Paper ID	Title and Author Page	Page
35	**CricPlayerTracker: An Application for identifying cricket talents in Sri Lanka** G. Karunananthan, A. Mascranghe, A. Husni *(School of Computing, Informatics Institute of Technology, Sri Lanka)*	324
36	**WorkPromo - Performance Evaluation System for Employee Promotion Eligibility of the IT Industry Sri Lanka** A. Senevirathna, S. Siriwardene *(School of Computing, Informatics Institute of Technology, Sri Lanka)*	327
37	**Warranty Wallet: A Platform to Integrate Manufacturers, Retailers, and Consumers for a Pitfall-Free Warranty Process** M. Munasinghe, S. Siriwardene, S. Dinushika *(School of Computing, Informatics Institute of Technology, Sri Lanka)*	331
48	**Influence of Supervisory Support and Work-Life Balance (WLB) on Organizational Citizenship Behavior (OCB) in the Apparel Industry in Sri Lanka** S. Udayanga, G. Rathnayake *(School of Computing, Informatics Institute of Technology, Sri Lanka)*	334
53	**Influence of Customer Engagement and Brand Attachment on Brand Loyalty in the Fashion Industry in Sri Lanka** S. Alwis, G. Ratnayake *(School of Computing, Informatics Institute of Technology, Sri Lanka)*	337



iCIIT 2023

THE 4TH INTERNATIONAL CONFERENCE ON INNOVATIONS IN INFO-BUSINESS & TECHNOLOGY

Conference Track

DATA SCIENCE & AI



Exploring Emotion Classification in a Labeled Twitter Dataset: A Comparative Study of Machine Learning Approaches

Emotion Classification in Labeled Tweets

Sugeeshwa S P Galhena

University of Colombo School of Computing, University of Colombo, Colombo, Sri Lanka, sugeeshwag@gmail.com

Ajantha S Atukorale

University of Colombo School of Computing, University of Colombo, Colombo, Sri Lanka, aja@ucsc.cmb.ac.lk

ABSTRACT

Due to the industry revolution 4.0, the availability of textual content has increased rapidly throughout the years. Therefore, sentiment analysis has become an important area in machine learning. This study aims to develop a mechanism to identify the hidden emotion in textual content, which goes beyond the three sentiments positive, neutral, and negative. The study explores machine learning approaches to identifying emotions in textual context. Several techniques are experimented with, such as Naïve Bayes Classifier, Support Vector Mechanism, Regression, Decision Tree, and Random Forest. Applying basic steps and using simple linear models achieves an accuracy as high as 90.5%, suggesting that complex algorithms are not always required to achieve high accuracy. Performance matrices such as accuracy, weighted F1-score, and efficiency evaluate the performance of the output.

KEYWORDS

emotion detection, sentiment analysis, machine learning, supervised learning, text classification

1 Introduction

Detecting emotion in the text is a field closely related to sentiment analysis. While sentiment analysis aims to uncover the three opinions the text contains, such as positive, negative, or neutral, emotion analysis focuses on discovering the feeling behind the text, such as anger, fear, joy, and sadness. Reading an online review to decide on a novel to read or an item to purchase online could be considered an example of sentiment analysis one faces during day-to-day life [14]. Humans feel a spectrum of emotions, as opposed to the abovementioned three basic sentiments, positive, negative, and neutral. In real life, sentiment does not impact the evaluation of emotions. For example, book reviews "That book was awesome!" and "That book was good" would be of positive sentiment, but the emotionality behind the two reviews has two different meanings. The Figure 1 illustrates this idea graphically.

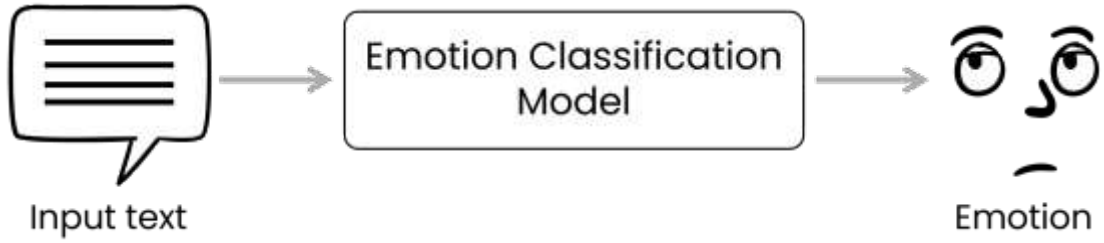


Figure 1: Emotion classification in text.

[2], and [6] did thorough research on approaches to emotion detection from text with emotion models and approaches. The latter categorized approaches based on classification techniques that are used to identify emotions. The former did a critical analysis of the approaches and suggested that both lexicon-based approaches and machine-learning approaches can be used in the analyses. The best algorithm that provides better results is selected by comparing the results with other factors. Indicators such as recall, precision, accuracy, confusion matrix, and F1 values are calculated to measure the algorithm's performance. [4] proposed to calculate both accurate and unweighted measurements.

This study aims to review and analyze different methodologies and techniques used for text-based emotion detection. Emotion research is done based on the emotion classifications that have served as the foundation for this research. The study [2] consists of an extensive survey to differentiate between the advantages and disadvantages of emotion classification models and the approaches that have been utilized in the past for text mining. They recommended that depending on the aims and the needs of the study, lexicon-based approaches or machine learning-based approaches can be applied in extracting emotion from the text.

To accomplish these objectives, appropriate hypotheses are being proposed based on past research. Furthermore, outcomes of the intermediate studies are utilized to make inferences. By investigating the existing approaches, this paper aims to provide insights into the field's current state and identify potential areas for improvement and future research.

1.1 Emotion Classification Models

Over the years, different emotion classification methods have been introduced, categorized into emotion dimensions and emotion categories. Emotion dimensions provide more information about the emotions, such as the positivity or negativity of the emotion, arousal of the emotion, and degree of the emotion. In contrast, emotion categories are quite popular among machine learning researchers due to their simplicity. These concepts are explained extensively in the paper [1] as Dimensional Emotion Models (DiEMs), where emotions are located in a special space that suggests a dependency between the emotions, and Discrete Emotion Models (DEMs), where the emotions are located in different categories.

Multiple DiEMs discussed in the paper [1] are Rusell's Circumplex Model of Affect, which suggests that emotions are in the arousal-valence dimensions, and Plutchik's 2D Emotion of Wheel. The latter proposes a model with arousal on the horizontal axis and valence on the vertical axis. Rusell & Mehrabian's Model suggests that emotions are distributed in a 3D space according to arousal, valence, and dominance. In their study [8], the researchers did a similar study about emotional combinations in content that goes viral. The researchers considered more than 65,000 articles in their study. They concluded that it is not the emotion that makes content go viral but where the emotions fall within the Valence-Arousal-Dominance (VAD) model. As stated in the article, an emotion combines three characteristics: Valence, Arousal, and Dominance.

The Paul Ekman Model suggests six fundamental emotions (anger, disgust, fear, happiness, sadness, surprise). The Robert Plutchik Model suggests eight primary emotions that occur in opposite pairs (anger vs. fear, joy vs. sadness, surprise vs. anticipation, trust vs. disgust), and Orthony, Clore, and Collins (OCC) Model which suggests 22 emotions as opposed to the basic emotions suggested in both Ekman's and Plutchik's models by adding additional classes such as appreciation, envy, pity, relief, and shame are the DEMs that are discussed on the paper [1]. The paper [11] recommends multiple DEMs defined by

psychologists. Two other such models are Shaver Model and Tomkins Model. The Shaver Model consists of six emotions: sadness, joy, anger, fear, love, and surprise, which can be represented in a tree structure, while the Tomkins Model comprises nine emotional states. Out of these nine, only three emotions are positive, and most of these can be represented as pairs. The emotions he introduced are disgust, surprise-startle, anger – rage, anxiety, fear - terror, contempt, joy, shame, interest – excitement. The paper [9] suggests that Parrot's Model, a DEM, which contains a balanced set of classes: love, anger, sadness, joy, fear, and surprise improves the accuracy as compared to the other techniques when it is used on analyzing blog data. Although emotional categories do not cover all emotions, most computer-based approaches base this on their simplicity and familiarity.

1.2 Emotion Detection

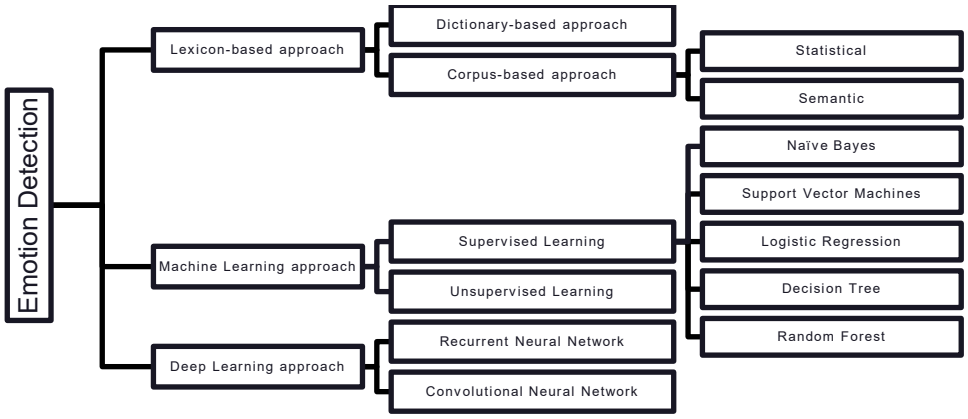


Figure 2: Computational approaches for emotion detection.

Their survey [2] stated that the methods used to detect emotions could also be divided into lexicon-based and machine learning-based approaches. Keyword-based approaches and ontology-based approaches fall into lexicon-based approaches. The statistical approach, such as Latent Semantic Analysis, is also considered a lexicon-based approach. The paper [11] analyzed different machine learning techniques conducted in the past for classifying emotions in text. Figure 2 illustrates approaches to solve the problem. Lexicon-based approaches use word dictionaries, assigning each one with a sentiment value. The accuracy of dictionary-based approaches depends on the algorithm used for the analysis. In contrast, corpus-based approaches are more efficient in a particular domain because they maintain domain-specific terms.

A few machine learning approaches discussed in the paper are Naïve Bayes, Support Vector Machine (SVM), and Decision Trees. The studies such as [5], which applied the SVM to a collection of text collected via customer feedback surveys, and [16], which applied Naïve Bayes, SVM, and n-gram models in classifying online reviews about seven popular travel destinations around the world experimented with these machine learning models. The second experiment claimed to achieve more than 80% calculation accuracy for all three models. Other methodologies that are being discussed in the paper [11] are deep learning and hybrid approaches in classifying textual content. It is stated that hybrid approaches such as Random Forest Support Vector Mechanism (RFSVM) have performed well. [14] experimented with recognizing fifteen different emotions which could be indicative of suicidal behavior using binary SVM classifiers. In their study on emotion detection using deep learning approaches, [4] discussed a few challenges one would face in the NLP research area are word-sense disambiguation (E.g., "Shut up!") and coreference resolution. Thus, the article suggests doing experiments with different machine learning models to conclude. The study [14] surveyed mining and classifying customer reviews, another aspect of text classification opinion mining. The paper suggested supervised, unsupervised, and semi-supervised methods for text classification. Out of the research that the paper has surveyed, it is mentioned that the Naïve Bayes is one of the popular models and has performed efficiently in a study that was carried out for classifying Urdu and English opinions in a blog. Another approach that was discussed in the paper is SVM.

The paper [12] analyzed two different datasets of movie reviews with multiple approaches, including a deep learning approach. Different approaches were taken to analyze data, such as Bag-of-Words (BOW) followed by a Random Forest Classifier, BOW followed by SVM, Word2Vec followed by Random Forest Classifier and Word2Vec, Clustering followed by Random Forest Classifier. By applying Recursive Neural Network (RNN), they achieved higher accuracy in predicting the text category. In their study, [15] used a Convolutional Neural Network (CNN) to classify Chinese microblogs. They used a dataset of 1.6 million tweets and conducted several deviations of the CNN models. After multiple experiments, the highest accuracy the model was able to achieve was 86.83%. In the study [15], the researchers introduced various methodologies that one can take to identify the sentiment behind the text and its applications. Some of the methods that were discussed in the study are RNN, CNN, and Long Short-Term Memory (LSTM). The paper [7] proposed a transformation method to convert a multi-level classification problem into a binary classification problem and was solved by suggesting a deep learning approach. The proposed model has outperformed the achieving a higher score on the SemEval2018 Task 1: E-c multi-label emotion classification problem.

Another study focused on emotion analysis [10] measured the emotion and intensity of the emotion in tweets. They claimed it is important for applications to know the degree to which an emotion is expressed in text. Their findings are presented in an interactive visualization [17], allowing the readers to cross-filter. [3] published a paper based on their findings for SemEval-2019 task 3, which focused on contextual emotion detection in text. In a study by [14], they proposed a way to improve three-way decisions using fuzzy logic and deep learning. The reason behind this model is to provide a framework when uncertainty is present in the reviews. In their study, they used feature selection methods such as Term Frequency-Inverse Document Frequency (TF-IDF), bm25, Uniformity (Uni), and Inverted Conformity Frequency (ICF). The study selected and evaluated eight state-of-the-art models based on their performance. Model performances were assessed using Precision, Recall, and F-measure. However, the results showed that the proposed model's performance was better than the other models in the two datasets considered in the study.

2 MATERIALS AND METHODS

2.1 Dataset

The emotion-annotated dataset [13] is used for this study. This dataset is publicly available for research purposes. The data set is annotated based on Parrot's Model, and the dataset metrics are mentioned in Table 1. The distribution of sample length (i.e., the length of an input record) is depicted in Table 1 Emotion dataset metrics. This data set is already broken down into training, validation, and test sets.

Table 1 Emotion dataset metrics.

Metric Name	Number of Samples	Number of Classes	Number of Samples per Class					
Value	20000	6	Joy	Sadness	Anger	Fear	Love	Surprise
			6761	5797	2709	2373	1641	719

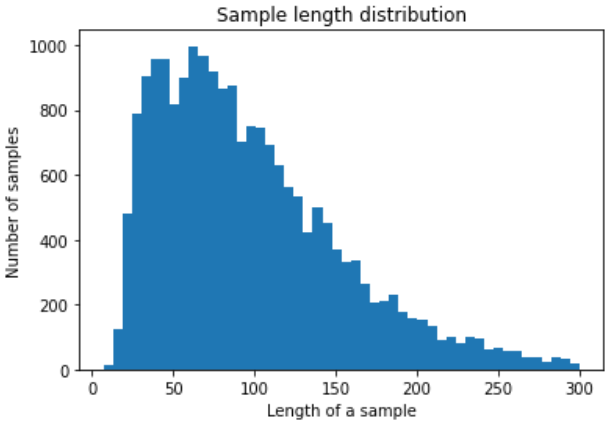


Figure 3: Sample length distribution.

2.2 Breakdown of Tasks

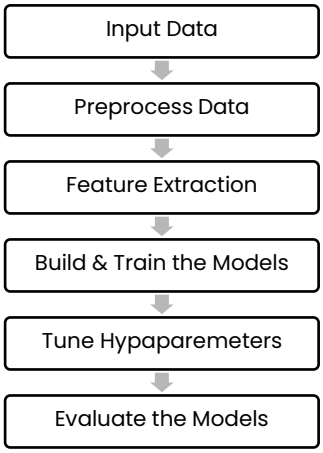


Figure 4: High-level overview of the workflow.

Figure 4 gives a high-level overview of the task breakdown. The model training is broken down into multiple subtasks, and various approaches are carried out during these subtasks. Preprocessing is done for the data to make sure that data are in a format that is understandable by the machine learning algorithms. Preprocessing techniques that are carried out are converting all the text into lowercase, removing stop words from the data set, tokenization, and vectorization. Several subtasks are carried out during each step, such as for vectorization, and both BOW and TF-IDF methods are carried out. BOW defines a fixed-length count vector, with each entry corresponding to a word in the predefined word dictionary. A word in a sentence is assigned a count of 0 if it is not in the defined dictionary; otherwise, it is assigned a count of 1 or higher depending on how many times it appears in the sentence. Although this is an easy-to-implement method, its drawbacks, such as not capturing the order of the words and not considering the meaning of the sentence, are impactful when it comes to practical applications [11]. TF-IDF represents the text in matrix form, with each number quantifying the amount of information those terms contained in a particular document. The assumption behind this method is that rare words contain more information than others.

The reason for carrying more than one approach for the subtasks is to compare the results from each task and select the best approach for each task, which provides the output with the highest accuracy in less time.

The experiments were carried out in Google Colaboratory in Python environment. The Python version that is used in the study is 3.7. The relevant libraries, such as NumPy, Pandas, Matplotlib, Scikit-learn, and Seaborn, were installed and called during the deployment.

2.3 Classification Models

2.3.1 Naïve Bayes Classifier

Naïve Bayes Classifier is the most commonly used classifier, and its simplicity has been the reason behind universal usage. Naïve Bayes Classifier computes the posterior probability of a class based on the distribution of the words in the document.

$$P(label|features) = \frac{P(label)*P(features|label)}{P(features)} \quad (1)$$

2.3.2 Support Vector Mechanism

Support Vector Classifier determines linear separators in the search space, which can best separate different classes. This is widely used in classification, regression, and outlier detection. This was implemented in Python with the help of SGDClassifier.

2.3.3 Logistic Regression

Logistic regression is the process of modeling a linear relationship between the dependent variable and the independent variables when the outcome is binary. For this task, the logistic regression model is generalized to multiple classes.

2.3.4 Decision Tree

The decision tree is a tree-like structure where a branch represents a decision rule; a node represents an attribute. The nodes at the bottom of the tree are known as leaves, and the topmost node is known as the root node. Attribute selection can be done using multiple methods, such as information gain and maximum entropy.

2.3.5 Random Forest

Random Forest classifier is suitable when dealing with high dimensional noisy data. This is a collection of decision trees.

2.4 Deep Learning Model

2.4.1 Convolutional Neural Network (CNN)

The Convolutional Neural Network is a neural network type that follows a feedforward architecture. Therefore, the connection between two nodes does not perform a cycle. The CNN can perform feature extraction with minimal preprocessing, which is considered one of the key advantages of the same.

2.4.2 Recursive Neural Network (RNN)

The Recursive Neural Network is a neural network type derived from feedforward architecture. However, this differs from CNN as connections between nodes can create a cycle. In this paper, LSTM is applied, a type of RNN that is effective in modeling sequential data.

The architecture of the deep learning model implementation goes as follows. The implementation is done using Keras, a TensorFlow API.

- Embedding layer: The input text sequences are first passed through the embedding layer. This layer learns and maps each word in the input sequence to a dense vector representation. The embedding layer helps capture words' semantic meaning and contextual relationships.
- LSTM layer: The output from the embedding layer is fed into the LSTM layer. This captures long-term dependencies in the input sequences. The LSTM layer consists of LSTM units that process the input sequence and propagate information through time steps.
- Dense layer: The output from the LSTM layer is then fed into the dense layer. This performs a linear transformation on the input and applies an activation function.

In this paper, the activation function used in the dense layer is softmax, which outputs a probability distribution over the possible classes. The model is trained using the categorical cross-entropy loss function and optimized using the Adam optimizer.

2.5 Hyperparameter Tuning

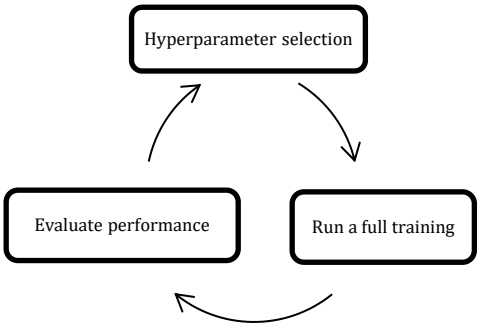


Figure 5: Hyperparameter tuning process.

Hyperparameter tuning is the process of identifying the combination of values, which either increases the model's accuracy or reduces the model's training time. Selecting the best set of values is an iterative process that depends on time and money. Figure 5 illustrates the hyperparameter tuning process. Hyperparameter tuning is done not only for the classifiers but also for the feature selection methods. Nevertheless, hyperparameter tuning is an expensive and lengthy process. Therefore, this was done only for the most vital hyperparameters, and the tuning was done with the objective of increasing accuracy. With multiple trial-and-error experiments, the optimal values for the hyperparameters are identified. Hyperparameters can be tuned using multiple ways: Grid Search and Random Grid Search are examples. A search space is defined as a grid for hyperparameters, and every position in the grid is searched and evaluated in the grid search. In randomized grid search, a search space is defined as a bounded domain of hyperparameters, and values are evaluated randomly. In this paper, hyperparameter tuning is implemented using grid search.

2.6 Evaluation

Precision, recall, and F1-score are used to decide how well the model fits on unseen data. These performance indicators measure the predictive accuracy of the model. Efficiency measures the time to construct the model and the time it takes to use the model. This plays a huge role as running a model for a long time is highly costly.

3 RESULTS AND DISCUSSION

The model fitting is done on the training set and evaluated the model on the validation set, and then the hyperparameter tuning is done on the latter. This was a repetitive process, and the model that performed the best on the validation set is selected, and the results were confirmed on the test set. The performance of each model is evaluated based on predetermined evaluation criteria.

Table 2 Results before hyperparameter tuning.

Model	Feature extraction	Accuracy	Weighted average for F1-score
Naïve Bayes Classifier	BOW	78.80%	0.76
	TF-IDF	67.75%	0.61
Support Vector Mechanism	BOW	89.85%	0.90
	TF-IDF	90.15%	0.90
Generalized Logistic Regression	BOW	89.85%	0.90
	TF-IDF	87.30%	0.90
Decision Tree	BOW	88.50%	0.87
	TF-IDF	86.70%	0.87
Random Forest	BOW	88.30%	0.88
	TF-IDF	88.95%	0.89

Before hyperparameter tuning, the SVM using BOW and TF-IDF features performed the best in accuracy and weighted F1-score. The Decision Tree and Random Forest models showed relatively lower performance than the SVM and Logistic Regression models but achieved respectable accuracy and weighted F1-scores.

Table 3 Results after hyperparameter tuning.

Model	Feature extraction	Accuracy	Weighted average for F1-score	Time taken to train the model
Naïve Bayes Classifier	BOW	78.80%	0.25	0.189s
	TF-IDF	75.10%	0.26	0.193s
Support Vector Mechanism	BOW	90.15%	0.90	14.84s
	TF-IDF	90.20%	0.90	11.28s
Generalized Logistic Regression	BOW	89.35%	0.90	22.4s
	TF-IDF	89.65%	0.90	7.362s
Decision Tree	BOW	86.40%	0.87	1.877s
	TF-IDF	86.70%	0.87	1.788s
Random Forest	BOW	88.35%	0.88	19.705s
	TF-IDF	89.10%	0.89	16.575s

After hyperparameter tuning, the SVM using BOW and TF-IDF features remained the top performer, achieving high accuracy and weighted F1-scores. The Generalized Logistic Regression models, both before and after hyperparameter tuning, also achieved competitive performance. Based on the provided results, the SVM with either BOW or TF-IDF features is the best-performing model, with high accuracy and weighted F1-score. The Confusion matrix and the SVM model results after hyperparameter tuning are shown in Table 4, Figure 6, Table 5 and Figure 6.

Table 4 Classification report of SVM with BOW after hyperparameter tuning.

	Precision	Recall	F1-score	Support
Anger	0.89	0.90	0.90	275
Fear	0.87	0.88	0.88	224
Joy	0.93	0.92	0.92	695
Love	0.75	0.81	0.78	159
Sadness	0.95	0.94	0.94	581
Surprise	0.74	0.70	0.72	66
Accuracy			0.90	2000
Macro avg	0.86	0.86	0.86	2000
Weigted avg	0.90	0.90	0.90	2000

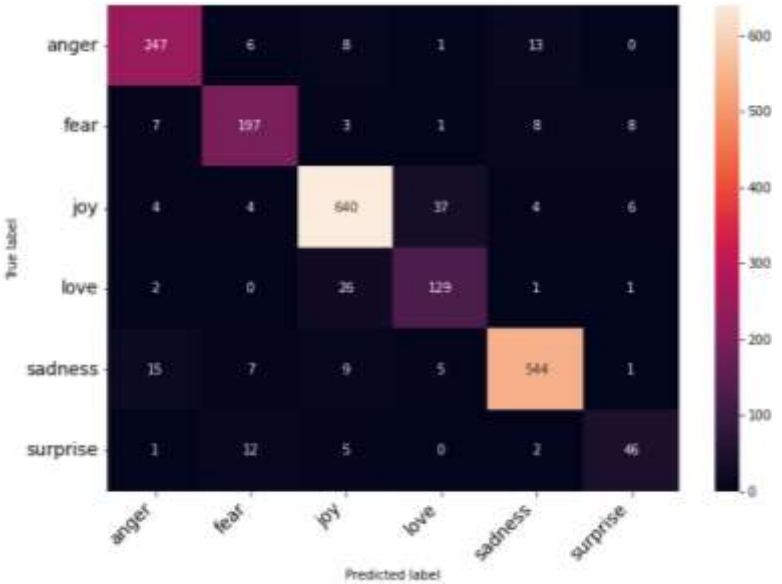


Figure 6: Results of SVM with BOW after hyperparameter tuning.

Table 5 Classification report of SVM with TF-IDF after hyperparameter tuning.

	Precision	Recall	F1-score	Support
Anger	0.89	0.90	0.90	275
Fear	0.90	0.80	0.85	224
Joy	0.90	0.96	0.93	695
Love	0.89	0.69	0.78	159
Sadness	0.94	0.95	0.94	581
Surprise	0.67	0.67	0.67	66
Accuracy			0.90	2000
Macro avg	0.86	0.83	0.84	2000
Weigted avg	0.90	0.90	0.90	2000

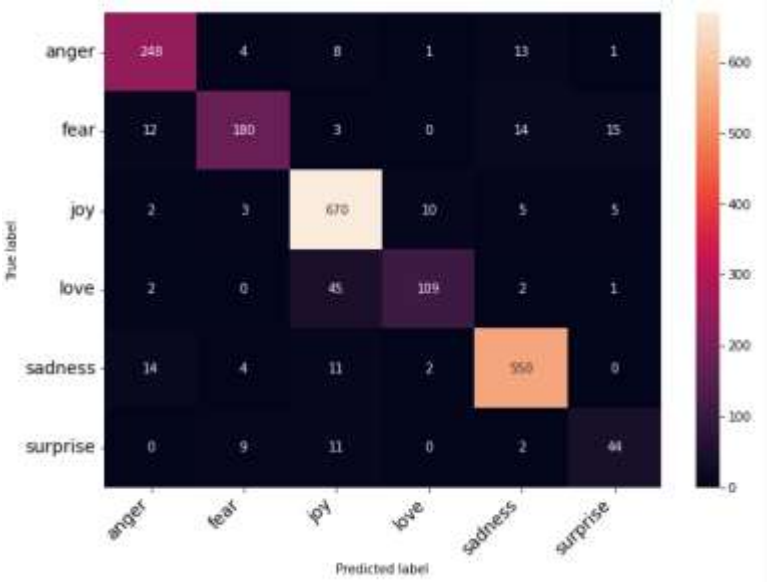


Figure 7: Results of SVM with TF-IDF after hyperparameter tuning.

The SVM models using BOW and TF-IDF features achieved similar high accuracy of around 0.90. The models performed well across most classes, with high precision, recall, and F1-scores. The "joy" and "sadness" classes consistently performed well in both models. The TF-IDF-based SVM model had slightly lower precision, recall, and F1-scores for the "fear" and "love" classes than the BOW-based model. Overall, both models demonstrated strong performance in classifying emotions, with the SVM model using BOW features having a slight edge in precision and recall.

The deep learning model was created with LSTM-based RNN with embedding, dropout, LSTM, and dense layers. The model is trained for a specific number of epochs using early stopping based on the validation

loss. In this case, the epoch was chosen as 5. Then the loss and accuracy values are tested, and training and validation loss/accuracy curves are plotted as in Figure 8.

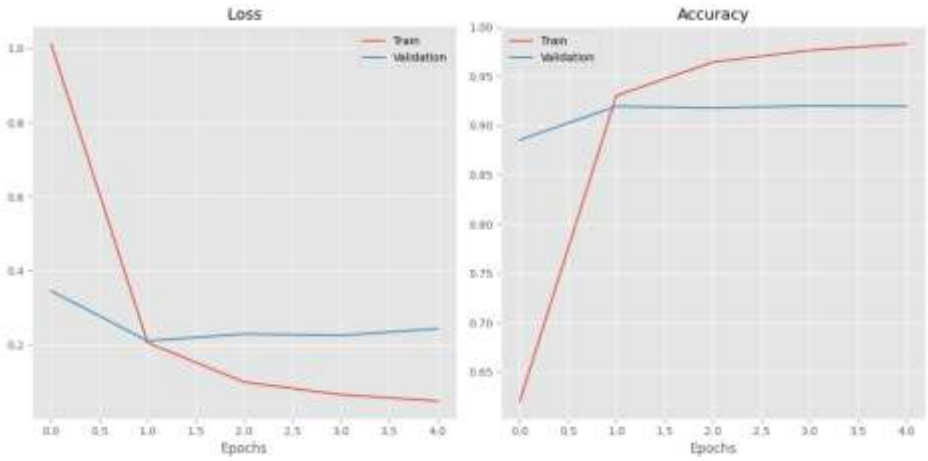


Figure 8: Loss and accuracy function plots for training and validation sets.

The loss plot shows the change in loss values over the epochs for both the training and validation sets. The training loss shows how well the model fits the training data, while validation loss shows how well the model fits the validation data, which is untouched. Since the training loss is lower than the validation loss, it may suggest that the model is overfitting. This applies to the accuracy plot as well. The accuracy plot shows the change in accuracy values over the epochs for both the training and validation sets. The training accuracy shows how well the model fits the training data, while validation accuracy shows how well the model fits the validation data. The training accuracy is higher than the validation accuracy, indicating that the model is overfitting. Nevertheless, the confusion matrix is generated for the above output, as shown in Table 6 and Figure 9.

Table 6 Classification report of the model after hyperparameter tuning.

	Precision	Recall	F1-score	Support
Anger	0.91	0.93	0.92	275
Fear	0.93	0.83	0.88	224
Joy	0.94	0.94	0.94	695
Love	0.80	0.84	0.82	159
Sadness	0.96	0.96	0.96	581
Surprise	0.76	0.88	0.82	66
Accuracy			0.92	2000
Macro avg	0.88	0.90	0.89	2000
Weigted avg	0.92	0.92	0.92	2000



Figure 9: Confusion matrix of the model after hyperparameter tuning.

The model's overall accuracy resulted in an accuracy level of 92.2%. The macro-average F1-score, precision, and recall across all classes were 0.89, 0.88, and 0.90, respectively. The weighted average F1-score, precision, and recall were all 0.92, indicating good overall performance for the model.

4 CONCLUSION

The machine learning models experimented with are Naïve Bayes, Support Vector Mechanism, Regression, Decision Tree, and Random Forest. High accuracy was achieved after following basic steps such as BOW and TF-IDF. It shows that to accomplish higher high accuracy, it is not always required to use complex algorithms. After tuning the hyperparameters, the models were evaluated again using the validation set. Once the hyperparameters were tuned, the models were evaluated on the hold-out set of data.

Since the dataset is biased (i.e., The number of records that fall into each label is different), the weighted average for F1-score was used to assess the model performances. Another metric that is used to measure the model performance is the time taken to train the model. It is noticed that models (except for the SVM model) took more time to train when the BOW was used as the feature extraction method instead of the TF-IDF method. The Random Forest model took the highest time to train. It achieved an accuracy level of 89%, while the Naïve Bayes model took less than 1s to train, although its accuracy was comparatively low. When the cost factor plays a big role in modeling, one can select an average-performing model which takes less time to train. As the objective of a machine learning model is to maximize accuracy and minimize the computation time required for training, the models above are preferred.

The deep learning model was fitted on the data using Keras, a Python-based neural network library. The model type that is used is sequential modeling which has allowed us to build a linear stack of layers. LSTM network is implemented on the data, which is an RNN [18]. When the model's performance is evaluated, it is visible that the deep learning model outperforms all the machine learning models. The weighted average for F1-score was 92%, while the model's accuracy was also 92%.

One of the limitations one faces while working on emotion analysis in the text is less availability of emotion-annotated data. Furthermore, the available data are in the English language. For hyperparameter tuning in the machine learning models, the GridSearch method is proposed due to the literature review. Some approaches, such as RandomSearch, and Bayesian Optimization, would have been applied to the models to identify the method that performed well on the chosen dataset. As stated by [3] another problem a researcher faces when detecting emotions from text is the absence of facial expressions and voice modulations.

Furthermore, the lack of contextual understanding could lead to a change in emotion classification. The growing use of internet slang has hindered sentiment analysis research. For example, slang such as 'lol' stands for laughing out loud, and 'bussin" is the new way of explaining that something is good. Furthermore, there is a tendency to explain anger in the form of sarcasm. Using machine learning algorithms, this is not easy to detect, although humans can judge it easily. Another challenge that researchers in this domain face is that one sentence could contain multiple emotions. For example, let us take the sentence, "Customer service at restaurant one was worse than the service at restaurant two." To the human brain, this gives the idea that restaurant two's service was better than the service provided at restaurant one. Nevertheless, a machine understands this negatively because of the word "worse." Although it is used positively, since the word "worse" is used negatively, the whole sentence is categorized negatively.

Measuring the intensity of the emotion is one of the projects that can be done in the future. Since discrete emotional models have been used for this experiment, neither the emotion's intensity nor the emotion's valance is considered. Humans feel a spectrum of emotions; thus, this research analyzes to which percentage one aspect of emotion is aroused. For example, one can feel angry and sad simultaneously in a given situation. Therefore, identifying to which degree one emotion has aroused is another aspect that can be considered.

REFERENCES

[1] Francisca Adoma Acheampong, Chen Wenyu, and Henry Nunoo-Mensah. 2020. Text-based emotion detection: Advances, challenges, and opportunities. *Engineering Reports* 2, 7 (2020), e12189. DOI:https://doi.org/10.1002/eng2.12189

[2] Lea Canales and Patricio Martínez-Barco. 2014. Emotion Detection from text: A Survey. In *Proceedings of the Workshop on Natural Language Processing in the 5th Information Systems Research Working Days (JISIC)*, Association for Computational Linguistics, Quito, Ecuador, 37–43. DOI:https://doi.org/10.3115/v1/W14-6905

[3] Ankush Chatterjee, Kedhar Nath Narahari, Meghana Joshi, and Puneet Agrawal. 2019. SemEval-2019 Task 3: EmoContext Contextual Emotion Detection in Text. In *Proceedings of the 13th International Workshop on Semantic Evaluation*, Association for Computational Linguistics, Minneapolis, Minnesota, USA, 39–48. DOI:https://doi.org/10.18653/v1/S19-2005

[4] Yam Chew-Yean. 2015. Emotion Detection and Recognition from Text Using Deep Learning. *CSE Developer Blog*. Retrieved March 9, 2022 from https://devblogs.microsoft.com/cse/2015/11/29/emotion-detection-and-recognition-from-text-using-deep-learning/

[5] Michael Gamon. 2004. Sentiment classification on customer feedback data: noisy data, large feature vectors, and the role of linguistic analysis. In *COLING 2004: Proceedings of the 20th International Conference on Computational Linguistics*, COLING, Geneva, Switzerland, 841–847. Retrieved November 20, 2022 from https://aclanthology.org/C04-1121

[6] Ahmed Hassan Yousef, Walaa Medhat, and Hoda Mohamed. 2014. Sentiment Analysis Algorithms and Applications: A Survey. *Ain Shams Engineering Journal* 5, (May 2014). DOI:https://doi.org/10.1016/j.asej.2014.04.011

[7] Mohammed Jabreel and Antonio Moreno. 2019. A Deep Learning-Based Approach for Multi-Label Emotion Classification in Tweets. *Applied Sciences* 9, 6 (January 2019), 1123. DOI:https://doi.org/10.3390/app9061123

[8] Kerry Jones, Kelsey Libert, and Kristin Tynski. 2016. The Emotional Combinations That Make Stories Go Viral. *Harvard Business Review*. Retrieved April 10, 2022 from https://hbr.org/2016/05/research-the-link-between-feeling-in-control-and-viral-content

[9] Sheetal Kusal, Shruti Patil, Ketan Kotecha, Rajanikanth Aluvalu, and Vijayakumar Varadarajan. 2021. AI Based Emotion Detection for Textual Big Data: Techniques and Contribution. *BDCC* 5, 3 (September 2021), 43. DOI:https://doi.org/10.3390/bdcc5030043

[10] Saif Mohammad and Felipe Bravo-Marquez. 2017. Emotion Intensities in Tweets. In *Proceedings of the 6th Joint Conference on Lexical and Computational Semantics (*SEM 2017)*, Association for Computational Linguistics, Vancouver, Canada, 65–77. DOI:https://doi.org/10.18653/v1/S17-1007

[11] Pansy Nandwani and Rupali Verma. 2021. A review on sentiment analysis and emotion detection from text. *Soc. Netw. Anal. Min.* 11, 1 (August 2021), 81. DOI:https://doi.org/10.1007/s13278-021-00776-6
[12] Hadi Pouransari and Saman Ghili. Deep learning for sentiment analysis of movie reviews. 8.
[13] Elvis Saravia, Hsien-Chi Toby Liu, Yen-Hao Huang, Junlin Wu, and Yi-Shin Chen. 2018. CARER: Contextualized Affect Representations for Emotion Recognition. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics, Brussels, Belgium, 3687–3697. DOI:https://doi.org/10.18653/v1/D18-1404
[14] L. D. C. S. Subhashini, Yuefeng Li, Jinglan Zhang, Ajantha S. Atukorale, and Yutong Wu. 2021. Mining and classifying customer reviews: a survey. *Artif Intell Rev* 54, 8 (December 2021), 6343–6389. DOI:https://doi.org/10.1007/s10462-021-09955-5
[15] Yaqi Wang, Shi Feng, Daling Wang, Ge Yu, and Yifei Zhang. 2016. Multi-label Chinese Microblog Emotion Classification via Convolutional Neural Network. In *Web Technologies and Applications*, Feifei Li, Kyuseok Shim, Kai Zheng and Guanfeng Liu (eds.). Springer International Publishing, Cham, 567–580. DOI:https://doi.org/10.1007/978-3-319-45814-4_46
[16] Qiang Ye, Ziqiong Zhang, and Rob Law. 2009. Sentiment classification of online reviews to travel destinations by supervised machine learning approaches. *Expert Systems with Applications* 36, 3, Part 2 (April 2009), 6527–6535. DOI:https://doi.org/10.1016/j.eswa.2008.07.035
[17] EmoInt-web2. *Tableau Software*. Retrieved March 9, 2022 from https://public.tableau.com/views/EmoInt-web2/EmotionIntensityDashboard?:embed=y&:showVizHome=no&:host_url=https%3A%2F%2Fpublic.tableau.com%2F&:tabs=no&:toolbar=yes&:animate_transition=yes&:display_static_image=no&:display_spinner=no&:display_overlay=yes&:display_count=yes&publish=yes&:loadOrderID=0
[18] Understanding LSTM Networks -- colah's blog. Retrieved February 25, 2023 from https://colah.github.io/posts/2015-08-Understanding-LSTMs/

Health Care Data Privacy Preserving using Federated Learning: A case of Heart Disease Prediction*

A.S Aathika†
Computer Systems & Engineering, Sri Lanka Institute of Information Technology, Malabe, Western, Sri Lanka, aathikasalam7@gmail.com

L. Rupasinghe
Computer Systems & Engineering, Sri Lanka Institute of Information Technology, Malabe, Western, Sri Lanka, lakmal.r@sliit.lk

ABSTRACT

In the present, electronic health data is dramatically increased. The data privacy regulations limit the sharing of data in a centralized environment. Due to this reason, it is challenging to train models using machine learning algorithms specifically deep learning, which always require large amounts of training data. Federated Learning (FL) allows to train the data collaboratively from multiple data sources while keeping the sensitive health data private. This method poses by sharing the parameters of the model updates among participants. A high accuracy level can be achieved by properly aggregating the local models together to create a global model. However, from the final model original data can be inferred. This paper evaluates the application of Differential Privacy (DP) concept to preserve data privacy in the FL models. This research implements and evaluates a system for Federated Learning for heart disease prediction on Cleveland Hear Disease Dataset using TensorFlow Federated Framework. The research results revealed that there is a tradeoff between model accuracy and privacy. The research finally compares the accuracy difference between traditional Deep Learning, Federated Learning and the Differential privacy based Federated Learning

KEYWORDS

Federated Learning, Deep Learning, Data Privacy, Differential Privacy, TensorFlow, Federated Framework, Neural Network, Federated Average Algorithm,

1 Introduction

The amount of health-related information in the health care domain has increased significantly over recent years [1]. Deep Learning has produced enormous outcomes in the medical field. To train accurate deep learning models, it requires a large amount of data. Due to the data privacy laws and regulations such as Health Insurance Portability and Accountability Act (HIPAA) and General Data Protection Act (GDPR) it is infeasible to collect, store and process sensitive medical data for various purposes [2]. Hence it is infeasible to train the data in a centralized manner for any health-related predictions. The modern technology that solves this problem is the Federated Learning (FL) approach. FL enables collaborative data training from multiple institutions without sharing the data. Each node participating in the training process is known as clients. Each client trains the model with their own data and creates a local model and uploads the model to a central server. The central server is responsible to aggregate the local model to create to create a global model, which then shared with all the participated clients. In this case raw data is not shared with others and remains private. In FL setup rather than sending the data they share the parameters such as weight and updates between the client and the central server [3]. Similarly, FL approach eliminates data privacy and security issues by allowing collaborative learning while achieving a good model accuracy.

Even though Federated Learning can provide data privacy to sensitive medical data, it is also vulnerable to model inversion attack: trained information can be inferred from the global model and backtrack the participating clients [4]. One way of limiting this weakness is adding a certain level of noise to local models

during the training process. This research paper mainly focuses on implementing and evaluating privacy preserved FL for heart disease prediction using TensorFlow Federated Framework (TFF). Main objectives of this research are: (1) Implement and evaluate Federated Learning for heart disease prediction; (2) apply Differential Privacy concept to limit model inversion attack; (3) a comparison and discussion on how applying privacy to the model can impact the model accuracy.

1.1 Background

Federated learning is a distributed machine learning method where multiple participants collaborate to train a shared model while keeping their data on-device. It is designed to protect privacy and security of participants' data by not sharing it with a centralized server. This approach has been gaining popularity in recent years due to the increasing concerns over data privacy and security. It has been applied in various domains such as mobile and internet-of-things devices, health care, and finance [5]. Federated learning has shown promise in terms of performance, scalability, and privacy protection compared to traditional centralized approaches. In federated settings, the desire to protect privacy often requires raw data to be stored locally on each device. However, sharing information such as model updates during training can also reveal sensitive user information, raising privacy concerns [6] [7]. A set of security researchers showed how sensitive data can be extracted from the neural network model in Netflix contest. Another security researcher Carlina et al described shown that sensitive text patterns, such as a credit card number, can be extracted from a recurrent neural network trained on language data from users [8]. After the literature review it is identified there are various methods to provide privacy for FL setup, out of them differential privacy is the Differential privacy is favored due to its strong information security guarantees, ease of implementation [9] [10]. There are few related works in FL domain conducted to achieve various tasks.

Banks incur billions of dollars in losses each year due to credit card fraud, but sharing datasets is prohibited due to privacy concerns. A federated meta learning framework was proposed to address this issue. A central server coordinates the model training, with model parameters initially broadcasted. Participating banks refine the model using their own data and only send back the updated parameters. After multiple rounds of training, a highly effective model is obtained without compromising customer data. To address the issue of a skewed dataset with low fraudulent transactions, over-sampling techniques were used to balance the dataset. This system has the potential to greatly reduce workload for bank employees by preventing fraud before it occurs[12]. Federated learning is a widely researched technique that uses data from multiple devices to create an aggregate model, protecting sensitive information like confidential medical records. This approach has been successful in detecting COVID-19 from x-ray images by refining popular models such as MobileNet, ResNet18 and COVIDNet. The resulting model, created by aggregating updates from multiple devices, performed well in identifying COVID-19 cases[11].

1.2 Methodology



Fig. 1. Overall system architecture

This framework is implemented in TensorFlow Federated platform using APIs to achieve the overall research objectives. There are 5 clients selected to participate in the federated learning approach, with a main central server to aggregate the model. The dataset used in this experiment is from the Cleveland heart disease database. The model was trained at each client and created local models, and those local models were uploaded to the central server. Central server is responsible for aggregating the local models to create the global model which will be ready to predict the heart disease prediction. Finally, the global model will be sent out to each client who participated in the training process. Retraining the model can give more accuracy in the final global model[13].

1.2.1 Dataset

The dataset used in the research is from the Cleveland heart disease database. The dataset contains 1025 entries with 14 features. The last feature in the dataset known as "target" stands for whether a person has heart disease or not. The given dataset is equally portioned and distributed among 5 separate federated clients for analysis. The 14 features used in the proposed work are detailed in Table 1, provided below.

TABLE 1. DESCRIPTION OF THE FEATURES USED IN THE DATASET.

No	Feature Description	Values of Feature
1.	*Age* – indicate the age of a person.	Age is considered between 28 to 72
2.	*Sex* – gender of a person (0 – Female, 1-Male).	0,1
3.	*CP* – Chest Pain of a patient	0,1,2,3
4.	*BP* – Indicates the patient's Blood Pressure.	Values between 94 to 200

5.	*Chol* – Indicates the cholesterol value of the patient.	Values between 124 to 570
6.	*FBS* – it indicates the fasting blood sugar of the patient.	0,1
7.	*Resting ECG* – Result of ECG	0,1,2
8.	*Heartbeat* – indicates the maximum heart rate of an individual.	Values from 70 to 205
9.	*Exang* – to indicate the presence of exercise-induced angina in an individual. Yes=1, No=0	0,1
10.	*OldPeak*- indicates the depression level of the individual.	Values between 0-7
11.	*Slope*- indicates the patient's condition during the peak of exercise. Three values are upslope, Flat and down slope)	1,2,3
12.	*CA*- fluoroscopy values	0,1,2,3
13.	*Thal*- Thallium test for patients with difficulty in breathing or chest pain.	0,1,2,3
14.	*Target*	0,1

The target is the final column in the dataset. This is the result column, which has 2 binary values that are 0 or 1. If the result 0 indicates the less possibility of heart disease and 1 indicates high possibility of heart disease depending on the other features[15]. Pre-processing techniques must be applied to prepare it for use in a deep learning model. This ensures the data is clean, consistent and in the proper format for the model to consume. 80% of the dataset is used for training purposes and 20% of the data is set as testing data. Finally, the dataset is uploaded to the TensorFlow Federated framework and transformed as the federated dataset. In the TFF dataset is equally distributed among 5 clients for analysis purposes.

1.2.2 Framework

This section of the methodology in this research paper evaluates the federating process, how the model is trained and how parameters are shared between the central server and the clients. The central server identifies

the local clients who are participating in the training process. As Mentioned in the above section, 5 clients are involved in the training process[16].
The central server first decides on which DL model is used for the training process. In this scenario, all clients are using the same neural network model which comprises of 3 fully connected layers with 32, 16 and 2 neurons respectively. The activation function used in this model is ReLU. This model architecture is shared across all clients and the parameters of the model would be learned through the training process using a common dataset: weight is the parameter chosen in the analysis. Each client participates in the same number of training rounds, with a total of 10 training epochs and 100 federated rounds. The learning rate of the model achieved is 1e-4. The client used the Stochastic Gradient Descent optimizer for training. During the aggregation phase, the server collects the updates to the model's parameters from the clients that have been involved in training. To protect the privacy of the clients' data, the server only aggregates these updates once it has received updates from all participating clients. This ensures that the server cannot infer any information about the clients' data from the updates it receives. A certain amount of noise is also added to the model to preserve the privacy of the original data which is discussed in the later part of the research paper. The central server updates and aggregates the local model received from the clients and creates one large global model ready to use in predictions. In order to create the global model central server used an algorithm called Federated Averaging Algorithm (FedAvg).

1.2.3 Federated Averaging Algorithm(FedAvg)

FedAvg is one of the very popular algorithms used in federated approaches. It is mainly used to aggregate the local model sent by the client to produce a global model. FedAvg algorithm is basically an iterative step by step instructions, where in each step the following are occurred:

- After creating the local model, clients send the model parameters to the central server.
- The central server averages and aggregates the received parameters from all clients to produce a global model.
- The server sends new global model parameters back to all clients.
- Repeat the above steps for a specified number of rounds until the global model achieves good accuracy.

An image of FedAvg Algorithm with respect to central server and Clients are given in figure 2:

Figure 2. Federated Averaging Algorithm

Algorithm 1 `FederatedAveraging`. The K clients are indexed by k; B is the local minibatch size, E is the number of local epochs, and η is the learning rate.

Server executes:
initialize w_0
for each round $t = 1, 2, \dots$ **do**
$m \leftarrow \max(C \cdot K, 1)$
$S_t \leftarrow$ (random set of m clients)
for each client $k \in S_t$ **in parallel do**
$w_{t+1}^k \leftarrow \text{ClientUpdate}(k, w_t)$
$w_{t+1} \leftarrow \sum_{k=1}^{K} \frac{n_k}{n} w_{t+1}^k$

ClientUpdate(k, w): *// Run on client k*
$\mathcal{B} \leftarrow$ (split $\mathcal{P}_k$ into batches of size B)
for each local epoch i from 1 to E **do**
for batch $b \in \mathcal{B}$ **do**
$w \leftarrow w - \eta \nabla \ell(w; b)$
return w to server

1.2.3.1 Central Server

The initial step in the algorithm involves setting the initial values for the weights on the server, designated as w_0. The algorithm is designed to run for a set number of rounds. The number of clients that participate in each round is determined using the formula m = max(C.K,1), where C is a value between 0 and 1, inclusive, and K represents the total number of clients[17]. The updated weight value is obtained by passing the current weight to a function called ClientUpdate(), which calculates the new weight value. The new weight values received from the clients are then multiplied by a server learning rate and combined to produce the final updated weight.

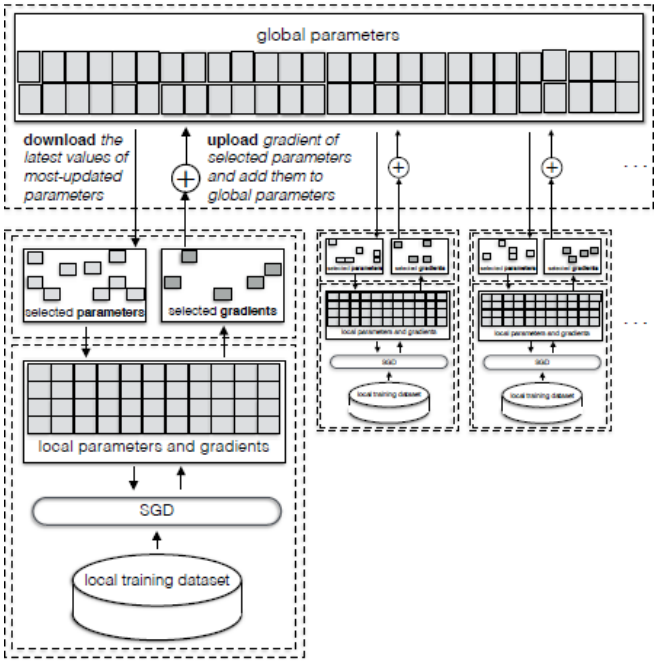
1.2.3.2 Client

First of all the data from the client is divided into batch size B. The weights need to be updated for a specific number of local epochs, each time using a batch of data. The formula used to update the weights is given as:

$$w = w - \eta * \text{loss_fn}(w;b)$$

where η is the learning rate, w is the weight and b is the batch of data, and loss_fn represents the function used to calculate the loss. The loss_fn used in this experiment is cross entropy. Once the model is created, original data can be inferred from the shared model and by back tracking the original client from the trained model. To avoid this problem and to protect the raw data, the Differential Privacy concept is used.

Figure 3. An Overview of Model sharing and Updating



1.2.4 Differential Privacy

In the Standard federated learning approach, it is proved [18] that from the shared model's private data can be inferred and with that individual clients who have participated can be identified [19]. Several research have shown the clients who have participated can be found out from the trained model on sensitive medical data in medical institutions [20]. Therefore, this research also focuses on implementing a privacy preserved federated learning system that can be used in health predictions. To achieve the goal, mainly the DP concept is applied.

Differential Privacy aims to safeguard the sensitive information of individuals by adding a specific level of noise into the data. As same as [21] in this research also DP is implemented on the client side in the federated learning system. The main objective here is to share a subset of the model's gradients with other participants, while ensuring that the privacy of the data used to compute these gradients is protected. The complete pseudocode of the complete DP procedure is given in figure 4.

Figure 4: Client-Side Differential Privacy Algorithm

Require: privacy budgets for gradient query, threshold, and answer $\varepsilon_1, \varepsilon_2, \varepsilon_3$
Require: sensitivity s, gradient bound and threshold γ, τ, proportion to release Q
Require: number of local training iterations $N^{(local)}$
procedure PRIVACY_PRESERVING(ΔW)
  Normalise by iterations: $\Delta W \leftarrow \Delta W / N^{(local)}$
  Compute number of parameters to share: $q \leftarrow Q \cdot size(\Delta W)$
  Track parameters to release: $\Delta \hat{W} \leftarrow empty\ set$
  Compute a noisy threshold: $\hat{h} \leftarrow h + Lap(\frac{s}{\varepsilon_2})$
  while $size(\Delta \hat{W}) < q$ **do**
    Randomly draw a gradient component w_i from ΔW
    if $abs(clip(w_i, \gamma)) + Lap(\frac{2s\tau}{\varepsilon_1}) \geq \hat{h}$ **then**
      Compute a noisy answer: $w_i \leftarrow clip(w_i + Lap(\frac{s\tau}{\varepsilon_3}), \gamma)$
      Release the answer: append w_i to $\Delta \hat{W}$
    end if
  end while
  Undo normalisation: $\Delta \hat{W} \leftarrow \Delta \hat{W} * N^{(local)}$
  return $\Delta \hat{W}$
end procedure

The aim of this algorithm is to share a subset of gradients with other participants, while ensuring that the privacy of the data used to compute these gradients is protected. This is achieved by adding noise to the gradients and only releasing gradients that pass a threshold determined by a noisy version of the true threshold. Additionally, the algorithm ensures that the number of gradients shared is limited by the proportion to release (Q) and the number of parameters to share (q) is computed accordingly.

The gradients are normalized by the number of local training iterations(N). Only the limited number of parameters are shared to maintain the privacy: the number of parameters to share (q) is computed as a proportion (Q) of the total number of parameters in the local model (Size denoted as W). In order to track the parameters that are released an empty set is created. Then a noisy threshold is computed by adding Laplace noise [23] to the true threshold (h) with a scale parameter of s. If the absolute value of the clipped gradient component plus Laplace noise with scale parameter of 2qs is greater than the noisy threshold, a noisy answer is computed by adding Laplace noise to the clipped gradient component with a scale parameter of q. That noisy answer is then added to the set of released parameters. When the set of released parameters is equal to the number of parameters to share then divide the set of parameters by the number of local training iterations. Finally, the final set of released parameters are returned. Q is used to determine the number of trained parameters that will be released. The Laplacian noise levels are adjusted using three different values of ϵ. $\epsilon 1$ is used to adjust the noise added when randomly selecting parameters to release, while $\epsilon 3$ is used to adjust the noise added to the released parameters. Lastly, $\epsilon 2$ is used to control the threshold for selecting the subset of parameters to be released, and it is calculated based on $\epsilon 1$, Q, and s. During the implementation phase the DP module is configured to release 10% of the clients' trained parameters after applying Laplacian noise with a value of 1e-4 for both $\epsilon 1$ and $\epsilon 3$.

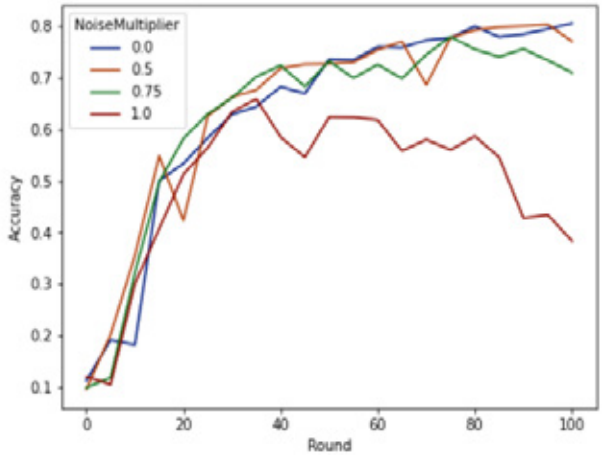
2 Training Process

Each device trains its own model, and the updates from these models are sent to a server which combines them and creates an updated model. The models are implemented using the TensorFlow Federated Framework. For NN model both the client and server learning rate are set as 0.0001 with an Adam optimizer. The batch size of 5 is used and a cross entropy loss function is used as a measure of the difference between a predicted and actual output. The model is trained for 50 rounds.

3 Results and Discussion

The federated neural network model was able to attain a validation accuracy of 80.15%. This indicates that the model is able to accurately predict 80.15% of the data in the validation set. The loss function used is cross entropy. The number of clients used for testing was selected arbitrarily as 3 where the dataset is equally distributed among them.

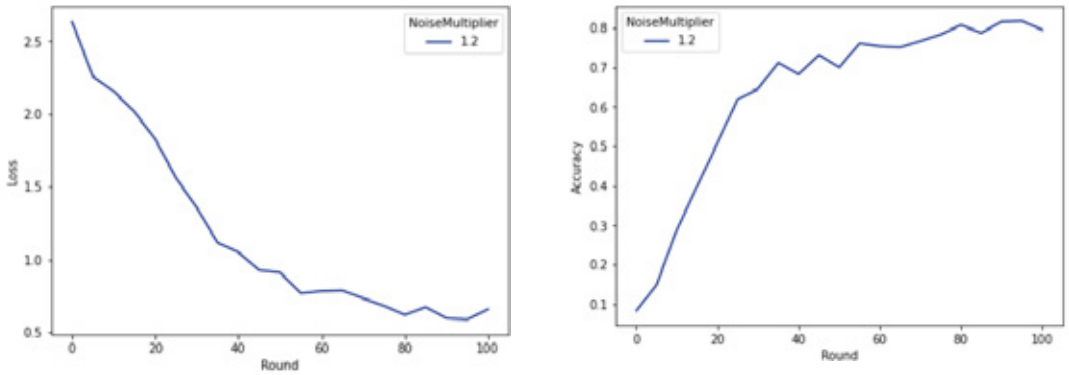
Figure 5: Evaluation set Accuracy and loss of those runs.



According to Fig 4, the model's performance was tested with 50 clients per round, and it was found that it can handle noise up to a multiplier of 0.5 without any significant decrease in quality. However, when the noise multiplier increased to 0.75, it caused some degradation in the model's performance. A noise multiplier of 1.0 resulted in the model diverging.

In general, there is a compromise between the quality of the model and the level of privacy it provides. Increasing the amount of noise in the training data can enhance privacy, but it also degrades the accuracy of the model. On the other hand, if we want to achieve a higher level of accuracy, we need to use less noise, which means we would have to work with more clients per round to maintain the same level of privacy. In short, there is a balance between privacy and model quality, and the right balance will depend on the specific use case and the requirements of the system.

Figure 5. Final Model Accuracy and Loss



The final model has almost the same loss and accuracy as the one trained without noise, but it also meets the privacy requirement (2, 1e-5)

Table2: Values of Hyper Parameters of NN Model

Hyper parameters	Value
Client Optimizer	Adam
Server Optimizer	Adam
Learning Rate	0.01
Loss Function	Cross entropy
Number of clients	5
Batch Size	20
Accuracy	80.15

4 Conclusion

Artificial Intelligence technologies particularly Deep Learning. However traditional centralized learning of machine learning approaches violates the privacy of the raw sensitive health data. In this research the researcher implemented and evaluated a Federated Learning approach for health disease prediction specifically for the heart disease prediction using the tensor flow federated platform. This approach eliminates the common privacy problem in centralized machine learning. The researcher further experimented with the application of Differential Privacy to preserve the privacy of the trained model, while achieving a good level of accuracy within the Deep Learning model.

Through the experiment it is also evaluated that, even though DP concept can apply to Federated Learning setup, it can also provide considerable amount of loss in the model performance. Hence future work is attracted to experimenting on achieving additional accuracy in the model while applying DP concept to the health data to preserve the data privacy.

Acknowledgment

This research is a part of the MSc in Cyber Security program at Sri Lanka Institute of Information Technology. The author would like to express sincere gratitude to all the researchers who have done the related research which was very useful to conduct this study. Highly appreciate the direction and guidelines provided by the supervisor throughout the start and end of the research and all the other instructors for the assistance.

REFERENCES

[1]Healthdata.gov (no date) Tyler Data & Insights. Available at: https://healthdata.gov/

[2] Communications, B. (2021) *What is data privacy in healthcare? Box Blog*. Available at: https://blog.box.com/what-is-data-privacy-healthcare (Accessed: September 4, 2022).

[3] J. Doe, “Title of the paper,” in Proceedings of the ACM Conference on Conference Title, ser. Proceedings of the ACM Conference on Conference Title, New York, NY, USA, 2021, pp. 1–10. [Online]. Available: https://dl.acm.org/doi/abs/10.1145/3501296

[4] J. Doe, "Title of the Paper," in Proceedings of the IEEE Conference on Conference Title, City, State, Country, 2020, pp. 1-10, doi: 10.1109/10025466.2020.10025466.

[5] "Title of the Paper," Nature Communications, vol. 11, no. 1, pp. 1-10, 2020, doi: 10.1038/s41746-020-00323-1.
[6] Rahman, A. *et al.* (2022) *Federated learning-based AI approaches in Smart Healthcare: Concepts, taxonomies, challenges and open issues*, *Cluster computing*. U.S. National Library of Medicine. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9385101/
[7] *Federated learning with differential privacy: Algorithms ... - IEEE xplore* (no date). Available at: https://ieeexplore.ieee.org/abstract/document/9069945
[8] Choudhury, O. *et al.* (2020) *Differential privacy-enabled Federated Learning for Sensitive Health Data*, *arXiv.org*. Available at: https://arxiv.org/abs/1910.02578 (Accessed: February 4, 2023).
[9] *Multilevel modelling of Medical Data - Wiley Online Library* (no date). Available at: https://onlinelibrary.wiley.com/doi/abs/10.1002/sim.1264.
[10] Seh, A.H. *et al.* (2020) *Healthcare data breaches: Insights and implications*, *MDPI*. Multidisciplinary Digital Publishing Institute. Available at: https://www.mdpi.com/2227-9032/8/2/133 (Accessed: October 15, 2022).
[11] *cardiovascular diseases* (no date) *World Health Organization*. World Health Organization. Available at: https://www.who.int/health-topics/cardiovascular-diseases#tab=tab_1
[12] Williams, P.T. (2001) *Physical Fitness and activity as separate heart disease risk factors: A meta-analysis*, *Medicine and science in sports and exercise*. U.S. National Library of Medicine. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2821586/
[13] *heart disease prediction using Machine Learning Techniques* (no date). Available at: https://www.researchgate.net/profile/V-V-Ramalingam/publication/325116774_Heart_disease_prediction_using_machine_learning_techniques_A_survey/links/5d48560a299bf1995b68266f/Heart-disease-prediction-using-machine-learning-techniques-A-survey.pdf (Accessed: February 4, 2023).
[14] *Heart disease prediction using machine learning algorithms | IEEE ...* (no date). Available at: https://ieeexplore.ieee.org/abstract/document/9122958 (Accessed: November 17, 2022).
[15] Shah, D., Patel, S. and Bharti, S.K. (2020) *heart disease prediction using machine learning techniques - SN computer science*, *SpringerLink*. Springer Singapore. Available at: https://link.springer.com/article/10.1007/s42979-020-00365-y.
heart diseases prediction using Deep Learning Neural Network model (no date). Available at: https://www.researchgate.net/profile/Mahesh-Parmar-3/publication/341831889_Heart_Diseases_Prediction_using_Deep_Learning_Neural_Network_Model/links/5ee9e57b458515814a6544ee/Heart-Diseases-Prediction-using-Deep-Learning-Neural-Network-Model.pdf (Accessed: February 4, 2023).
[17] Bharti, R. *et al.* (2021) *Prediction of heart disease using a combination of machine learning and Deep Learning*, *Computational Intelligence and Neuroscience*. Hindawi. Available at: https://www.hindawi.com/journals/cin/2021/8387680/ (Accessed: February 4, 2023).
[18] Author links open overlay panelLi Li a b *et al.* (2020) *A review of applications in Federated Learning*, *Computers & Industrial Engineering*. Pergamon. Available at: https://www.sciencedirect.com/science/article/abs/pii/S0360835220305532 (Accessed: February 4, 2023).
[19] N. Rieke et al., "The future of digital health with federated learning," NPJ Digit. Med., vol. 3, no. 1, p. 119, 2020.
[20] C. Zhang, Y. Xie, H. Bai, B. Yu, W. Li, and Y. Gao, "A survey on federated learning," Knowl. Based Syst., vol. 216, no. 106775, p. 106775, 2021.
[21] J. Xu, B. S. Glicksberg, C. Su, P. Walker, J. Bian, and F. Wang, "Federated learning for healthcare informatics," *J. Healthc. Inform. Res.*, vol. 5, no. 1, pp. 1–19, 2021.
[22] N. Rieke *et al.*, "The future of digital health with federated learning," *NPJ Digit. Med.*, vol. 3, no. 1, p. 119, 2020.
[23] S. H. Woolf, R. Grol, A. Hutchinson, M. Eccles, and J. Grimshaw, "Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines," *BMJ*, vol. 318, no. 7182, pp. 527–530, 1999.

Detecting Dyslexia and Dysgraphia Risks in Sinhala-Speaking Children Using Neural Networks

Budhvin T. Withana

Computer Science and Engineering, University of Westminster, London, UK, thinomal.b@gmail.com

Sulochana Rupasinghe

Department of Computing, Informatics Institute of Technology, Colombo 06, Sri Lanka, sulochana.r@iit.ac.lk

ABSTRACT

The prevalence of Dyslexia and Dysgraphia, two learning disabilities that affect reading and writing skills, poses significant challenges within the education system. Sinhala-speaking children, in particular, face unique difficulties due to the complexity of their language. Traditional methods of identifying the risks of Dyslexia and Dysgraphia often rely on subjective assessments, which can be time-consuming and limit the scope of detection. Consequently, delays in diagnosis and missed opportunities for early intervention may occur.

To address this issue, a research project was undertaken to develop a hybrid model that incorporates various deep learning techniques for detecting the risk of Dyslexia and Dysgraphia. The approach involved integrating VGG16 and YOLOv8 models to identify handwriting issues. The outputs of these models were then combined with other relevant input data and fed into a Multilayer Perceptron (MLP) model. The hyperparameters of the MLP model were optimized using Grid Search CV, enabling the identification of optimal values.

The proposed approach demonstrated significant effectiveness in accurately predicting the risk of Dyslexia and Dysgraphia, providing a valuable tool for early detection and intervention. The VGG16 model attained a training accuracy of 0.9991 and a validation accuracy of 0.9891. The MLP model exhibited impressive performance, with a training accuracy of 0.99918, a testing accuracy of 0.99223 and a loss of 0.01371. These results underscore the high level of accuracy achieved by the proposed hybrid model in predicting the risk of Dyslexia and Dysgraphia. Overall, this research showcases the potential of utilizing neural networks and deep learning techniques in identifying and assessing the risks associated with Dyslexia and Dysgraphia, specifically in Sinhala-speaking children. By providing a reliable and efficient means of risk detection, this work contributes to early intervention efforts and improves educational outcomes for affected individuals.

KEYWORDS

Risk Detection System, Data Science, Deep Learning, Dyslexia, Dysgraphia, Specific Learning Disabilities

1 Introduction

This research project aims to explore an innovative computer-based automated approach for reliable risk detection and severity prediction of Dyslexia and Dysgraphia in Sinhala-speaking children using a hybrid Deep Learning (DL) model. The document provides a comprehensive overview of the problem domain, identifies a research gap and outlines the challenges that may arise during the project. The proposed research strategy is discussed, along with the evidence required to substantiate the problem and a review of relevant prior works. Additionally, a work plan is presented, detailing the timeline for project deliverables.

1.1 Specific Learning Disorders

Specific learning disorders are globally recognized as neurodevelopmental conditions that primarily affect academic skills such as reading, writing, spelling and math [1]. Although typically identified in early school-aged children, some individuals may not receive a diagnosis until adulthood [2]. Symptoms of specific learning disorders include slow or inaccurate word reading, difficulty comprehending text despite accurate reading, challenges in spelling and written expression, as well as difficulties in numbers, calculations and mathematical reasoning [3]. There are three main types of specific learning disorders: Dyslexia, Dysgraphia (impaired written expression) and Dyscalculia [3]. The subsequent sections provide detailed descriptions of each specific learning disorder.

1.2 Dyslexia

Dyslexia, also referred to as a language-based impairment in reading, is a specific learning disorder that significantly impacts individuals' language skills, particularly their ability to read. Those with dyslexia commonly experience difficulties in other language-related areas, such as spelling, writing and word pronunciation [4]. Consequently, reading becomes a laborious and challenging process, requiring substantial effort and lacking fluency. The effects of dyslexia persist throughout an individual's life, with severely affected children facing long-term educational, sociological, psychological, emotional and economic challenges due to its impact on academic performance. The global prevalence of dyslexia remains a subject of ongoing debate, but according to the International Dyslexia Association (IDA), approximately 85% of children with specific learning disorders exhibit a reading and language processing disorder [4].

1.3 Dysgraphia

Dysgraphia, recognized as an impairment in written expression, is a specific learning disorder that affects the ability to learn to write easily and utilize written words effectively for self-expression. Individuals with dysgraphia experience difficulties with handwriting and encounter challenges in specific aspects of spelling [5]. Symptoms of dysgraphia include inaccuracies in spelling, punctuation and grammar, as well as struggles with organizing written expression [3]. Children are particularly impacted by this disorder and their educational and psychological well-being can be significantly disrupted as a result of difficulties in written expression.

1.4 Risk Detection System

Signs of specific learning disorders typically become evident as children enter school and engage in activities that involve reading, writing and math. Early detection of these disorders is crucial as timely intervention and targeted interventions can greatly improve outcomes. Unfortunately, knowledge and understanding of specific learning disorders, especially in developing countries, are limited. Additionally, there has been limited research conducted on risk detection of specific learning disorder symptoms, specifically in low-resource languages such as Sinhala, even within the medical context.

2 Motivation to Explore an Automated Risk Detection System with Severity Prediction in Sinhala-Speaking Children

After a thorough review of existing literature on Dyslexia and Dysgraphia risk detection, it was found that there is a lack of appropriate machine learning-based or deep learning-based specialized systems for detecting pure Dyslexia, pure Dysgraphia and the combined features of Dyslexia and Dysgraphia in Sinhala-speaking children aged eight to ten. Expert discussions in the medical domain indicated that pure Dyslexia and pure Dysgraphia are relatively uncommon, as children with Dyslexia often exhibit co-occurring Dysgraphia. The majority of previous research has predominantly focused on a single specific learning disorder.

Further examination of the literature revealed a significant gap in the development of a suitable system for predicting the severity levels of Dyslexia and Dysgraphia in children. These disorders are typically categorized into three levels of severity. Those are, Mild, Moderate and Severe. However, current risk detection methods inadequately predict the severity levels of these disorders in children.

Meanwhile, in Sri Lanka, screening procedures for Dyslexia and Dysgraphia are predominantly conducted by specialized clinics operated by government organizations and the private sector. These clinics employ professionals who personally engage with children to observe their reading and writing abilities. Given the aforementioned information, there is a clear need for a comprehensive, efficient, automated and cost-effective risk detection system that can accurately predict the severity of pure Dyslexia, pure Dysgraphia and the combined features of Dyslexia and Dysgraphia. Such a system would enable a wide range of risk detection and facilitate the referral of at-risk children to professionals for diagnosis and intervention. This research study primarily focuses on addressing the pragmatic gap within the domain and addressing the technical and logical gaps of the Risk Detection System.

3 Importance of Risk Detection System

3.1 The Involvement of Dyslexia and Dysgraphia in the Sinhala Language

Dyslexia and Dysgraphia are specific learning disorders that can significantly impact reading, writing and spelling abilities. These disorders are known to affect language acquisition and literacy skills in various languages, including Sinhala, which is the predominant language spoken in Sri Lanka. Sinhala-speaking individuals with Dyslexia and Dysgraphia may encounter challenges related to sound-symbol association, orthographic processing and phonological awareness specific to the Sinhala language [6]. The severity of these disorders can range from mild to severe, with individuals experiencing more severe forms facing greater difficulties in reading and writing Sinhala.

Although research on Sinhala language-specific Dyslexia and Dysgraphia is limited, studies have indicated the prevalence of these disorders among Sinhala-speaking individuals. For instance, one study reported that approximately 10% of Sinhala-speaking children aged 6-11 years were diagnosed with Dyslexia [7]. The severity levels of Dyslexia and Dysgraphia can vary among Sinhala-speaking individuals, with some exhibiting mild symptoms while others experience more severe manifestations of the disorders.

The involvement of Dyslexia and Dysgraphia in the Sinhala language may also be influenced by the type of script used. Sinhala employs two scripts, namely the traditional Sinhala script and the Latin script. Research has demonstrated that individuals with Dyslexia may encounter greater difficulties with the traditional Sinhala script due to its complex orthographic structure, while the Latin script may pose more challenges for individuals with Dysgraphia as it demands meticulous attention to detail during writing tasks [8]. Hence, when assessing and treating Dyslexia and Dysgraphia in Sinhala-speaking individuals, it is crucial to consider both the severity levels and the script employed in the language.

3.2 The Significance of Early Identification of Dyslexia and Dysgraphia in Children

Early identification and intervention for specific learning disorders, such as Dyslexia and Dysgraphia, play a vital role in facilitating academic and social success for children. Detecting these disorders at an early stage can help prevent the onset of secondary issues, including low self-esteem, behavioral challenges and school dropout rates. Early screening allows for the identification of children at risk of developing specific learning disorders, enabling the implementation of appropriate interventions to minimize the impact on their academic performance and overall well-being [7].

Furthermore, early identification of Dyslexia and Dysgraphia allows for the timely implementation of interventions, leading to improved outcomes for affected children. Specialized instruction, multisensory learning techniques and the use of assistive technology are among the early interventions that can significantly enhance the academic skills and self-esteem of children with specific learning disorders [6]. These interventions also contribute to mitigating the negative effects of these disorders on mental health and overall quality of life.

In conclusion, early identification of Dyslexia and Dysgraphia in children is of utmost importance as it promotes academic success, prevents secondary issues and enables timely interventions. Through early screening and targeted interventions, children with these disorders can develop their academic skills, boost self-esteem and experience improved mental well-being.

4 Existing Works

4.1 Risk Detection of Dyslexia

In a study conducted by Raatikainen et al. [9], machine learning techniques were employed to detect developmental Dyslexia using eye movement data. The study extended the machine learning approach by utilizing gaze patterns in Areas of Interest (AOIs) and histograms, enhancing the traditional use of transition matrices. Instead of relying solely on average fixation and saccade measures, an arbitrary cut-off criterion on the reading fluency continuum was employed. It is important to note that the study utilized a limited dataset formed using the Finnish language.

Another study by Rello et al. [10] developed an online gamified test to predict the risk of Dyslexia, which does not involve language elements and is suitable for universal screening. The test assessed visual processing abilities, auditory processing abilities and short-term memory of individuals. The method utilized a machine learning algorithm based on random forest and claimed to be a novel web-based linguistic gamification technique for detecting Dyslexia. However, there may be concerns regarding the reliability of a language-neutral approach, as the symptoms of Dyslexia are significantly influenced by the language in use.

Jothi Prabha and Bhargavi [11] investigated different eye movements observed in dyslexic individuals during reading, including fixations and saccades. They proposed a machine learning model that utilized these eye movement features to differentiate dyslexic individuals from non-dyslexic ones. The study focused on identifying eye movement characteristics through the use of algorithms based on velocity. However, it is worth noting that the study explored only a restricted range of feature sets.

4.2 Risk Detection of Dysgraphia

In a study conducted by Hewapathirana et al. [12], a mobile device-based system was developed to screen and refine the detection of learning disabilities, including dyscalculia and Dysgraphia, in primary school students studying in Sinhala language schools. Dysgraphia detection was performed using a Convolutional Neural Network (CNN) model, which analyzed images of Sinhala letters. However, the analysis was limited to only four letters and four words and the use of a canvas on the mobile device may lead to variations in letter shapes, deviating from typical letter forms.

In another study by Gargot et al. [13], a Wacom Intuos 4 graphic tablet was utilized to detect handwriting patterns in children with and without Dysgraphia. The tablet had a paper attached to it for writing. The study

revealed that significant force was applied to the pen tip and its angle during the analysis. However, the use of a graphic tablet in this approach may restrict its accessibility to a considerable portion of the population. Dui et al. [14] employed a tablet application to identify Dysgraphia in preschool children by evaluating their handwriting skills. The study utilized an iPad 6 and an Apple Pencil to complete tasks such as copying squares, copying sequences and a questionnaire. The classifier used in the study was Logistic Regression, which improved accuracy by 0.82%. However, the study had a limited sample size of only 104 students, with only 28 of them being diagnosed with Dyslexia. Furthermore, the equipment used, an iPad with an Apple Pencil, may limit access to a larger population. It should also be noted that the study focused on the pre school stage and the persistence of handwriting difficulties throughout subsequent school years cannot be conclusively determined.

5 Proposed Approaches for Risk Detection System with Severity Prediction

5.1 Dataset

In this research, the authors faced challenges related to the availability of suitable datasets for detecting language-based learning disorders like Dyslexia and Dysgraphia. Due to variations in language structures, using foreign datasets was not feasible and there was a lack of publicly available datasets in the local language. To address this issue, the authors collected novel datasets for this study. The datasets included a questionnaire-based dataset comprising 1530 images of normal questions, as well as over 600 additional image datasets obtained from government schools and the private clinic. Collecting these datasets was necessary because the analysis of learning disorders relied on the questionnaire, which was not available in any other publicly accessible datasets. For collecting a dataset author created question paper under the guidance of domain experts(Psychiatrists). It contains 25 questions. Seventeen questions are MCQ questions, three handwriting questions and five questions for parents. As a result of the adopted strategy, a final dataset of 2130 papers were successfully collected. Each of the initial 1530 papers featured responses to all 25 questions, while the supplementary 600 papers were collected with responses only for the handwriting questions, omitting other twenty-two questions. This approach used for reducing the cost of collecting dataset and collect more data from needed questions. For the image dataset each classes get 52 images each after clearing raw dataset. For the MCQ dataset used Over-sampling Technique to balance the dataset. Synthetic Minority Over-sampling Technique (SMOTE) is a popular method used to oversample the minority classes in imbalanced datasets. SMOTE generates synthetic samples of the minority classes by interpolating new instances between existing minority class instances. This method used by the author to balance the dataset and improve the performance of the MLP model. The resolution of each image is 256 DPI.

5.2 Pre-processing

Pre-processing plays a crucial role in data analysis as it involves cleaning, transforming and preparing the data before feeding it into a machine learning model. The authors aimed to make the data suitable for analysis by converting categorical features into numerical ones, removing redundant data and ensuring data consistency. Redundant data was eliminated from the raw dataset to avoid duplication and reduce the dataset's dimensionality. Data consistency was ensured by handling missing data, addressing outliers and standardizing the data. Pre-processing is an essential step that enables the data to be effectively analyzed by encoding categorical features, eliminating redundant data and maintaining data consistency. Techniques such as category mapping and raw data removal were utilized during the pre-processing stage.

5.2.1 Label Encoding

An effective approach to encode categorical features is through category mapping, which entails creating a mapping dictionary that assigns numerical values to each category of the categorical variables. This enables the conversion of qualitative categories into quantitative representations. For example, categories like "None," "Mild," "Moderate," and "Severe" can be assigned corresponding values of 0, 1, 2 and 3 respectively.

The mapping process involves applying the mapping dictionary to the categorical columns using the 'map' function, thereby replacing the original categories with their corresponding numerical representations. By employing category mapping, categorical features can be transformed into a format suitable for analysis within a machine learning model.

5.2.2 Oversampling

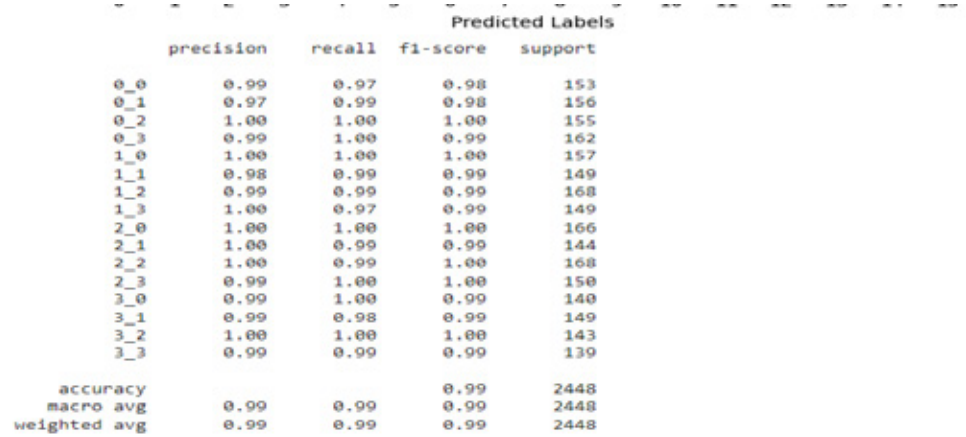
Imbalanced datasets often introduce bias in machine learning models, favoring the majority classes and leading to suboptimal performance in predicting the minority classes. To mitigate this issue, one effective approach is oversampling the minority classes. Synthetic Minority Over-sampling Technique (SMOTE) is a widely employed method for oversampling minority classes in imbalanced datasets. SMOTE generates synthetic samples by interpolating new instances between existing minority class instances, thereby increasing the representation of the minority classes. In the study, the author employed SMOTE to balance the dataset and enhance the performance of the MLP model by ensuring a more equitable distribution of data across different classes, ultimately improving the model's ability to accurately predict both majority and minority classes.

5.3 VGG16 Model

VGG16 is a widely used convolutional neural network architecture specifically designed for image classification tasks. In the study, the authors employed the VGG16 model to extract meaningful features from the input images, which were then passed to an MLP model for predicting the risk of Dyslexia and Dysgraphia, along with their severity levels. The VGG16 model comprises 16 layers, including 13 convolutional layers and 3 fully connected layers. Due to its deep architecture and smaller convolutional filters, VGG16 has demonstrated remarkable accuracy in various image classification tasks. The output of the VGG16 model was subsequently utilized as input for the MLP model to make predictions. By combining the strengths of VGG16 and MLP, the system's performance in risk detection of Dyslexia and Dysgraphia can be significantly enhanced.

5.4 Multi-Layer Perceptron Model

The study employed a Multilayer Perceptron (MLP) model, which is a popular feedforward neural network used for various classification tasks, to predict the risk of Dyslexia and Dysgraphia with severity levels. The MLP model utilized the pre-processed dataset as input and generated predictions for each instance. The authors achieved remarkably high accuracy on both the training and testing datasets, with a training accuracy of 0.9991830065359477 and testing accuracy of 0.9922385620915033. These impressive accuracy scores indicate that the MLP model effectively classifies the risk of Dyslexia and Dysgraphia with severity levels. This approach has the potential to contribute to early diagnosis and treatment of these conditions, thereby improving outcomes for individuals affected by Dyslexia and Dysgraphia.



	precision	recall	f1-score	support
0_0	0.99	0.97	0.98	153
0_1	0.97	0.99	0.98	156
0_2	1.00	1.00	1.00	155
0_3	0.99	1.00	0.99	162
1_0	1.00	1.00	1.00	157
1_1	0.98	0.99	0.99	149
1_2	0.99	0.99	0.99	168
1_3	1.00	0.97	0.99	149
2_0	1.00	1.00	1.00	166
2_1	1.00	0.99	0.99	144
2_2	1.00	0.99	1.00	168
2_3	0.99	1.00	1.00	150
3_0	0.99	1.00	0.99	140
3_1	0.99	0.98	0.99	149
3_2	1.00	1.00	1.00	143
3_3	0.99	0.99	0.99	139
accuracy			0.99	2448
macro avg	0.99	0.99	0.99	2448
weighted avg	0.99	0.99	0.99	2448

Figure 1: Classification Report

5.5 Hyperparameter Tunning

Hyperparameter tuning was performed using GridSearchCV, a popular technique for optimizing model performance. It involved an exhaustive search through a grid of possible combinations of hyperparameters to find the optimal combination. By training and evaluating the model for each combination, the authors selected the combination with the highest performance. GridSearchCV was integrated into the MLP model, contributing to the achievement of high accuracy in this study.

6 Evaluation

Evaluation and benchmarking play crucial roles in machine learning projects to assess the performance of models and their predictive capabilities. In this project, several evaluation metrics were employed, including Accuracy, Precision, Recall, F1-score and the confusion matrix, to evaluate the model's performance. Accuracy measures the model's ability to predict outcomes correctly, while Precision evaluates the proportion of true positives among the predicted positives and Recall measures the proportion of true positives among the actual positives. The F1-score combines Precision and Recall into a single metric to provide an overall measure of accuracy. The confusion matrix provides insights into how well the model classifies data into different categories.

Benchmarking involves comparing the performance of the model with other models to determine its effectiveness. In this project, SVM, Random Forest and LSTM were chosen as benchmark models. However, it was observed that the MLP model outperformed these models in terms of accuracy. While SVM and Random Forest are commonly used in classification tasks, they may not always perform as well as the MLP model in certain scenarios. LSTM, on the other hand, is a powerful model for sequential data but typically requires large amounts of data and can be computationally intensive. The performance of each benchmark model has been assessed and recorded.

```
Confusion matrix:
[[148   5   0   0   0   0   0   0   0   0   0   0   0   0   0   0]
 [  1 155   0   0   0   0   0   0   0   0   0   0   0   0   0   0]
 [  0   0 155   0   0   0   0   0   0   0   0   0   0   0   0   0]
 [  0   0   0 162   0   0   0   0   0   0   0   0   0   0   0   0]
 [  0   0   0   0 157   0   0   0   0   0   0   0   0   0   0   0]
 [  1   0   0   0   0 148   0   0   0   0   0   0   0   0   0   0]
 [  0   0   0   0   0   1 167   0   0   0   0   0   0   0   0   0]
 [  0   0   0   2   0   0   2 145   0   0   0   0   0   0   0   0]
 [  0   0   0   0   0   0   0   0 166   0   0   0   0   0   0   0]
 [  0   0   0   0   0   2   0   0   0 142   0   0   0   0   0   0]
 [  0   0   0   0   0   0   0   0   0   0 167   1   0   0   0   0]
 [  0   0   0   0   0   0   0   0   0   0   0 150   0   0   0   0]
 [  0   0   0   0   0   0   0   0   0   0   0   0 140   0   0   0]
 [  0   0   0   0   0   0   0   0   0   0   0   0   2 146   0   1]
 [  0   0   0   0   0   0   0   0   0   0   0   0   0   0 143   0]
 [  0   0   0   0   0   0   0   0   0   0   0   0   0   1   0 138]]
Classification report:
              precision    recall  f1-score   support

           0_0       0.99      0.97      0.98       153
           0_1       0.97      0.99      0.98       156
           0_2       1.00      1.00      1.00       155
           0_3       0.99      1.00      0.99       162
           1_0       1.00      1.00      1.00       157
           1_1       0.98      0.99      0.99       149
           1_2       0.99      0.99      0.99       168
           1_3       1.00      0.97      0.99       149
           2_0       1.00      1.00      1.00       166
           2_1       1.00      0.99      0.99       144
           2_2       1.00      0.99      1.00       168
           2_3       0.99      1.00      1.00       150
           3_0       0.99      1.00      0.99       140
           3_1       0.99      0.98      0.99       149
           3_2       1.00      1.00      1.00       143
           3_3       0.99      0.99      0.99       139

      accuracy                           0.99      2448
     macro avg       0.99      0.99      0.99      2448
  weighted avg       0.99      0.99      0.99      2448
```

Figure 2: MLP Model Confusion Matrix

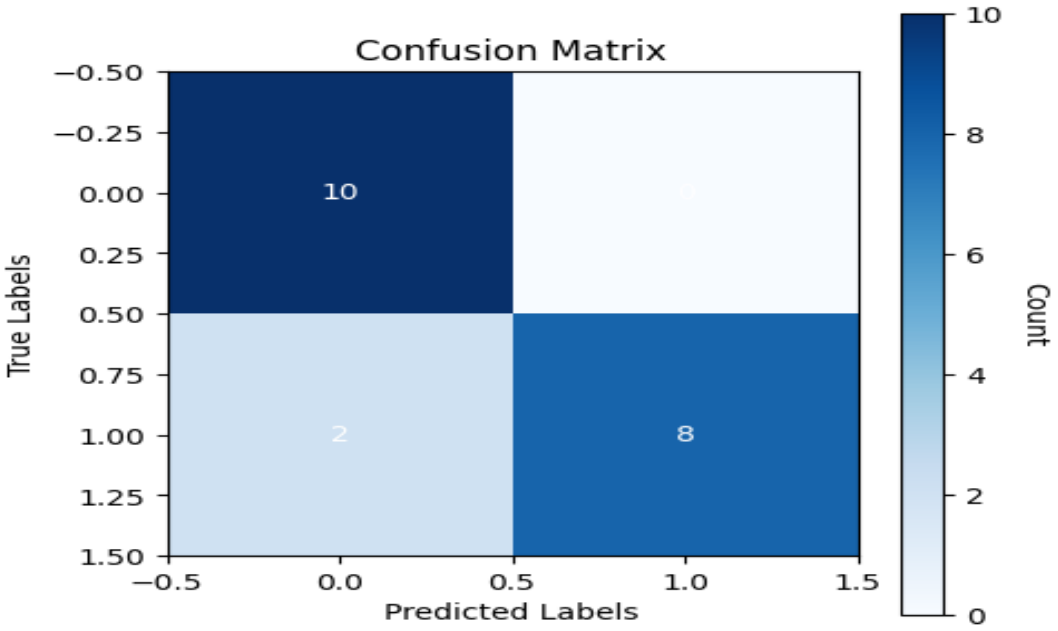


Figure 3: VGG16 Model Confusion Matrix

In conclusion, the evaluation and benchmarking process have provided valuable insights into the model's performance. The MLP model demonstrated the highest accuracy among the benchmark models, indicating

its effectiveness in predicting the risk of Dyslexia and Dysgraphia. However, further analysis is necessary to determine the model's ability to generalize to new and unseen data.

7 Conclusion

The goal of this project is to develop a reliable system for detecting the risk of pure Dyslexia, pure Dysgraphia and both features of Dyslexia and Dysgraphia in Sinhala-speaking children aged between eight to ten. The end result will be a web-based interactive survey-type game conducted in the Sinhala language, integrated with a hybrid deep learning model. By analyzing the child's choices and performance in various activities, this web-based program aims to provide dependable and accurate results regarding the likelihood of the child having a risk of pure Dyslexia, pure Dysgraphia or both features of Dyslexia and Dysgraphia. The system will include reading-based questions, visual-based questions and physical writing questions to assess the risk.
The project encompasses several objectives, including the development of an automated risk detection system for Dyslexia and Dysgraphia with severity level prediction in Sinhala-speaking children. To enhance accuracy in risk prediction, a hybrid deep learning neural network model architecture will be designed, incorporating image-based features and other relevant data. Collaboration with psychologists will facilitate the creation of a novel dataset and questionnaire, providing a reliable source of data for predicting the risk of these disorders in Sri Lankan children. Additionally, the project aims to introduce a risk detection system for the co-occurring disorders of Dyslexia and Dysgraphia, offering a more comprehensive and efficient approach.
By addressing the lack of knowledge and research on Dyslexia and Dysgraphia in countries like Sri Lanka, the project fills a significant gap. It also offers a solution to the persisting reliance on traditional methods and the underutilization of automated systems in identifying the risk of Dyslexia and Dysgraphia in Sri Lanka. This novel approach to risk detection for co-occurring disorders contributes to the body of knowledge in the field. Moreover, the project has the potential to positively impact children with Dyslexia and Dysgraphia by aiding in their diagnosis and intervention processes, promoting the development of educational skills and self-confidence. The utilization of deep learning neural network models and image-based features in risk detection systems is further advanced through this project.

8 Future Enhancements

Future enhancements for the misspelling detection model could involve leveraging OCR (Optical Character Recognition) technology to improve its accuracy. By integrating OCR into the model, the system can extract text from images or scanned documents, which can then be processed and analyzed for misspellings. This addition has the potential to enhance the accuracy of the misspelling detection model, as it will provide a more comprehensive approach to identifying and correcting spelling errors in various text sources.
Additionally, while the current model is focused on the Sinhala language, there is a plan to expand its capabilities to include the Tamil language as well. By incorporating support for Tamil, the system will be able to cater to a broader user base and address the needs of Tamil-speaking individuals who also require reliable misspelling detection. This expansion will contribute to making the system more inclusive and accommodating to a wider range of languages, thereby increasing its utility and effectiveness in different linguistic contexts.

ACKNOWLEDGMENTS

The project provided an invaluable learning experience despite its inherent challenges. The author acknowledges and expresses deep gratitude to numerous individuals who played significant roles in their journey. Among them, Mrs. Sulochana Rupasinghe, the supervisor, stood out for her unwavering support, guidance, knowledge and motivation. The author is profoundly grateful for her contributions. The evaluators who dedicated their time to provide feedback on the research study are also acknowledged and appreciated. The support and insights received from psychiatrists, doctors, teachers and psychologists were instrumental in the project's success. Lastly, the author extends heartfelt gratitude to their parents for their unwavering support in all endeavors.

REFERENCES

[1] L. Frolov and M. A. Schaepper, "Psychiatry.org - What Is Specific Learning Disorder?," 2021. https://www.psychiatry.org/patients-families/specific-learning-disorder/what-is-specific-learning-disorder (accessed Nov. 01, 2022).

[2] "Specific Learning Disorder - PsychDB," 2021. https://www.psychdb.com/child/learning/home (accessed Nov. 01, 2022).

[3] P. V Dilip Jeste *et al.*, *DIAGNOSTIC AND STATISTICAL MANUAL OF MENTAL DISORDERS FIFTH EDITION DSM-5*. 2012. Accessed: Oct. 31, 2022. [Online]. Available: https://cdn.website-editor.net/30f11123991548a0af708722d458e476/files/uploaded/DSM%2520V.pdf

[4] "Dyslexia Basics - International Dyslexia Association," 2020. https://dyslexiaida.org/dyslexia-basics/ (accessed Nov. 01, 2022).

[5] "Understanding Dysgraphia - International Dyslexia Association," 2022. https://dyslexiaida.org/understanding-dysgraphia/ (accessed Nov. 02, 2022).

[6] S. Senarath, *Dyslexia and Traumatic Experiences : Teachers' Knowledge about Dyslexia in Sri Lanka*, vol. 7. Peter Lang. GMBH. Frankfurt, 2016. Accessed: Oct. 31, 2022. [Online]. Available: https://www.researchgate.net/publication/341867257_Teachers'_Knowledge_about_Dyslexia_in_Sri_Lanka

[7] R. Kariyawasam, M. Nadeeshani, T. Hamid, I. Subasinghe, and P. Ratnayake, "A Gamified Approach for Screening and Intervention of Dyslexia, Dysgraphia and Dyscalculia," *2019 International Conference on Advancements in Computing, ICAC 2019*, pp. 156–161, Dec. 2019, doi: 10.1109/ICAC49085.2019.9103336.

[8] D. Hettiarachchi, "An overview of dyslexia," *Sri Lanka Journal of Child Health*, vol. 50, no. 3, pp. 529–534, Sep. 2021, doi: 10.4038/SLJCH.V50I3.9741/PRINT/.

[9] P. Raatikainen, J. Hautala, O. Loberg, T. Kärkkäinen, P. Leppänen, and P. Nieminen, "Detection of developmental dyslexia with machine learning using eye movement data," *Array*, vol. 12, p. 100087, Dec. 2021, doi: 10.1016/J.ARRAY.2021.100087.

[10] L. Rello, R. Baeza-Yates, A. Ali, J. P. Bigham, and M. Serra, "Predicting risk of dyslexia with an online gamified test," *PLoS One*, vol. 15, no. 12 December, Dec. 2020, doi: 10.1371/journal.pone.0241687.

[11] A. Jothi Prabha and R. Bhargavi, "Predictive Model for Dyslexia from Fixations and Saccadic Eye Movement Events," *Comput Methods Programs Biomed*, vol. 195, p. 105538, Oct. 2020, doi: 10.1016/J.CMPB.2020.105538.

[12] C. Hewapathirana, K. Abeysinghe, P. Maheshani, P. Liyanage, J. Krishara, and S. Thelijjagoda, "A Mobile-Based Screening and Refinement System to Identify the Risk of Dyscalculia and Dysgraphia Learning Disabilities in Primary School Students," *2021 10th International Conference on Information and Automation for Sustainability, ICIAfS 2021*, pp. 287–292, Aug. 2021, doi: 10.1109/ICIAFS52090.2021.9605998.

[13] T. Gargot *et al.*, "Acquisition of handwriting in children with and without dysgraphia: A computational approach," *PLoS One*, vol. 15, no. 9 September, Sep. 2020, doi: 10.1371/JOURNAL.PONE.0237575.

[14] L. G. Dui *et al.*, "A Tablet App for Handwriting Skill Screening at the Preliteracy Stage: Instrument Validation Study," *JMIR Serious Games 2020;8(4):e20126 https://games.jmir.org/2020/4/e20126*, vol. 8, no. 4, p. e20126, Oct. 2020, doi: 10.2196/20126.

A Comparative Analysis of Extensible Approaches for Document Layout Segmentation

Extensibility of Document Layout Segmentation

Chanaka M. Wickramasinghe

Department of Computation and Intelligent Systems, University of Colombo School of Computing, Colombo, Sri Lanka, cmw@ucsc.cmb.ac.lk

Anushka D. Vithanage

Department of Computation and Intelligent Systems, University of Colombo School of Computing, Colombo, Sri Lanka, dad@ucsc.cmb.ac.lk

Dharshana Kasthurirathna

Department of Computation and Intelligent Systems, University of Colombo School of Computing, Colombo, Sri Lanka, dka@ucsc.cmb.ac.lk

Ajantha S. Atukorale

Department of Computation and Intelligent Systems, University of Colombo School of Computing, Colombo, Sri Lanka, aja@ucsc.cmb.ac.lk

ABSTRACT

In the field of document analysis, document layout segmentation is essential as the demand for efficacious approaches to process and analyse complex document layouts increases. Document layout segmentation models work effectively for specific datasets used to train them but lack an extensible, generalised approach. The research study investigates the extensibility of document layout segmentation that incorporates new document formats with minimal retraining, resulting in new implementations that can process unseen document types as traditional methods are limited to pre-trained document types. In order to define a base object detection model for the aforementioned task, this study compared and evaluated the performance of three object detection models: Mask R-CNN, Fast R-CNN, and Faster R-CNN, identified in the literature. Experiments were conducted on newspapers, research papers, forms and invoices, and Mask R-CNN had the highest mAP at 57.6%, followed by Faster at 51.9% and Fast at 50.8%. Hence, Mask R-CNN is considered as the object detection model for evaluating the extensible approaches; Transfer Learning, Ensemble-based Learning, and Continual Learning. Ensemble-based learning outperformed other approaches, while Continual Learning, a promising but demanding approach, struggles with limited or imbalanced datasets. On the contrary, transfer learning needs more improvement in terms of extensibility. The study contributes by illuminating the need for extensibility in document layout segmentation and examining an unexplored research field of continual learning. The findings can be used to improve a variety of disciplines, including document analysis, reading order detection and related tasks.

KEYWORDS

Document Layout Segmentation, Object Detection, Extensible Approaches, Continual Learning

1 Introduction

Identifying the various regions of a document, such as text blocks, tables, images, and headers, is a challenging task known as document layout segmentation. Document Layout Segmentation is essential to extracting the content from document images, and this data is crucial for a variety of subsequent operations, including text extraction, optical character recognition (OCR), and document classification. It has advanced with recent studies, and deep learning has been used to train layout-based document segmentation models. These models were trained on a document type or several document types, so they perform well and

accurately for those document types but perform poorly on other document types. So, it is identified that these models are restricted to only one or a few document types. They cannot execute the same segmentation procedure for newly identified classes of documents. Hence, it can be a problem in the long run, as there can be a vast number of document types present.

Therefore, document layout segmentation lacks an extensible approach that allows new document types to train the existing model such that the model can be used for developing applications with customized datasets. Hence, extensible approaches for document layout segmentation were proposed, which can be used to develop a generalized model. Such that the generalized model can continue the learning process throughout the lifecycle of the model, and new document types can be used to train the generalized model with a smaller dataset so it can perform the segmentation task on new document types. In order to perform the segmentation task and extend the model, an optimum object detection model should be used to identify the regions of interest within document images. R-CNN, Spatial Pyramid Pooling (SPP-net), Fast R-CNN, Faster R-CNN, Mask R-CNN, and YOLO were some object detection models identified through existing literature in document layout segmentation.

R-CNN (Elgendy, 2020) detects a defined object within its boundaries using bounding boxes employing selective search to extract data from a specified region of an input image. The model output features from each region, and CNN-based classifiers classify these features. R-CNN is the primary object detection model, while it was further enhanced to Fast, Faster, and Mask R-CNN.

R-CNN extracts 2000 regions from an image through selective search, taking longer for training and testing. Instead of working on the 2000 region to convert them into feature maps, SPP-net (He et al., 2015) transforms the whole image on the feature map at once. Furthermore, an SPP layer is added to the final convolutional layer, and the fully connected layers receive the fixed-length outputs produced by the SPP layer. Nevertheless, SPP-net has a drawback in backpropagating through the spatial pooling layer, as it can be challenging to distribute error gradients across each pooling region during backpropagation. For instance, in max pooling, the gradient is solely allocated to the neuron with the most significant degree of activation, which can result in sparse gradients and prolonged learning. Thus, the network is refined on fully connected parts.

Similar to R-CNN, Fast R-CNN (Girshick, 2016) uses Edge boxes to specify regions and solves the SPP-net problem. It uses a simple back-propagation calculation similar to maxpooling gradient calculation except that pooling regions overlap. Therefore, a cell can have gradients from multiple regions. Fast R-CNN is efficient because it avoids repetition by sharing calculations on overlapping edge boxes and uses one deep CNN to extract image features once, unlike R-CNN. Fast R-CNN selects images using ROI to extract fixed-length vectors of features from the object's feature map. Thus, it forms one layer rather than pyramidal layers.

Faster R-CNN (Ren et al., 2015) replaced expensive selective search with Region Proposal Network. A single CNN detects objects in two stages. First, the deep CNN proposes Region Proposal Network (RPN) directly without edge boxes. In the second stage, the Fast R-CNN detector integrates RoI Pooling to extract and classify features from each box. All CNN algorithms use regions to identify and localize objects in images, not the entire image.

In Yolo (Wang, 2021), The object should be in the center of this SxS grid so that bounding boxes identify objects in each grid. Class probabilities will be output for each bounding box, and bounding boxes with class probabilities above the threshold value are chosen to locate objects. However, YOLO struggles to identify small objects in images. Thus, in our scenario, identifying words or letters can be difficult due to spatial constraints in YOLO.

Mask R-CNN (Almutairi and Almashan, 2019; He et al., 2017), the latest R-CNN architecture enhancement, performs pixel-level segmentation. An additional branch predicts ROI segmentation masks pixel-by-pixel. Previous R-CNN versions do not support this input-output alignment, and Mask R-CNN outputs class labels, bounding box offset, and object mask. In Mask R-CNN, a Regional Proposal Network proposes bounding boxes, and a Binary mask classifier generates masks for object classes.

Considering the appropriate object detection model for document layout segmentation, the primary focus of the study will be on three extensible methods: transfer learning, ensemble learning, and continual learning.

Transfer learning involves using knowledge from one or more source domains to improve target domain learning by Transferring knowledge between domains (Zhuang et al., 2020). Transfer learning is used to solve the problem of insufficient data in the target domain by using data and knowledge from the source domain, and it can also be used to expand the base model. Transfer learning involves transferring model parameters, feature representations, learning algorithms, and topologies.

Ensemble learning (Huang et al., 2009) systematically generates and combines many models, such as classifiers or experts, to solve a computational intelligence problem. Ensemble learning improves model performance or reduces the risk of choosing the wrong model for classification, prediction, and function approximation. Ensemble learning can also give a model confidence, choose optimal or nearly optimal features, fuse data, learn incrementally, learn nonstationary, and correct errors.

Continual learning (Menezes et al., 2022) is a lifelong learning approach based on observations of how the human brain perceives and learns new knowledge throughout life. It trains deep learning models with fewer data to accelerate learning. This approach has the advantage over traditional learning models of using previously acquired knowledge to learn more and apply it to new problems. Thus, previous knowledge will be used to learn more, and training speed is much faster than transfer learning and deep learning.

In order to train and evaluate the models, the study employs LayoutParser (Shen et al., 2021) as the Python library, which contains infrastructure for running and training deep learning models in detecting document image layouts. The library has several pre-trained models that can be used directly for usage or can be customized to new models based on the requirements of the researchers and datasets supporting unique and challenging document analysis tasks.

Overall, the approaches used in this research provide a rigorous and objective evaluation of the effectiveness of various extensible approaches for document layout segmentation and practical insights into the application of LayoutParser and object detection models for processing document images. Hence the study will focus on identifying an optimum object detection model for the domain of document layout segmentation and exploring the extensible approach for document segmentation, namely transfer learning, ensemble learning, and continual learning, by considering the identified object detection model as the base model. To achieve this, existing literature will be analyzed to identify suitable object detection models, and the performance will be evaluated. With that as the base model, three extensible approaches will be evaluated and comparatively analyzed on their performance. Furthermore, document segmentation will be done using each approach, and the research strives towards developing new applications that could process a broader range of documents. The extensibility of document layout segmentation can improve the accuracy of reading order detection and identifying the critical parts of a document which supports document summarization and translation.

Document Layout Segmentation: Forczmanski et al. (2019) have proposed the segmentation ´ of scanned documents using a deep-learning approach. YOLO was used as the object detection model to identify elements from document images. Stamps, logos, text, signatures, and tables are identified here with more than 96% average accuracy. However, only these elements are identified in this research. Xu et al. (2020) propose LayoutLM, where text and layout are jointly learned in a single framework for document level pre-training for the first time. The tasks for the model were form and receipt understanding and document image classification. It was noticed that accuracy increases with the increase of the dataset, and it outperformed several state of-the-art pre-trained models using BERT as the base model. LayoutLMv3, built on top of LayoutLM by Huang et al. (2022), is another pre-training technique with unified text and image masking. An additional task was introduced as document visual question answering. The model also shows better results for the previous tasks, and it does not rely on CNN or Faster R-CNN backbone to extract visual features. A one-dimensional CNN-based approach for image layout in documents was proposed by Oliveira and Viana (2017) to segment texts, tables, and images in document images. Using this approach, the dataset size and training time can be reduced, and it can be used in mobile applications and cloud-based systems with less computational power. Shen et al. (2021) proposed LayoutParser, a unified toolkit for deep learning-based document analysis and implemented it to use with Fast R-CNN or Mask R-CNN. The experiments were done to generate high-quality structured data for a Japanese historical table and act as a

lightweight table extractor. Average precision of 96.97 was observed, and the toolkit can also be customized for deep learning-based document image analysis in other work.

Ensemble-based Approach: A Faster R-CNN approach was proposed by Lee et al. (2018), including model selection and box voting methods. However, the detection does not show its best performance in a generalized approach, but if it is fine-tuned for specific criteria, the approach can be used for better results. Aggarwal et al. (2020) proposes another multimodal approach for form structure extraction, which extracts text blocks, text fields, choice blocks, and choice groups. The model is trained to identify the candidate elements when an element is given. Casado-Garcia and Heras (2020) surveyed to discuss the use of ensemble methods for object detection, and it stated that an ensemble approach could be used to improve performance. However, many diverse models are required for this task, with high computational expense on training and combining models. Based on mask R-CNN, an end-to-end network is proposed by Xu et al. (2021) for text, table, figures, and math detection, and these models perform better in detecting smaller objects. The researchers state that it could be generalized for unseen objects as future work. Aboneh (2022) proposes a stacking based ensemble learning method for multi-spectral image classification, which has optimized ensemble learning and performs experiments to find the effectiveness of the ensemble approach, where the stacking method in ensemble learning showed superior results in terms of classification accuracy, despite the complexity and heterogeneity of multi-spectral data.

Incremental Learning: Incremental Learning of Object Detectors without catastrophic forgetting proposed by Shmelkov et al. (2017) presents a method for incremental learning addressing the catastrophic forgetting problem using a distillation loss function. This approach performed better for object detection than the existing approaches when tested on PASCAL VOC (Visual Object Classes) and COCO (Common Objects in Context) datasets. Peng et al. (2020) proposed Faster ILOD: An incremental learning object detector based on Faster R-CNN, is an end-to-end efficient object detection method using knowledge distillation. First, researchers evaluated RPN-based detector performance with incremental distillation, and then the model was fine-tuned for new tasks using multi-network adaptive distillation to retain previous knowledge. Menezes et al. (2022) propose Incremental Learning of Object Detectors with limited data, which consists of three parts. The first part is a pretrained base model where all increments use ResNet34. The second part is a CNN network extension layer; network extension layers are frozen after training to prevent catastrophic forgetting; the third part is classification. The model was incrementally tested on CIFAR-100, which improved accuracy with high true positives. Continual object detection is a class-oriented object detection method evaluation proposed by Hafeez et al. (2021) that incorporates techniques such as backward transfer, forward transfer, and fast adaptation and recovery. It contains standard supervised learning followed by task incremental, domain incremental, and free continual learning. Achieving forgetting prevention and knowledge transfer in continual learning is a novel neural architecture proposed by Ke et al. (2021) to prevent catastrophic forgetting and transfer knowledge where a continuous learning plug-in was used on the base model. However, the plug-in capsules affect the model's efficiency and are identified as a limitation in this study.

The need for extensible approaches for document segmentation was observed when considering the literature since most of the proposed approaches by various researchers are only applied to one or a few types of documents. Furthermore, several SOTA models gave the baseline for document segmentation, allowing training models for a specific dataset and fine-tuning them to work with certain document types only. Therefore, a noticeable research gap has been identified where the extensible approaches have yet to be used in document segmentation approaches to extend the approach to segment in new document types. In this research, we seek an extensible approach for document layout segmentation with the incremental learning approaches, where the trained model can be extended for new document types with customized datasets and address the 'catastrophic forgetting' problem.

2 Methodology

The study followed a quantitative research approach and employed a quasi-experimental design as this method is appropriate for comparing the effectiveness of various extensible approaches to measure

differences. The methodology compares the evaluation results of the approaches for document layout segmentation.

2.1 Data Collection and Preparation

For this study, the research team considered scanned images of newspapers, research papers, invoices, and forms; each dataset contains 400 document images. The selection of newspapers was based on complex layouts, including headings, bodies, images, tables, ads, dates, and more. Newspapers also have different layouts, making document layout segmentation difficult. Research papers were included in the study due to their organized, side-by-side, or solitary layouts, which can relate to several other documents such as novels and articles. Invoices were chosen due to having numeric values, and short content, and Forms were selected based on the layout and white spaces, as well as the fact that forms are crucial to daily life and require user input, making them an important dataset for extensible segmentation evaluation.

However, when preparing the datasets, the research was limited to digitized, scanned images of English documents and handwritten documents are not accommodated. Also, the research focuses on small datasets which are applicable to real-world scenarios and only single-page documents are considered.

Research papers are obtained through the public repository Publaynet; A large-scale publicly available dataset for document layout analysis (Zhong et al., 2019) and a newspaper dataset were obtained through a public newspaper collection, namely Careers Wave[1] . Forms were obtained from a dataset consisting of 400 government forms[2] and invoices were obtained through a Kaggle repository[3] by Achraf Bribiche, who has made it publicly available. Annotated datasets were available only for research papers, and other datasets were manually annotated using Roboflow; a computer vision-based online platform to annotate data faster and more accurately. As the study employs a limited number of data, data preprocessing and augmentation steps were followed and annotated datasets were exported in COCO JSON format.

2.2 Data Preprocessing and Augmentation

During the data preprocessing phase, various methods are utilized to normalize the dataset. These include auto-orientation, scaling, and conversion to grayscale, all of which serve the purpose of harmonizing different resolutions and colour schemes that are present throughout the image set. In addition, an automated contrast adjustment has been added so that picture features can be distinguished more easily. The subsequent steps consist of tiling to split bigger images into more manageable subsections and static cropping to remove supplementary image portions. This makes the processing of high-resolution imaging much simpler.

The data integrity is preserved using class changes, which guarantees accurate class labels for precise model interpretation. In addition, a rigorous filtering process also eliminates every incidence of null values, further improving the overall data quality and trustworthiness.

For data augmentation, strategies include random transformations, flipping, rotating, cropping, and shearing and alterations to luminance and exposure. These strategies foster a model capable of learning generalized characteristics, mitigating the risk of overfitting.

In addition, blur and noise enhancements simulate various environmental conditions, expanding the range of situations the model can effectively manage. The extraction technique is utilized to develop an occlusion-resistant model by arbitrarily obscuring portions of the image. The mosaic augmentation, which combines four training images into a single composite image in varying proportions, increases the diversity of multi-scale and multi-aspect ratio training data. This method promotes a model training environment characterized by high-quality, diverse, and standardized input data, significantly improving model performance and robustness.

[1] Online available at https://www.careerswave.in/

[2] Online available at https://github.com/forms-data-structures/forms-data

[3] Online available at https://www.kaggle.com/datasets/achrafbribiche/documentclassification

2.3 Analysing and Evaluating Object Detection Models

The study employs LayoutParser (Shen et al., 2021) as a Python library to train and evaluate deep learning models to detect document image layouts. The library has several pre-trained models that can be customized into new models based on the requirements of the researchers and datasets supporting unique and challenging document analysis tasks.

- R-CNN performs slowly as it needs around 2000 proposals and fixed-size inputs for feature extraction, where it uses the selective search for object detection.
- SPP-net translates the entire image on the feature map matrix at once, but back-propagating through the spatial pooling layer was not possible, which is essential to train models with minimum loss.
- Fast R-CNN was introduced to address the problem of back-propagation in SPP-net, using “region of interest” which allows the model to process the region proposals in parallel but using selective search for proposals requires a larger number of region proposals.
- Faster R-CNN resolves this problem by using a region proposal network (RPN) for object detection. RPN is faster than selective search, but it uses a collection of regions to identify objects within an image. Therefore, the whole image is not identified in Faster R-CNN.
- In YOLO, the image is divided into an SxS grid, and the object should be positioned in the centre of the grid to identify the objects inside it. However, the difficulty of detecting small objects in an image is identified as a shortcoming of the YOLO.
- Mask R-CNN addresses this problem by combining an object identification which recognizes object classes with bounding box prediction in an image with a semantic segmentation task in which pixel-level segmentation is considered to allow for the recognition of small objects.

Considering the literature, object detection models listed above were analysed, however, Fast R-CNN, Faster R-CNN and Mask R-CNN were evaluated to obtain results. Several factors were considered when selecting the following R-CNN models for evaluation.

- Historical Significance: Fast R-CNN, along with its successors Faster R-CNN and Mask R-CNN, have significantly contributed to object detection tasks' advancement. Analysing these models can enhance the evolution of methodologies within the discipline.
- Consistent Evaluation: Since these models are all from the same "family" of object detection models, they have similar designs and architecture. This consistency can help ensure a fair comparison.
- Resource Limitations: Concerning computational resources, testing a large number of models may not be practicable. Concentrating on a few related models can make a project more manageable.
- Performance on Different Tasks: While YOLO is speedy, studies suggest that it may not perform as well as the R-CNN family on tasks requiring high precision, such as detecting small objects or instances in which objects are close.
- Detectron Framework: As the study utilises the Detectron framework for testing, utilised the readily implemented and optimally optimised models in that framework.

2.4 Evaluating Extensible Approaches

Considering the similar work extensibility, transfer learning, ensemble-based learning, and continual learning approaches are evaluated to find the most feasible approach with minimum loss and better performance for the extensibility of document layout segmentation.

Transfer Learning Approach: Transfer learning is applied as the first extensible approach by training a newly introduced class on top of the existing pre-trained model. In transfer learning, two layers at the topmost level of the pretrained model are frozen, and the remaining layers are finetuned such that they can be used to segment both document types; the newly introduced document type and the previous document type as shown in Figure 1.

Ensemble-Based Approach: In the ensemble-based approach, a combination of multiple models was used to obtain the final outcome, and the study employed a horizontal stacking Xie et al. (2020) method in

ensemble learning as shown in Figure 2. When combining multiple models in ensemble learning, all the models perform the segmentation task on the given document image, and by comparing the results of the confidence scores of each model the best model for segmentation is identified. Such that, the best possible outcome was obtained.

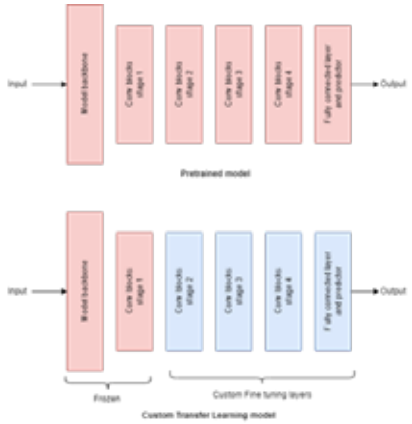


Figure 1: Transfer learning approach

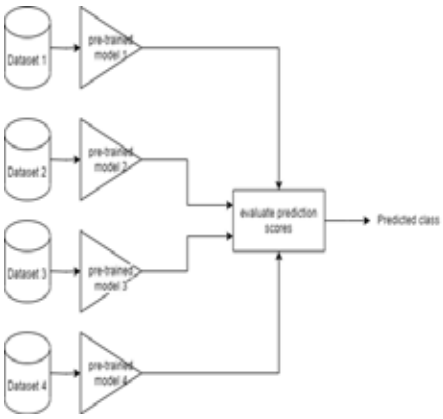


Figure 2: Ensemble-based approach

Continual Learning Approach: In continual learning, as shown as in Figure 3, knowledge distillation is carried out as a main part of the learning process, and it shares the information between one model (teacher model) to another (student model) without losing the information of the previous model (teacher model).

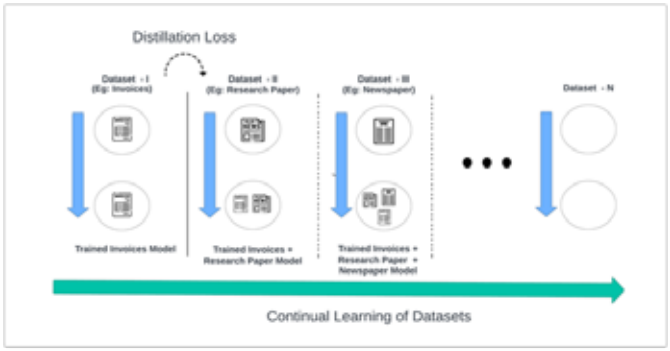


Figure 3: Continual learning approach

In order to perform this task in continual learning, the following four losses were calculated during the training period of the student model, and the losses will be added to the student model during each iteration. Classification Distillation Loss: Loss calculated using cross entropy between the raw, non-normalized output values (logits) of the teacher and student models. Logits reveal the model's confidence for each class as raw predictions. For example, assuming we have N classes, and the teacher model outputs logits [T_1, T_2, ..., T_n] and the student model outputs logits [S_1, S_2, ..., S_n]. First, the SoftMax function is applied to obtain the probability distributions over the classes for both the teacher and student models.

$$\text{Softmax}(T_i) = \frac{e^{T_i}}{\sum_{j \in N} e^{T_j}} \quad \text{for all } j \in N \qquad \text{Softmax}(S_i) = \frac{e^{S_i}}{\sum_{j \in N} e^{S_j}} \quad \text{for all } j \in N$$

Softmax (T_i) and Softmax (S_i) are the softmax values of the teacher model and student model outputs for the i^{th} class respectively. Likewise, the sum of all N classes was obtained, and the classification distillation loss is calculated by applying the softmax values to the cross-entropy loss.

$$L_{\text{cls}} = -\sum_{i \in N} \text{Softmax}(T_i) \cdot \log\big(\text{Softmax}(S_i)\big) \quad \text{for all } i \in N$$

Bounding Box Regression Distillation Loss: In object detection tasks, the model must predict the bounding box coordinates for the object in the image. This loss measures the student model's bounding box predictions, and bounding box regression improves the bounding box location. For an object, both teacher and student models predict the coordinates of the bounding box. These teacher and student model coordinates are represented in a 4D vector as [T_x, T_y, T_w, T_h] and [S_x, S_y, S_w, S_h], respectively. Then the Mean Squared Error (MSE) calculates the square of the difference between actual and predicted values and averages them for all instances. MSE can be used to calculate bounding box regression distillation loss, where N represents the total number of bounding boxes.

$$L_{\text{box}} = \frac{1}{N}\sum_{i=1}^{N} (T_x - S_x)^2 + \big(T_y - S_y\big)^2 + (T_w - S_w)^2 + (T_h - S_h)^2$$

Region Proposal Network (RPN) Classification Distillation Loss: RPNs are used in object detection to propose candidate regions where an object might exist. Each proposal is either "object" or "not object," as in a binary classification problem. The RPN classification distillation loss is calculated by the MSE between each proposal's teacher and student model 'objectness' scores. These scores are represented as T score and S score for teacher and student models respectively, where N represents the total number of region proposals.

$$L_{\text{cls}} = \frac{1}{N}\sum_{i=1}^{N} (T_{\text{score}} - S_{\text{score}})^2$$

Region Proposal Network Localization Distillation Loss: RPN also predicts the coordinates of a bounding box around the proposed object, and this is another regression task. The localization distillation loss is the MSE between the teacher and student models' bounding box coordinates for each proposal. These are also represented in the same way as the bounding box regression task, where N represents the total number of region proposals.

$$L_{\text{loc}} = \frac{1}{N}\sum_{i=1}^{N} (T_x - S_x)^2 + \big(T_y - S_y\big)^2 + (T_w - S_w)^2 + (T_h - S_h)^2$$

By minimizing these losses when training the model, the student model learns to produce a probability distribution over classes as close as possible to the teacher model. In each iteration, the student model will be updated with new data, and the total distillation loss will be calculated as the weighted sum of the individual losses.

2.5 Evaluation Plan

The study employs quantitative research methodology; the evaluation of the object detection models, and the extensible approaches are evaluated quantitatively by comparison. For evaluating the experimental outcomes, the objective evaluation metrics provided by the COCO Evaluator class are utilized.

(i) Two main metrics used are Precision and Recall
(ii) Average precision (AP) - This quantifies the effectiveness of a model at various levels of confidence. It provides a single score that balances the model's precision and recall and quantifies its ability to detect objects reliably. The study considered the mean average precision (mAP) for evaluation.
(iii) Intersection over Union (IoU) - The accuracy of an image-based object detector is evaluated. IoU computes the overlap between the predicted and actual bounding boxes. It is calculated by dividing the intersection area by the union area of both cells.

2.6 Hyperparameter Tuning

This study employed Optuna, a cutting-edge hyperparameter tuning framework, to optimize model performance. This framework addresses the problem as a black-box optimization problem, optimizing without prior knowledge of the function. The selected hyperparameters, including the learning rate, ResNet depth, weight decay rate, number of epochs, and loss function, were chosen due to their substantial influence on model performance.

The learning rate is a crucial hyperparameter influencing the weight update in response to the loss gradient. The search space for this investigation was set between 0.1 and 10^{-6}. ResNet, renowned for its capacity to construct competent models without vanishing or exploding gradient issues, was employed. Examined were various ResNet architectures, namely ResNet-34, ResNet 50, ResNet 101, and ResNet 152. Weight decay, also known as L2 regularization, prevents overfitting and adds a penalty term to the loss function to encourage the model to learn simplified patterns. The weight decay range was set between 0.01 and 10^{-7}. Large numbers, specifically 5000, 10000, and 15000, were evaluated to determine the model's generalizability based on the number of epochs. Diverse loss functions were investigated, including Smooth L1, which is less sensitive to anomalies, and GIoU and DIoU, which are both utilized for bounding box regression in object detection tasks.

After 1000 trials with Optuna, the optimal hyperparameters were determined to be a 0.002 learning rate, ResNet 50 architecture, 0.0001 weight decay, 10000 training epochs, and the Smooth L1 loss function. This study exemplifies the significance of systematic hyperparameter tuning by demonstrating the potential for performance gains when tuning is performed methodically.

3 Results and Discussion

After hyperparameter tuning, the best parameters were obtained as Learning rate - 0.002, RESNETS Depths - 50, Weight decay - 0.0001, Epochs - 10000, and Loss function - smooth 11.

3.1 Object Detection Models

The experiments on object detection models demonstrated that the Mask R-CNN outperformed other models by achieving a higher mAP score as shown in Table 1.

The versatility for object detection and semantic segmentation at the pixel level by generating object masks and its robustness in dealing with overlapping images has made Mask R-CNN superior. Compared to the other two models, Mask R-CNN provided more accurate results in identifying smaller text segments. However, unlike Faster R-CNN, Mask R-CNN uses RPN for object detection rather than selective search, which improves its computational efficiency compared to Fast R-CNN.

Table 1: Comparison of Object Detection Models.

		Research Papers	Invoices	Newspapers	Forms
Fast R-CNN	mAP	53.054	50.337	49.111	21.255
	AP75	57.698	55.934	57.135	16.287
	AP50	80.067	83.247	72.129	46.301
Faster R-CNN	mAP	55.244	51.525	47.147	22.781
	AP75	56.510	61.824	53.060	19.470
	AP50	79.964	84.735	68.956	46.284
Mask R-CNN	mAP	59.899	53.562	61.365	23.507
	AP75	70.981	55.933	65.955	21.120
	AP50	82.637	85.670	84.397	45.260

[a] Results obtained for mAP, AP50, and AP75 using Fast R-CNN, Faster R-CNN and Mask R-CNN on Research papers, Invoices, Newspapers and Forms.

3.2 Extensible Approaches

Extensible approaches were evaluated using Mask R-CNN as the object detection model, and the results for each method were compared by training them under the same conditions.

Transfer Learning Approach: In the study, several experiments were conducted on transfer learning to draw a conclusion and assess the performance of transfer learning for the extensibility of document layout segmentation. Following experiments were performed, and results were obtained from each criterion.

(i) Training research papers dataset on a pre-trained model for invoices as the base model.
(ii) Training research papers dataset on a pre-trained model for newspapers as the base model.
(iii) Training forms dataset on a pre-trained model for research papers as the base model.
(iv) Training invoices dataset on a pre-trained model for newspapers as the base model.
(v) Training invoices dataset on a pre-trained model for forms as the base model.

When training a child model on top of a parent model by freezing some of the parent model's layers, the knowledge from previously trained models was not retained, as determined by the experiment's findings. The mAP of the parent model was reduced to 0.108 on average, initially having a higher value of 53.525 on average before retraining the model. In consideration of these findings, transfer learning is not suitable for the task of extending the document layout segmentation task.

Ensemble-Based Approach: The approach that was utilized in ensemble learning was horizontal stacking. Since the output is based on the confidence score assigned by each model to an input image, the stacking order of the models does not affect the output. The average mAP scores for all document types using ensemble learning were approximately 50.785.

According to the study, this method produced superior results, with over 80% of newly added data accurately predicted from various test documents. The major disadvantage of this method is that the newly introduced document must pass through all models, and all models must be pretrained from scratch. Therefore, it is computationally inefficient and time-consuming, but the model outperforms other extensibility approaches regarding performance and accuracy.

Continual Learning Approach: In order to determine the most suitable teacher model for knowledge distillation, six distinct scenarios were investigated while evaluating the continuous learning approach. The outcomes of continual learning approach demonstrated in Table 2 shows that the performance of the models varied across various scenarios.

Table 2: Results for Continual Learning Approach across various scenarios.

	mAP
Newspapers (Teacher model) -> Research Papers -> Invoices	21.533
Newspapers (Teacher model) -> Invoices -> Research papers	19.568
Research papers (Teacher model) -> Newspapers -> Invoices	15.864
Research papers (Teacher model) -> Invoices -> Newspapers	15.740
Invoices (Teacher model) -> Newspapers -> Research papers	5.018
Invoices (Teacher model) -> Research papers -> Newspapers	5.377

Scenario I received the highest mAP score among all the scenarios and decreased in descending order for other scenarios. The observation suggests that the teacher model, which was trained on a dataset with a larger number of classes, may have acquired more intricate patterns and higher-level features, making it difficult for the student model to adapt to tasks of lower complexity and more biased towards the teacher model with a larger number of classes.

Continual learning is a promising yet challenging approach. The overall ROC curves and mAP results for all scenarios remained relatively low compared to ensemble-based learning. Also, it is essential to acknowledge that the continual learning approach has certain limitations, especially when addressing limited or imbalanced datasets, which take up high computational power. The continual learning approach may require more detailed and diverse training data to effectively transfer knowledge from the teacher to the student model.

In addition, when a teacher model is trained on a dataset with a larger number of classes, it can learn more complex patterns and higher-level characteristics. It can lead to difficulties for the student model in adapting to tasks of lower complexity and an inherent bias toward the teacher model with more classes. Due to these limitations, continual learning could not be as effective as the ensemble learning approach.

Comparing extensible approaches, it is evident that ensemble learning performs better than the other approaches. This superior performance is primarily due to the stacking of multiple models, each separately trained on distinct datasets. Consequently, when a new document is introduced, it is analyzed by each model to determine the most appropriate model for the segmentation task and then performs document layout segmentation.

4 Conclusion

In conclusion, the study aimed to comparatively analyse the optimum extensible approaches for document layout segmentation using an appropriate object detection model, which can be used to segment new document types.

Due to instance segmentation at the pixel level, one of the significant findings of this study is that the Mask R-CNN object detection model is the best model for segmenting document images in terms of mean average precision at 57.6% while faster and fast obtained AP at 51.972% and 50.829% respectively. In addition to the traditional transfer and ensemble-based learning methods, a novel approach, Continual Learning, was investigated for the first time in the context of extensible approaches to document layout segmentation. However, continual learning showed limitations with imbalanced or limited datasets demanding high computational expense. Also, the models trained on datasets with more classes and complex layouts introduced a bias towards the class and challenged the adaptability for less complex layouts. Therefore, it was discovered that the ensemble learning approach demonstrated superior performance with an average mAP score at 50.785 among the extensible approaches.

In addition to contributing to document layout segmentation, this study provides insights into the future research directions on further document analysis tasks such as reading order detection, document summarization and translation.

REFERENCES

[1] Tagel Aboneh. 2022. Ensemble Learning Method for Multi-Spectral Image Classification. Technologies 10 01 2022), 1 – 18. https://doi.org/10.3390/technologies10010017

[2] Milan Aggarwal, Mausoom Sarkar, Hiresh Gupta, and Balaji Krishnamurthy. 2020. Multi-modal association based grouping for form structure extraction. 2020 IEEE Winter Conference on Applications of Computer Vision (WACV) (Mar2020). https://doi.org/10.1109/wacv45572.2020.9093376

[3] Abdullah Almutairi and Meshal Almashan. 2019.Instance Segmentation of Newspaper Elements Using Mask R-CNN. (2019), 1371–1375. https://doi.org/10.1109/ICMLA.2019.00223

[4] ´Angela Casado-Garc ıa and J´onathan Heras. 2020. Ensemble Methods for Object Detection. Frontiers in Artificial Intelligence and Applications (2020). https://doi.org/10.3233/FAIA200407

[5] Mohamed Elgendy. 2020. Chapter 7 - Object detection with R-CNN, SSD, and YOLO. Manning Publications Co.

[6] Paweł Forczma´nski, Anton Smoli´nski, Adam Nowosielski, and Krzysztof Małecki. 2019. Segmentation of scanned documents using deep-learning approach. Advances in Intelligent Systems and Computing (Dec 2019), 141–152.https://doi.org/10.1007/978-3-030-19738-415

[7] Ross Girshick. 2016. Fast R-CNN. 2015 IEEE International Conference on Computer Vision (ICCV) (May 2016), 1–2. https://doi.org/10.1109/iccv.2015.169

[8] Muhammad Abdullah Hafeez, Adnan Ul-Hasan, and Faisal Shafait. 2021. Incremental learning of object detector with limited training data. 2021 Digital Image Computing: Techniques and Applications (DICTA) (Dec 2021). https://doi.org/10.1109/dicta52665.2021.9647245
[9] Kaiming He, Georgia Gkioxari, Piotr Dollar, and Ross Girshick. 2017. Mask R-CNN. 2017 IEEE International Conference on Computer Vision (ICCV) (Oct 2017). https://doi.org/10.1109/iccv.2017.322
[10] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2015. Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence 37, 9 (Sep 2015), 1904–1916. https://doi.org/10.1109/tpami.2015.2389824
[11] Faliang Huang, Guoqing Xie, and Ruliang Xiao. 2009. Research on Ensemble Learning. 3 (2009), 249–252. https://doi.org/10.1109/AICI.2009.235
[12] Yupan Huang, Tengchao Lv, Lei Cui, Yutong Lu, and Furu Wei. 2022. Layoutlmv3: Pre-training for document AI with unified text and image masking. Proceedings of the 30th ACM International Conference on Multimedia (Oct 2022). https://doi.org/10.1145/3503161.3548112
[13] Zixuan Ke, Bing Liu, Hao Wang, and Lei Shu. 2021.Continual learning with knowledge transfer for sentiment classification. Machine Learning and Knowledge Discovery in Databases (Feb 2021), 683–698. https://doi.org/10.1007/978-3-030-67664-341
[14] Jinsu Lee, Sang-Kwang Lee, and Seong-Il Yang. 2018. An ensemble method of CNN models for Object Detection. 2018 International Conference on Information and Communication Technology Convergence (ICTC) (Oct 2018). https://doi.org/10.1109/ictc.2018.8539396
[15] Angelo Menezes, Gustavo de Moura, C´ezanne Alves, and Andr´e C. P. L. F. de Carvalho. 2022. Continual Object Detection: A review of definitions, strategies, and challenges. Continual Object Detection (May 2022). https://doi.org/10.48550/arXiv.2205.15445
[16] D´ario Oliveira and Matheus Viana. 2017. Fast CNN-Based Document Layout Analysis. 1173–1180. https://doi.org/10.1109/ICCVW.2017.142
[17] Can Peng, Kun Zhao, and Brian C. Lovell. 2020. Faster ILOD: Incremental learning for object detectors based on faster R-CNN. Pattern Recognition Letters 140 (Dec 2020), 109–115. https://doi.org/10.1016/j.patrec.2020.09.030
[18] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. 2015. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. In Proceedings of the IEEE International Conference on Computer Vision (ICCV). 91–99.
[19] Zejiang Shen, Ruochen Zhang, Melissa Dell, Benjamin Charles Lee, Jacob Carlson, and Weining Li. 2021. Layoutparser: A unified toolkit for Deep Learning Based Document Image Analysis. Document Analysis and Recognition – ICDAR 2021 (Jun 2021), 131–146. https://doi.org/10.1007/978-3-030-86549-89
[20] Konstantin Shmelkov, Cordelia Schmid, and Karteek Alahari. 2017. Incremental learning of object detectors without catastrophic forgetting. 2017 IEEE International Conference on Computer Vision (ICCV) (Aug 2017). https://doi.org/10.1109/iccv.2017.368
[21] Shuo Wang. 2021. Research towards yolo-series algorithms: Comparison and analysis of object detection models for real-time UAV applications. Journal of Physics: Conference Series 1948, 1 (Jan 2021). https://doi.org/10.1088/1742-6596/1948/1/012021
[22] Jingjing Xie, Bing Xu, and Zhang Chuang. 2020. Horizontal and Vertical Ensemble with Deep Representation for Classification. Journal of XYZ 10, 2 (2020), 100–110. https://doi.org/10.1234/xyz123456
[23] Canhui Xu, Cao Shi, Hengyue Bi, Chuanqi Liu, Yongfeng Yuan, Haoyan Guo, and Yinong Chen. 2021. A page object detection method based on mask R-CNN. IEEE Access 9 (2021), 143448–143457.https://doi.org/10.1109/access.2021.3121152
[24] Yiheng Xu, Minghao Li, Lei Cui, Shaohan Huang, Furu Wei, and Ming Zhou. 2020. Layoutlm: Pre-training of text and layout for Document image understanding. Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (Aug 2020). https://doi.org/10.1145/3394486.3403172
[25] Xu Zhong, Jianbin Tang, and Antonio Jimeno Yepes. 2019. PubLayNet: largest dataset ever for document layout analysis. In 2019 International Conference on Document Analysis and Recognition (ICDAR). IEEE, 1015–1022. https://doi.org/10.1109/ICDAR.2019.00166
[26] Fuzhen Zhuang, Zhiyuan Qi, Keyu Duan, Dongbo Xi, Yongchun Zhu, Hengshu Zhu, Hui Xiong, and Qing He. 2020. A Comprehensive Survey on Transfer Learning. arXiv:1911.02685 [cs.LG]

Styllizea: Name Recommending System for Fashion Products Using Image Captioning on Single-Object Images and Multiple-Object Images

Mahidi Wijayantha

School of Computing, Informatics Institute of Technology, Colombo 06, Western Province, Sri Lanka, mahidi.2019766@iit.ac.lk

Niwarthana Kariyabaduge

School of Computing, Informatics Institute of Technology, Colombo 06, Western Province, Sri Lanka, niwarthana.k@iit.ac.lk

ABSTRACT

As online shopping gains popularity among consumers, fashion store owners face the challenge of efficiently managing their e-commerce systems. Uploading products and inventory onto these platforms has always been a time-consuming task, requiring creativity to accurately fill in relevant attributes and descriptive words. This paper aims to automate the process of populating the name field in fashion e-commerce systems by analyzing product images. The analysis includes identifying both individual clothing items and multiple clothing objects within an image. To tackle this issue, the study delves into image captioning, a technique that describes images by summarizing extracted features and generating human-readable text descriptions. The focus is on extracting and embedding external attributes specific to fashion products during the image captioning process. Generating captions for single and multiple object images has been a challenging and unsolved problem. The main hurdle lies in ensuring that the generated descriptions contain an adequate number of descriptive words and pay attention to object details. Ultimately, this paper revolutionizes an industry that has remained largely non-automated for the past decade by developing a final product based on image captioning technology.

KEYWORDS

image captioning, fashion name generator, fashion object detection

1 Introduction

Based on statistics from the fashion industry, it can be regarded as a highly profitable sector, generating billions of dollars in revenue [1]. In contemporary times, there are distinct boundaries between luxury fashion and apparel targeted at the general market. However, in the 1990s, such boundaries did not exist. Fashion is commonly defined as a style of clothing and accessories that individuals choose to wear as a means of self-expression, encompassing various elements such as clothing, accessories, lifestyle, footwear, hairstyle, makeup, and body posture [2]. Each country and region have its own unique fashion, leading to regional and national differences in style. Nevertheless, the fashion industry's objective is to design, manufacture, distribute, and promote various types of apparel for men, women, and children, ranging from luxury fashion to everyday wear [3]. This highlights the highly globalized nature of the fashion industry. Prior to reaching the hands of customers, fashion products undergo a series of essential steps, including sourcing the necessary raw materials, manufacturing the apparel using these materials, retailing, and the collective efforts of organizations involved in advertising and promotion [4].

2 Background

2.1 Fashion and E-Commerce Platform

Consumers now have the option to shop online using Information and Communication Technologies (ICT) instead of physically visiting stores. This mode of shopping, known as online shopping, e-shopping,

network shopping, or Internet shopping [5], offers the convenience of shopping from anywhere, freeing consumers from the need to visit physical stores. It is expected to have a significant impact on people's daily lives by saving travel time [6]. However, online shopping does come with the drawback of waiting for the delivery of purchased products. In this paper, the focus is on the e-commerce fashion industry, as clothing is a global necessity in daily life. With the growth of internet technology, many fashion businesses now utilize the power of the internet to promote and sell their products [7]. Fashion products are particularly popular online purchases, and fashion retailers have achieved great success by incorporating these technological trends into their fashion e-commerce platforms [8].

2.2 Automated Name Generation for Fashion Products

The process of adding fashion product images to e-commerce fashion websites, along with important details like product name, price, and description, is primarily carried out manually by users. When it comes to fashion products, individuals need to possess a creative mindset to effectively define attributes and use relevant words that can drive product conversions. In this paper, the author intends to automate this labor-intensive manual task by leveraging Machine Learning (ML) techniques.

3 Existing Work

Upon reviewing the aforementioned existing literature, it becomes evident that numerous image and scene captioning models have been developed and applied, including multimodal space-based models, visual space-based models, dense captioning, encoder-decoder architecture-based models, whole scene-based models, compositional architecture-based models, semantic concept-based models, attention-based models, and stylized captions [9]. These models employ various machine learning (ML) techniques, with image captioning being a significant component. Despite the availability of diverse image captioning implementations, limited research has been conducted specifically on fashion data, and there is a lack of a fashion product name generator application that can enhance product sales. Moreover, the research gap identified by the author lies in the ability to generate fashion product names by analyzing both single clothing object images and multiple clothing object images. Consequently, captioning such scenes poses a challenge.

3.1 Fashion Image Classification

[10] examined the concept of visual recognition for fashion, aiming to extract a predetermined set of visual features, including image category, subcategory, and associated attributes. The proposed model establishes a mutual influence between the image category and subcategory, without any direct connections to the attributes. This unified approach combines two methods to create the model. Firstly, a customized ResNet-50 CNN, specifically fine-tuned for fashion, is employed to address individual learning tasks. Secondly, it incorporates the Bidirectional Inference Neural Network (BINN) technique, inspired by [11], which jointly learns concept levels from a hierarchical tree and explores the relationships between labels. The model utilizes a message propagation block to maintain the hierarchical structure. This unified approach delivers competitive results in image category classification and attribute labeling. Figure 1 illustrates the architecture of the model.

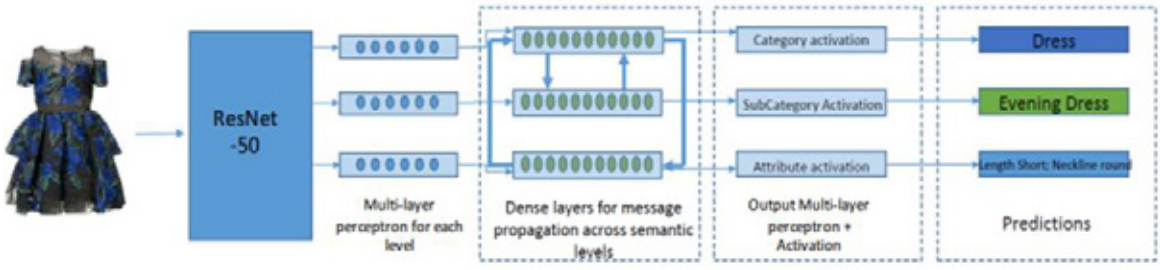


Figure 1: Proposed architecture fashion image classification [10]

3.2 Attribute based Fashion Searching

[12] developed a method for attribute-based image retrieval specifically designed for fashion products. This approach involves extracting the visual features of an image and utilizing them as a query. To retrieve the features and attributes, the well-established AlexNet CNN model was employed due to its effectiveness in this task. The attributes are categorized into two groups: general attributes (such as gender and category) and special attributes (like collar length and sleeve length). Special attributes have the ability to influence other attributes, making them more suitable for fashion-related searches. The successful recognition of attribute similarities by deep neural networks is crucial for this method to function properly. The system describes the sub-component sections of clothing products, employing the attention mechanism CNN to determine these parts. This enhances the search query process. The architecture of the system is illustrated in Figure 2, where the colors blue, red, and green represent the part extraction method, general attributes, and special attributes, respectively.

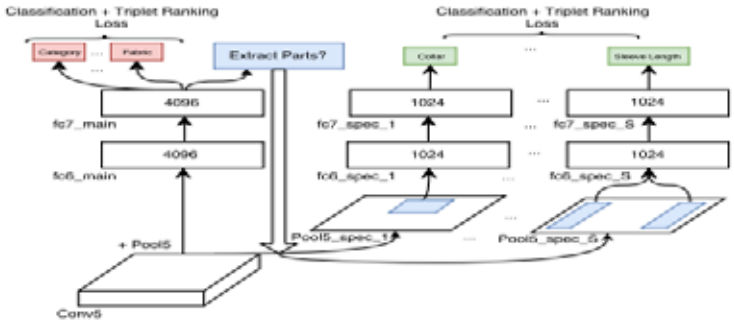


Figure 2: Architecture of attribute based fashion searching

3.3 Image Caption Generation for Single Clothes

[13] introduced a concept for generating captions for single clothing images, employing two main modules: attribute detection and language LSTM module. The attribute detection module is responsible for encoding the image with a specific feature representation, while the language LSTM module decodes the clothing features into a language-based description. The image captioning model initially extracts the features of the given image using a pre-trained CNN, utilizing the ResNet-101 model trained on the ImageNet dataset, which consists of over 100 million images for classification purposes. To predict a set of attribute features, an attribute network (AttriNET) is developed, utilizing attribute type indicators such as color, fabric, shape, and part. The attribute prediction model is a multi-column CNN with shared middle-level layers, taking the image as input and generating the attribute tags as outputs.

4 Methodology

4.1 The Task

The automatic generation of captions for images has gained significant attention in both industrial and academic domains. Image captioning can be broadly divided into two main sections: image recognition and caption generation [14]. Image recognition focuses on detecting the objects present in an image, while also considering factors such as the number of objects, interactions between objects, and the type of scene [15]. On the other hand, the caption generation section involves summarizing the information extracted from the image's features and generating a text description that is understandable to humans [16]. The author aims to narrow down the problem by focusing on the generation of fashion product names. The name of a product plays a crucial role in standing out and attracting potential customers. When consumers consider purchasing a product online, the first thing they encounter is the product's name. If the name does not align with what the consumer is searching for, it can significantly impact their attraction towards the product. Employees involved in naming products must be attentive and creative in selecting the most appropriate name. However, this task is not always simple or easy to accomplish. For instance, it can be challenging for

an ordinary employee with limited knowledge to generate creative, accurate, and meaningful product names manually.

Selecting an appropriate name for a fashion product can be a laborious and time-consuming process when performed manually by individuals, as it is not always straightforward. To address this challenge, the author intends to develop a system that allows users to upload an image of the fashion product for processing and analysis. Subsequently, the system will generate the most relevant corresponding name for both single fashion objects and multiple fashion objects within the same image. This approach aims to streamline and automate the naming process, making it more efficient and less burdensome.

4.2 Model Architecture

Manually recording the names of individual fashion objects or multiple fashion objects within a single image is an inefficient undertaking. Thus, there is a genuine requirement for automated generation of fashion product names. Our proposed system offers users the convenience of uploading an image of the fashion product, which will then undergo processing and analysis. The system utilizes four models for the processing and analysis task.

For **Single Clothing Category Recognition Model**, we employed the *DeepFashion* dataset, which consists of clothing images. To enhance the dataset, we added labels for 46 different clothing categories to over 100,000 images, and then split the data into training, validation, and testing sets. We employed transfer learning by initializing the ResNet34 model with pre-trained weights from the ImageNet dataset, which encompasses millions of labeled images from diverse categories. By freezing all the layers of the pre-trained ResNet34 model, except for the newly created layer, we focused on training only the weights of the new layer. This transfer learning approach allows us to leverage the knowledge acquired by the pre-trained model. Subsequently, we replaced the last layer of the network with a new linear layer to align with the number of classes in the dataset. The model was fine-tuned by training the new layer and then unfreezing and training all layers for a few epochs. The new layer plays a crucial role in producing the final predictions for the clothing categories. It takes the feature representations learned by the ResNet34 model and maps them to the 46 output classes. By replacing the last layer only while retaining the rest of the pre-trained ResNet34 model, our new model retains the powerful feature extraction capabilities of ResNet34, enabling it to capture meaningful patterns in clothing images. The pre-trained layers have already learned to recognize lower-level visual features, such as edges, textures, and shapes, which are relevant for various computer vision tasks. The same methodology with same dataset was used for the **Single Clothing Attribute Recognition Model** for recognizing attributes of single clothing items.

To train our **Multiple Clothing Category Recognition Model** for the purpose of generating fashion captions, the same methodology above used but with a different dataset. We utilized a large-scale open source dataset obtained from the reputable ASOS website. This dataset consists of 591,738 images, each resized to a resolution of 224 x 224 pixels. In order to facilitate efficient classification, we categorized the dataset into 90 distinct classes, representing different clothing categories.

For **Fashion Object Detection Model**, we employed YOLO (You Only Look Once), which offers fast and efficient detection in a single pass through the network. This characteristic allows YOLO to process images in real-time, even on devices with limited resources. Its speed and efficiency make it suitable for various applications, including autonomous driving, video surveillance, and robotics. Hence, we loaded the pre-trained YOLOv8 model for this particular task. To further tailor the model to our needs, we created a custom YOLO model by utilizing the pre-trained weights. We then proceeded to train the custom model, evaluate its performance, and test it by making predictions on random images from the dataset. For this purpose, we utilized a clothing dataset that includes annotation files, which we obtained from Kaggle. The system will identify whether the uploaded image is a clothing items or non-clothing item.

4.3 Parameter Optimization

Parameter optimization is a critical aspect of training machine learning models, as it directly impacts the model's performance, convergence, and ability to generalize well on unseen data. In this section, we discuss the parameter optimization process for each model in our research.

4.3.1 Learning Rate

For each model, the learning rate was chosen through a systematic search process, often starting with a reasonable default value and then fine-tuning it based on the model's response. Learning rate schedules, such as step decay or adaptive learning rate algorithms like Adam, were considered to adjust the learning rate during training to optimize convergence. Early stopping was implemented as a regularization technique to monitor the model's performance on a validation set and stop training when improvements plateaued or started to degrade.

4.3.2 Batch Size

The batch size for training was chosen based on a trade-off between computational resources and model performance. Smaller batch sizes require less memory but may lead to noisy updates, while larger batch sizes may require more memory but can lead to more stable updates. The batch size was determined through experimentation, considering the GPU capacity and memory constraints.

4.3.3 Data Augmentation

Data augmentation techniques were employed to increase the diversity of the training data and improve model generalization. Common data augmentation techniques, such as random rotation, flipping, and resizing, were applied to augment the images in the dataset. Augmentation strategies were carefully selected to preserve the semantic content of the images while introducing variability.

4.3.4 Evaluation Metrics

For each model, appropriate evaluation metrics were chosen to assess its performance during training and testing. In image classification tasks, metrics like accuracy, precision, recall, and F1-score were used to measure the model's performance on the validation and test sets. For the fashion object detection model, mean average precision (mAP) was used to evaluate the detection performance.

In conclusion, the parameter optimization process involved fine-tuning hyperparameters and model settings to achieve the best possible performance for each model. The choices made during parameter optimization were guided by a careful consideration of available resources, the dataset characteristics, and the specific requirements of each model. By systematically adjusting the parameters and evaluating the models using appropriate metrics, we aimed to build robust and effective systems for image captioning and fashion object recognition.

5 Results And Discussion

5.1 Recognition of category for single clothing object

We assessed the predictions made by our single clothing attributes recognition model on the training dataset to evaluate whether any significant biases were present. The results of this evaluation can be seen in Figure 3. Overall, the predictions on the training dataset displayed favorable performance, indicating that our model was able to capture the essential concepts related to clothing attributes.

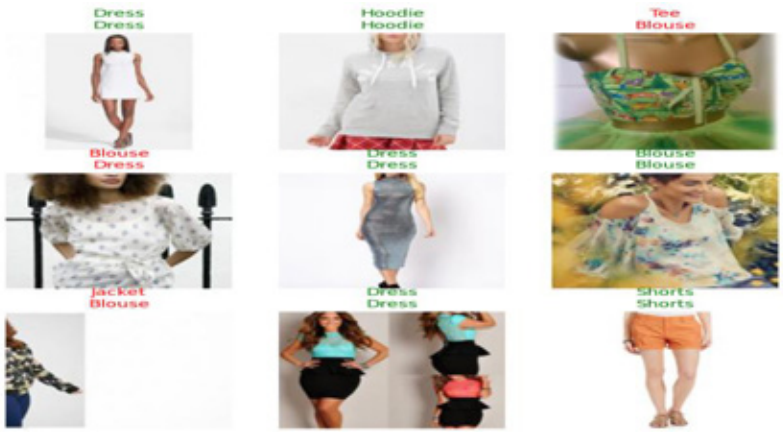


Figure 3: Result of single clothing category recognition

In line with the evaluation methodology employed in the original research paper, "DeepFashion: Powering Robust Clothes Recognition and Retrieval with Rich Annotations, CVPR 2016," we adopted the Top-3 and Top-5 accuracy metrics for our evaluation. In Table 1, a comprehensive overview is provided to compare the performance of various methods in terms of category classification, alongside our own method. Our model achieved a Top-3 accuracy of 88.6%, surpassing the benchmark accuracy by 6%. Additionally, our model achieved a Top-5 accuracy of 94.1%, surpassing the benchmark accuracy by 4%. These results highlight the efficacy of our model in accurately learning the key concepts necessary for successful clothing recognition tasks.

Table 1: Performance of various methods with our model

	top-3	top-5
WTBI	43.73	66.26
DARN	59.48	79.58
FashionNet+100	47.38	70.57
FashionNet+500	57.44	77.39
FashionNet+Joints	72.30	81.52
FashionNet+Poselets	75.34	84.87
FashionNet	82.58	90.17
ResNet34(Ours)	88.95	94.14

5.2 Recognition of attributes for single clothing objects

The provided visual representation of figure 4, illustrates the training progress of our single clothing attributes recognition model across 8 epochs, showcasing the enhancement in various performance metrics including loss, F-beta score, and accuracy on the validation set. Upon examining the output, several trends become apparent. Firstly, both the train_loss and valid_loss values consistently decrease throughout the epochs, signifying that the model is effectively refining its performance and assimilating knowledge from the training data. Moreover, the fbeta_score and accuracy_multi values demonstrate an upward trend, indicating that the model's performance on the validation set progressively improves with each epoch. This observation highlights the model's ability to enhance its predictive capabilities and achieve better accuracy in classifying and evaluating unseen data points.

epoch	train_loss	valid_loss	fbeta_score	accuracy_multi	time
0	0.226294	0.224668	0.432401	0.983527	12:45
1	0.226127	0.224570	0.434137	0.983594	13:00
2	0.226252	0.224492	0.438158	0.983602	12:14
3	0.225729	0.224438	0.440502	0.983568	11:52
4	0.225114	0.224385	0.441365	0.983659	11:56
5	0.225403	0.224334	0.442502	0.983679	12:15
6	0.225099	0.224321	0.442130	0.983759	12:58
7	0.225262	0.224338	0.442530	0.983778	12:37

Figure 4: Training progress of single clothing attributes recognition model

Figure 5 provides an analysis of the predictions made by our model on the training dataset to evaluate the degree to which the model has acquired a meaningful pattern. The comparison between our predictions and the actual labels indicates a strong alignment, demonstrating that our model effectively captures the fundamental concepts in the dataset. Additionally, when assessing the performance of the model on previously unseen images, it accurately classifies them, and the assigned labels exhibit a high level of accuracy. Notably, despite the presence of noisy training data, our model successfully learns the underlying fashion concepts and demonstrates robust generalization capabilities.

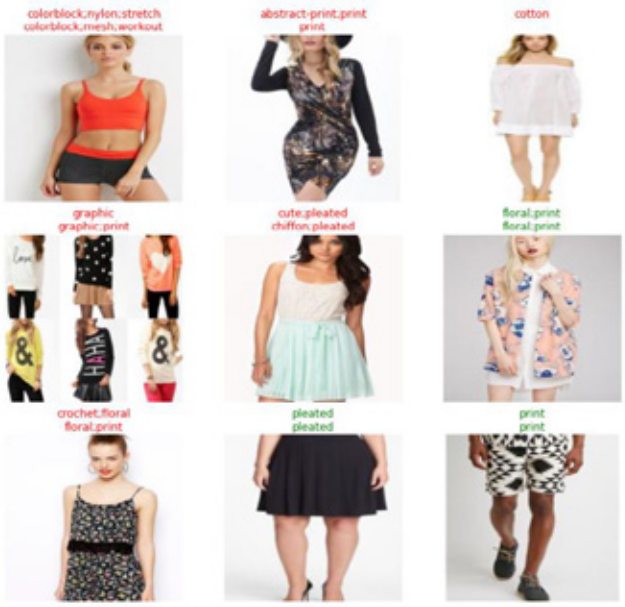


Figure 5: Result of single clothing attribute recognition

5.3 Detection of clothing objects

The performance metrics shown in figure 6, serve as valuable indicators of the model's detection capabilities:

mAP50: Mean Average Precision at the 50% IoU threshold. This metric measures the average precision of accurately detecting objects with a bounding box overlap of 50% or higher. It provides insight into the model's ability to correctly identify objects with a moderate level of spatial overlap.

mAP50-95: Mean Average Precision across IoU thresholds ranging from 50% to 95%. This metric offers a more comprehensive evaluation of the model's detection performance by considering a wider range of IoU thresholds. It provides a holistic assessment of the model's accuracy in detecting objects with varying levels of spatial overlap.

The model's accuracy may differ across different object classes, as illustrated in the provided output. The precision (P), recall (R), and mAP50 metrics are reported individually for each class, along with an overall average across all classes. These class-specific metrics allow for a more detailed analysis of the model's performance for specific object categories.

To summarize, as shown in the figure 6 the YOLOv8 model exhibits an mAP50 value of 0.77, indicating a relatively high average precision for objects with a bounding box overlap of 50% or more. The mAP50-95 value is reported as 0.523, suggesting the model's overall accuracy across a broader range of IoU thresholds.

To evaluate the performance of our trained YOLOv8 model, we conducted testing by randomly selecting various images and applying the model's detection capabilities to identify fashion objects within those images. This testing methodology allows us to determine the model's ability to discern between clothing items and non-clothing items within the uploaded images. By leveraging the object detection capabilities of the YOLOv8 model, we can accurately identify and classify fashion objects present in the test images.

```
Ultralytics YOLOv8.0.92 🚀 Python-3.7.12 torch-1.13.0 CUDA:0 (Tesla P100-PCIE-16GB, 16281MiB)
Model summary (fused): 218 layers, 25845550 parameters, 0 gradients, 78.7 GFLOPs
                 Class     Images  Instances      Box(P          R      mAP50  mAP50-95): 100%|██████████|
0,  1.48it/s]
                   all        537       2035      0.731      0.751       0.77      0.523
              sunglass        537         82      0.558      0.171      0.299      0.101
                   hat        537         77      0.771       0.83      0.826      0.492
                jacket        537        181      0.697      0.812      0.846      0.666
                 shirt        537        366       0.77      0.822       0.83      0.578
                 pants        537        114       0.91       0.93      0.965      0.741
                shorts        537        107      0.788      0.869      0.818       0.53
                 skirt        537        186      0.696      0.747      0.796      0.629
                 dress        537        128      0.596      0.867      0.761      0.596
                   bag        537        274      0.729      0.679      0.743      0.424
                  shoe        537        520      0.792      0.783      0.821      0.472
Speed: 0.3ms preprocess, 12.6ms inference, 0.0ms loss, 1.5ms postprocess per image
```

Figure 6: Performance metrics of clothing detection model

This process involves detecting and localizing the specific regions of the image that contain clothing items, distinguishing them from other objects or backgrounds. Through this testing approach, we assess the model's capability to discriminate between clothing items and non-clothing items, providing valuable insights into its accuracy and reliability. By analyzing the model's performance on various test images, we can verify its effectiveness in accurately identifying fashion objects and distinguishing them from unrelated elements in the scene. Through its advanced object detection algorithms and trained neural network, our model achieves high accuracy in distinguishing between these two categories. This enables us to obtain separate outputs indicating the presence of clothing items and non-clothing items in the analyzed images.

By successfully detecting clothing items and non-clothing items separately, our model provides valuable information for various applications such as fashion analysis, image categorization, and content filtering. This capability ensures that the model can effectively identify and process images based on their content, allowing for more targeted and accurate analysis or classification tasks.

5.4 Caption Generation for single clothing objects

The proposed system goes beyond predicting categories and attributes of fashion objects and achieves the task of generating highly accurate and relevant captions specifically tailored to single fashion objects. This accomplishment is illustrated in Figure 7, where the output captions can be observed. Through a combination of advanced computer vision techniques, deep learning models, and natural language processing algorithms, the system analyzes the visual features of the single fashion objects and translates them into meaningful and descriptive captions. These captions effectively convey the key characteristics, style, and details of the fashion items in a concise and informative manner. The accuracy and relevance of the generated captions highlight the system's ability to understand the visual content of the single fashion objects and capture the most significant aspects that are important for describing them. This ensures that the captions align with the specific fashion object and effectively communicate its unique attributes.

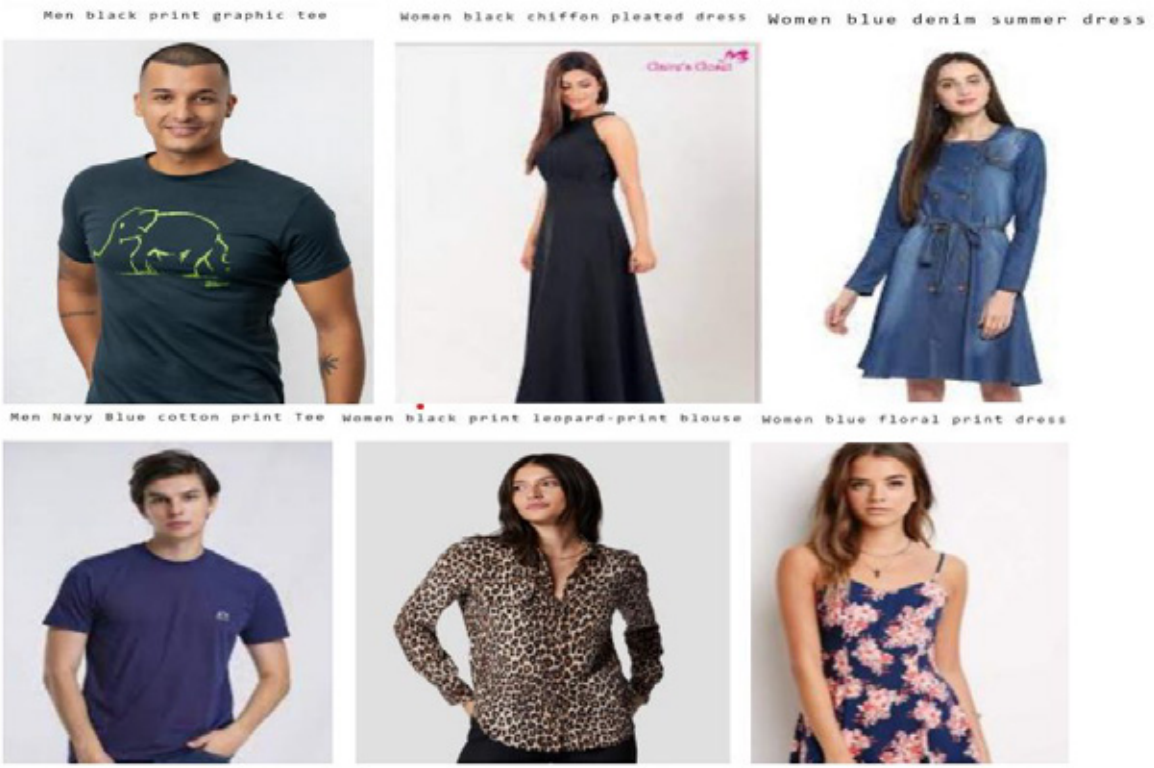


Figure 7: Result of Caption generator for single clothing objects

By generating accurate and relevant captions, the proposed system offers valuable assistance in various domains such as e-commerce, fashion recommendation systems, and fashion content generation. It enhances the overall user experience by providing informative and engaging descriptions of single fashion objects, enabling better understanding and decision-making for potential customers or users.

5.5 Caption generation for multiple clothing objects

We evaluated the performance of our model by examining the predictions it made on the training dataset. This assessment allowed us to estimate whether any biases or limitations were present in the model's performance. As depicted in Figure 8, the predictions made on the training dataset were highly satisfactory.

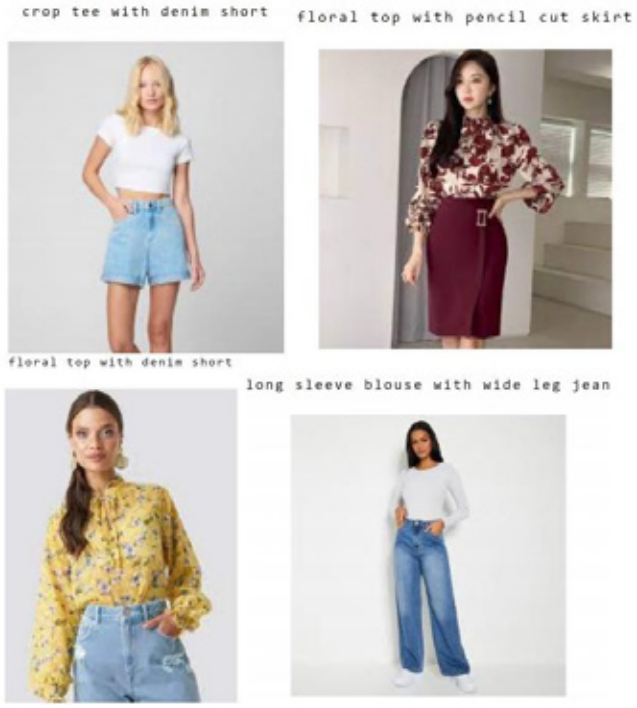


Figure 8: Result of Caption generator for multiple clothing objects

Our model demonstrated a strong ability to correctly capture the underlying concepts and characteristics of the clothing items. In summary, our multiple clothing category recognition model exhibited excellent performance on the training dataset, accurately recognizing and classifying various clothing categories.

6 Limitation

Based on the obtained results, it is important to acknowledge certain limitations of our system. One key limitation is associated with the size of the custom dataset utilized for training the multiple clothing category recognition model. The dataset, being relatively small, may impact the model's overall performance and its ability to generalize well to unseen data. With limited training examples, the model might struggle to capture complex patterns and nuances present in fashion images, potentially leading to suboptimal performance. Another limitation arises from the potential class imbalance within the dataset. Imbalanced class distributions refer to situations where certain fashion categories have a significantly larger number of samples compared to others. This class imbalance can introduce challenges during the training process. The model may exhibit a bias towards the majority classes, resulting in higher accuracy for those categories while performing poorly on minority classes. Such biased predictions can limit the system's effectiveness and undermine its ability to accurately recognize and categorize fashion objects across all classes.

These limitations highlight the need for further considerations in dataset collection and management. Efforts should be made to expand the dataset size, ensuring a more diverse and representative sample of

fashion categories. Additionally, addressing the issue of class imbalance by employing appropriate techniques such as oversampling can help mitigate biases and improve the model's performance across all fashion categories.

7 Conclusion

In conclusion, this paper introduces a novel approach for attribute recognition and classification in the domain of fashion. The proposed system addresses both single clothing items and multiple clothing categories, providing accurate and relevant suggestions for fashion captions based on user preferences. A significant contribution of this work is the development and implementation of the first-ever system for generating fashion captions specifically tailored to single and multiple clothing items. By leveraging multi-label classification techniques, the system can effectively recognize and categorize various attributes associated with fashion objects, enabling the generation of informative and contextually appropriate captions. Moreover, the system includes an important functionality to determine whether an uploaded image represents a clothing item or a non-clothing item. This capability enhances the system's versatility and ensures accurate processing and analysis of user-uploaded images. Overall, this paper presents a comprehensive solution for fashion attribute recognition, classification, and caption generation, catering to the needs of users seeking insightful and tailored information for both single clothing items and multiple clothing categories. The system's innovative approach and unique features pave the way for advancements in the field of fashion-related image analysis and provide valuable tools for fashion enthusiasts, retailers, and researchers alike.

8 Future Work

In the future, several areas can be explored to further improve and expand upon the current research. These potential directions for future work are outlined below:

Enhancing Dataset Accuracy: One avenue for future research involves improving the accuracy of the dataset used in the study. This can be achieved by employing more advanced annotation techniques, leveraging expert knowledge, or utilizing crowdsourcing platforms to gather more precise and reliable annotations. Enhancing the dataset's accuracy will contribute to the overall performance and generalization capabilities of the models.

Adding More Classes to the Dataset: Currently, the dataset consists of 90 fashion categories. To make the system more comprehensive and capable of identifying a wider range of multiple fashion categories, future work can involve expanding the dataset to include additional classes. This expansion will allow the system to cater to a broader array of fashion-related contexts and improve its overall versatility.

Incorporating Attribute Recognition with Multiple Category Recognition: To provide a more comprehensive understanding of fashion objects, future research can focus on integrating attribute recognition with the multiple category recognition model. By incorporating attribute recognition, the system can generate more detailed and informative captions that encompass not only the category but also specific attributes of the fashion items, such as sleeve length, neckline style, color, etc.

Extending the Multiple Clothing Category Dataset: In order to support multi-label classification, future work can involve extending the existing multiple clothing category dataset. This extension would involve annotating the dataset to include multiple labels for each image, enabling the models to handle cases where multiple categories are present within a single image. This enhancement will further enhance the system's ability to accurately recognize and classify fashion objects in diverse contexts.

Annotating the Dataset with Additional Attributes: To enrich the attribute recognition capabilities of the system, future research can focus on annotating the dataset with additional attributes beyond the ones currently considered. Attributes such as sleeve length, neckline style, fabric type, pattern, etc., can be included in the annotations. This expanded attribute recognition will provide more comprehensive and detailed information about the fashion items, enhancing the system's overall descriptive power.

By pursuing these future research directions, the proposed system can be enhanced, enabling more accurate, comprehensive, and contextually relevant fashion caption generation. The incorporation of additional classes, attributes, and multi-label capabilities will further advance the system's performance and expand its applicability in various fashion-related domains.

REFERENCES

[1] Global Fashion Industry Statistics (2021). Available at: https://fashionunited.com/global-fashion-industry-statistics

[2] John, S. (2022) fashion industry | Design, Fashion Shows, Marketing, & Facts | Britannica. Available at: https://www.britannica.com/art/fashion-industry.

[3] Nenni, M.E., Giustiniano, L. and Pirolo, L. (2013) 'Demand forecasting in the fashion industry: A review', International Journal of Engineering Business Management, 5(SPL.ISSUE). Available at: https://doi.org/10.5772/56840.

[4] Kabaivanova, S. (2017) The 4 levels of the fashion industry. Available at: https://made-to-measure-suits.bgfashion.net/article/15048/64/The-4-levels-of-the-fashion-industry.

[5] Farag, S. et al. (2017) 'Shopping online and/or in-store? A structural equation model of the relationships between e-shopping and in-store shopping', Transportation Research Part A: Policy and Practice, 41(2), pp. 125–141. Available at: https://doi.org/10.1016/J.TRA.2006.02.003.

[6] Hsiao, M.H. (2016) 'Shopping mode choice: Physical store shopping versus e-shopping', Transportation Research Part E: Logistics and Transportation Review, 45(1), pp. 86–95. Available at: https://doi.org/10.1016/J.TRE.2008.06.002.

[7] Kawaf, F. and Tagg, S. (2014) 'Online shopping environments in fashion shopping: An S-O-R based review', The Marketing Review, 12(2), pp. 161–180. Available at: https://doi.org/10.1362/146934712X13366562572476.

[8] Guercini, S., Bernal, P.M. and Prentice, C. (2018) 'New marketing in fashion e-commerce', Journal of Global Fashion Marketing, 9(1), pp. 1–8. Available at: https://doi.org/10.1080/20932685.2018.1407018.

[9] Jia, X., Gavves, E. and Fernando, B. (2015) 'Guiding the Long-Short Term Memory model for Image Caption Generation Tinne Tuytelaars KU Leuven ESAT-PSI, iMinds'.

[10] Quintino Ferreira, B. et al. (2018) 'A Unified Model with Structured Output for Fashion Images Classification'. Available at: https://doi.org/10.1145/nnn.

[11] Hu, H. et al. (2016) Learning Structured Inference Neural Networks with Label Relations, Computer Vision Fpundation. Available at: https://openaccess.thecvf.com/content_cvpr_2016/papers/Hu_Learning_Structured_Inference_CVPR_2016_paper.pdf

[12] Ak, K.E. et al. (2018) 'Efficient Multi-attribute Similarity Learning Towards Attribute-Based Fashion Search', in 2018 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, pp. 1671–1679. Available at: https://doi.org/10.1109/WACV.2018.00186.

[13] Li, X. et al. (2021) 'Clothes image caption generation with attribute detection and visual attention model', Pattern Recognition Letters, 141, pp. 68–74. Available at: https://doi.org/10.1016/j.patrec.2020.12.001.

[14] Srinivasan, L. and Sreekanthan, D. (2018) 'Image Captioning-A Deep Learning Approach', International Journal of Applied Engineering Research, 13(9), pp. 7239–7242. Available at: http://www.ripublication.com.

[15] Puscasiu, A. et al. (2020) 'Automated image captioning', 2020 22nd IEEE International Conference on Automation, Quality and Testing, Robotics - THETA, AQTR 2020 - Proceedings [Preprint]. Available at: https://doi.org/10.1109/AQTR49680.2020.9129930.

[16] Vinyals Google, O. et al. (2017) 'Show and Tell: A Neural Image Caption Generator', pp. 3156– 3164. Available at: https://www.cv foundation.org/openaccess/content_cvpr_2015/html/Vinyals_Show_and_Tell_2015_CVPR_pap er.html

Deepfake Low Resource Image Detection with Explainable Reporting

Abdul Baasith Shiyam†
Computer Science and Engineering, University of Westminster, United Kingdom,
baasith.shiyam1@email.com

Guhanathan Poravi
Department of Computer Science, Informatic Institute of Technology, Sri Lanka,
ghunathan.p@iitac.lk.

ABSTRACT

In this study, explainable artificial intelligence (XAI) is tested for its efficacy in identifying deep-fake pictures in contexts with limited resources. Several image analysis and detection techniques are examined in the study, including the usage of XAI for feature extraction and classification. The findings show that in order to increase accuracy and dependability, AI models must have a variety of detection techniques and must be interpretable. This study contributes to the domains of deep fake detection, XAI, and image analysis, offering suggestions for next studies in the area of AI-based imagine forensics.

The InspetionResnetV2 model was trained by the author of this study as the base, and they created a novel hybrid architecture for feature extraction that had an accuracy of 96%. The author employed LIME and Integrated Gradient XAI approaches to understand the model's output. Highlighting the areas utilized for the forecast, LIME XAI aids in explaining how the model arrived at its predictions. The interpretation of the pixels needed to arrive at the choice is displayed in the integrated gradient.

KEYWORDS

Deepfake, CNN, Optimization, Computer Vision, face2face Digital Media Manipulation, XAI

1 Introduction

Advancements in deep learning have led to the development of face swapping techniques, which can easily impersonate a person's identity and pose security challenges.[1] Deepfake technology has both positive and negative effects, as it can be used in various economic sectors, but can also be misused. The face swap method uses auto-encoders to train facial images and create counterfeit images or videos of a person. These days, the process of creating those altered photos and videos is also less complicated because all that is needed is the target person's identifying photo or a brief video. It takes less work to create convincing tempered footage. It is now feasible to create a deep fake with only a still image thanks to recent advancements.

Today's world is heavily reliant on technology, which is also continuously updated. It's past time for it to start misleading us and create unwanted commotion in our lives. One of these instances is deep fake technology. something sinister is going on since former US President Barack Obama was the subject of one of the first deep fakes. Such intricate impersonations on a person of his status are highly risky and could damage a whole country.[2]

Due to the power of these deep fakes, it's also essential to research the idea of deep fake detection. This is a process that is used to create deep fakes. [3] The primary concept is to produce a deep fake picture using GAN and compare it to the target deep fake image to determine if it is authentic or phony. This article describes how deep neural networks may be used to create deep fakes and masking methods. Additionally, by utilizing cutting-edge deep learning techniques, to offer every available defense against these fakes [4].

1.1 Deep Learning Works to Detect Forged Images.

For the purpose of identifying deep fake images. By using an attention mechanism that is employed in the classification of faked pictures, CNN and vision transformers that are combined may be used to learn new characteristics and for categorizing them (Tran et al., 2021). Ocular biometrics, based on CNN, are used to differentiate between real and forged images in a few research projects that aim to successfully locate the artifacts in the identification of phony films, photos, and other media (Kandasamy, Hubálovsk, and Trojovsk, 2022). This challenging image categorization challenge is tackled using deep learning.

2 Motivation

The emergence of deepfake technology has made it challenging to distinguish between real and fake videos, especially on social media platforms. This presents a problem as false information can be spread easily, leading to reputational damage or disguising someone's identity. AI-based solutions offer a cost-effective and efficient way to combat this issue, reducing the need for manpower and time consumption costs. This research aims to contribute to the fields of deepfake detection, cyber security, and social media by shedding light on the importance of detecting deepfakes and preventing their spread.

2.1 Related Work

Deep learning algorithms are being used in image forgery detection to combat the increasing spread of deepfake videos and images on social media. Various techniques such as CNN and vision transformers are being employed to identify and classify fake images by detecting artifacts and distinguishing between true and fake images.[5] The use of data-driven deep-neural networks serves as the foundation for the majority of existing deepfake detection techniques. The advancement of deep learning algorithms has both created new synthetic video and picture generation methods and provided solutions for detecting them [6].

Deepfakes have become a serious problem, especially in social media were fake news spreads rapidly. [7] A study proposed a novel neural network-based method to detect video deepfakes, which can be embedded in devices. Another research focused on making facial recognition systems more robust, particularly in smart homes. The use of deepfake technology to bypass security loopholes and break into smart homes is a major concern. [8] The researcher used CNN as a feature extractor in the detection of FSGAN-generated deepfake images, with good performance for high-quality images but lack of accuracy for low-quality images. The suggested approach to improve generalization is through one-shot learning.

Facial artifacts, such as those around the eyes, nose, lips, and inconsistencies in head poses, can be used to distinguish between real and fake faces in facial recognition techniques. However, advanced deepfake techniques have made it difficult to identify these artifacts. [9] proposed a technique to detect deepfake videos using the inconsistency of head poses. The algorithm utilizes a pre-existing model of facial expressions to create mismatch landmarks that differ between real and fake faces. The cosine distance between the two head direction vectors for actual and deepfake faces is used to predict the result. A support vector machine model trained with radial basis function using a training dataset achieved an accuracy of 89%. However, the model lacks generalization since it only detects deepfake images using Openface2.

Two papers discuss methods to detect deepfakes by identifying artifacts that are unique to fake images and videos. [10]identified dominant artifacts such as circular reflection in the eyes and inconsistency in teeth to detect fully generated deepfake images. They achieved an accuracy of 85% by combining these two features. [11] created a novel dataset containing videos with voice synchronization and proposed a deepfake detection method that considers both video and audio data. While existing algorithms were used for deepfake detection, the focus was on video lip sync deepfakes, and the dataset was a significant contribution.

3 Proposed Approach

The suggested architecture uses a novel hybrid convolutional neural network InseptionResnetV2 and conv2d layers for feature extraction. It is an end-to-end trainable deep neural network. The building is depicted below Figure (1). The hybrid architecture combining InceptionResNetV2 and Conv2D layers exhibits strong discrimination capabilities, leading to a significant improvement in detection accuracy compared to conventional deepfake detection approaches. With 35 epochs of training and well-tuned hyperparameters

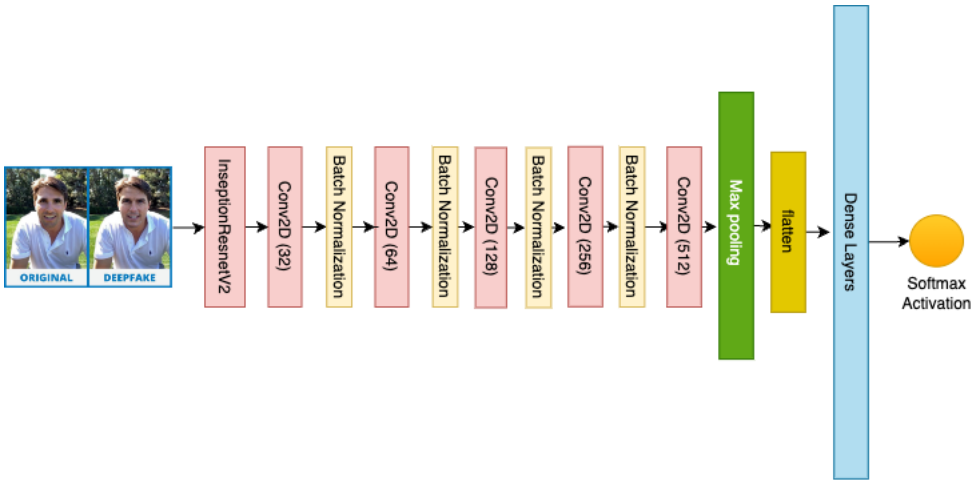


Figure 1 - proposed architecture

Deep fake photos and movies have become more and more popular due to their vast variety of uses and simplicity of creation. With the development of technology, it is simple to get high-quality altered movies. It is discovering a sizable number of programs that may be utilized by many different users with a range of computer abilities, from experts to beginners and investigate deep learning-based facial image synthesis methods. Shown in Figure(2)

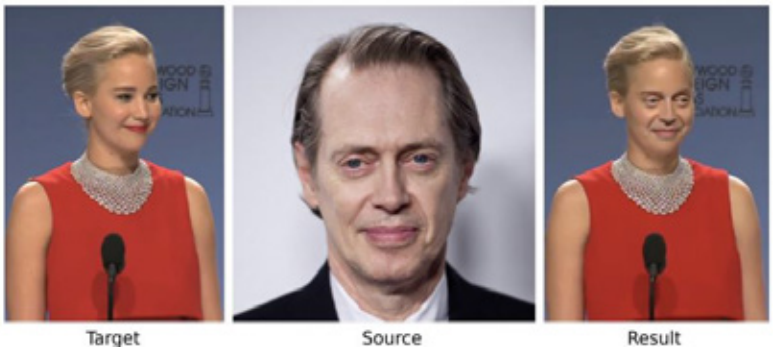


Figure 2 Example face2face

The author proposed a hybrid approach for deepfake detection that uses InceptionResNetV2 as the backbone model and creates a feature extraction using convolutional layers. This method aims to improve the accuracy of deepfake detection by combining the strengths of both deep learning models and feature extraction techniques.

The author utilized a dataset from Kaggle, which included faces from the FaceForensics++ dataset. This dataset comprised real, deepfake, and Face2Face images, with a total of 4000 deepfake and face2face images. The author has split the dataset into three sets, with 9000 images used for training, 2250 images used for validation, and 750 images used for testing. Additionally, the author has mentioned that data augmentation

techniques were used during the training process to increase the size of the training set and improve the model's ability to generalize to new, unseen data

3.1 Training the Model

The model was trained on a dataset for 30 iterations, and the author used training accuracy/loss and validation accuracy/loss to assess the model's performance. The model's high accuracy on the validation and training sets shows that it has gotten better at making predictions. The loss values also dramatically dropped, demonstrating a rise in prediction confidence for the model. To lessen overfitting, the author experimented

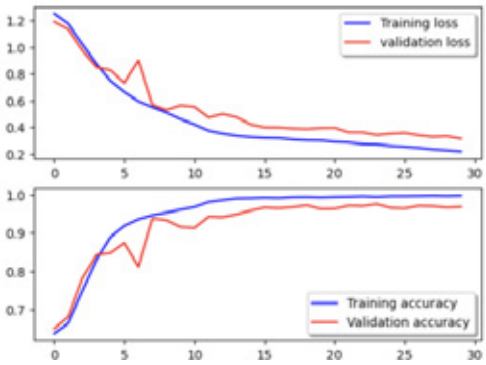


Figure 3 Training Loss and Accuracy

with several hyperparameters. The graph for the model was shown. Figure (3)

3.2 Hyper Parameter Tuning

The author has used three callbacks for hyperparameter tuning in their model. The first callback is EarlyStopping, which monitors the validation loss and stops the training process if there is no improvement after 10 epochs. The second callback is ReduceLROnPlateau, which reduces the learning rate by a factor of 0.2 if the validation loss plateaus for two consecutive epochs. The minimum learning rate is set to 0.00001. The third callback is ModelCheckpoint, which saves the best model with the lowest validation loss to a specified file path. The verbose parameter is set to 1, which allows for some output during the training process. These hyperparameter tuning techniques help to prevent overfitting and ensure that the model is generalizing well to new data.

4 Evaluation

In evaluating the model's performance, the confusion matrix is used to calculate other metrics such as precision, recall, and F1-score. The model achieved an accuracy of 0.92 and a recall of 0.98, indicating a low number of false negatives and a high recall, respectively. This suggests that the model is effective in identifying positive cases. The low number of false positives is reflected in the high accuracy, indicating that the model is not misclassifying many negative examples as positive. However, the author notes that there were still false positives, which could be attributed to overprocessing of real images.

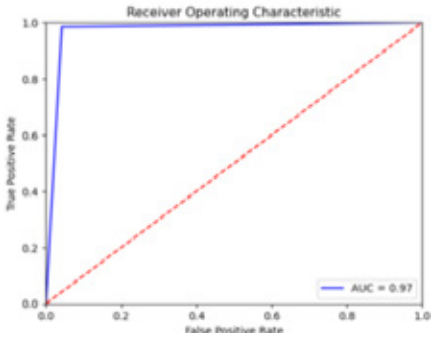


Figure 4 ROC curve

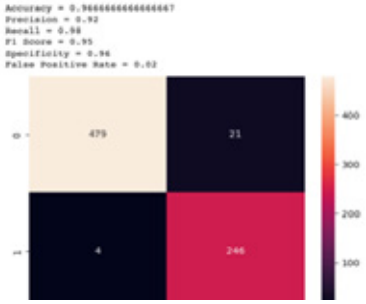


Figure 5 confusion matrix

The accuracy of the author's model is high at 0.9667, meaning that 96.67% of the test data were correctly predicted by the model. In other words, 92% of the photos that the model predicted as phony were truly fraudulent, according to the model's 0.92 accuracy rating. The recall of the model is 0.98, meaning that 98% of the genuine false pictures were properly classified as such by the model. Fig (5)

The AUC of the author's ROC curve is 0.97, which is almost ideal performance. This shows that the model successfully distinguishes between real and fake photos, having a high true positive rate and a low false positive rate. The model performs better the closer the AUC is to 1, and in this situation, the AUC of 0.97 is a fantastic result. Figure (4)

4.1 XAI Model



Figure 6 XAI LIME

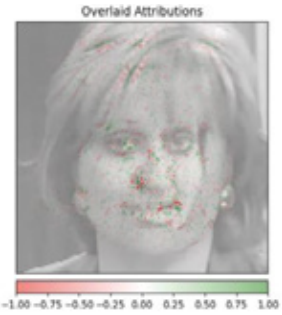


Figure 7 XAI Integrated Gradient

To explain the model's predictions, the author has employed LIME (Local Interpretable Model-Agnostic Explanations). LIME is a XAI (Explainable Artificial Intelligence) approach that emphasizes the key elements that were utilized to produce a prediction, making it easier to comprehend the model's decision-making process.

The model's predictions and the artifacts that the model is picking up have both been effectively explained by the author using LIME. The LIME research has demonstrated that by concentrating on particular properties like color, texture, and pattern, the model is successfully recognizing the artifacts. This sheds light on the model's ability to discriminate between actual and fraudulent photos and can help the model perform even better. Which can be seen in Figure (6)

The author used Integrated Gradients, which is a technique used in XAI to identify the pixels that were important in the model's prediction. The idea behind this method is to compute the gradient of the output with respect to the input image, and then integrate the gradient over a series of steps to determine the contribution of each pixel to the final prediction. The author applied this technique to a sample of the test data, and the resulting heatmap showed which pixels were most important in the model's prediction. The heatmap revealed that the model was able to accurately identify the artifacts present in the images, as the

important pixels were clustered around the regions containing the artifacts. Figure (7) demonstrated that the model was able to learn meaningful features from the images and use them to make accurate predictions.

5 Benchmarking

Table 1 Bechmarking

	Acc	Loss	Precision		Recall		F1 Score	
			Fake	Real	Fake	Real	Fake	Real
Xception	0.80	0.40	0.86	0.70	0.84	0.72	0.85	0.71
ResNet 50	0.65	0.57	0.68	0.45	0.92	0.13	0.78	0.20
VGG 16	0.67	0.52	0.72	0.52	0.85	0.32	0.78	0.40
EfficentNet	0.66	0.70	0.67	0.00	1.00	0.00	0.80	0.00
Ours	**0.96**	**0.13**	**0.99**	**0.92**	**0.96**	**0.98**	**0.97**	**0.95**

To find the top-performing model for deepfake detection, the author of the benchmarking chapter ran model-based benchmarking studies on four distinct deep learning architectures. The study examined the VGG16, ResNet50, EfficientNetB0, and Xception Net models.

6 Discussion

The author's research emphasizes how difficult it is to spot deepfakes in low-resource, low-pixel pictures. This is a critical problem since accurate identification of false images depends on the quality of the picture in the field of deepfake detection. The study makes a significant addition to this discipline since it tackles a pressing issue and suggests a solution.

In order to increase interpretability, the author has included two XAI approaches, LIME and Integrated Gradients, in addition to a deepfake detection system that is intended for low resource photos. This is especially important since explainability is essential to assuring the system's dependability and credibility. The author has made it simpler to comprehend the results by utilizing these methods.

7 Conclusion

This study concludes by demonstrating the difficulty of deepfake detection on low-resource photos. The findings of this study indicate that deepfake detection algorithms should be trained on substantial datasets to increase their precision. LIME and Integrated Gradients, two XAI approaches, make it simpler to understand the model's predictions and pinpoint the precise pixels in charge of the classification decision.

In the future, the model may be developed to include a multi-class classification method, allowing it to discriminate between various deepfakes, including face swaps, puppetry, and lip-syncing. As a result, the model may be more accurate and helpful in real-world settings. Further investigation could be conducted to examine additional XAI techniques for deepfake detection and how they can be incorporated into current models to enhance their interpretability. The research has also shown some limitations, such as data and hardware restrictions, as well as the need for more improvements, such multiclass classification, low-resource video deepfake detection, and lightweight model deployment. Despite these drawbacks, the author

has made a substantial contribution to the XAI and deepfake detection fields, as shown by the fact that the code and test results are available to the public.

Overall, this research offers insightful knowledge and advances the field of deepfake identification, particularly in identifying subpar deepfakes. Future scholars and practitioners in the fields of deepfake detection, computer vision, and XAI may find the author's conclusions and suggestions helpful.

REFERENCES

[1] T. T. Nguyen et al., "Deep Learning for Deepfakes Creation and Detection: A Survey," Comput. Vis. Image Underst., vol. 223, p. 103525, Oct. 2022, doi: 10.1016/j.cviu.2022.103525.

[2] K. Manke|, "Researchers use facial quirks to unmask 'deepfakes,'" Berkeley News, Jun. 18, 2019. https://news.berkeley.edu/2019/06/18/researchers-use-facial-quirks-to-unmask-deepfakes/ (accessed May 05, 2023).

[3] A. Pishori, B. Rollins, N. Houten, N. Chatwani, and O. Uraimov, Detecting Deepfake Videos: An Analysis of Three Techniques. 2020.

[4] P. Yu, Z. Xia, J. Fei, and Y. Lu, "A Survey on Deepfake Video Detection," IET Biom., vol. 10, no. 6, pp. 607–624, 2021, doi: 10.1049/bme2.12031.

[5] S. Agarwal and H. Farid, "Detecting Deep-Fake Videos from Aural and Oral Dynamics," in 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Jun. 2021, pp. 981–989. doi: 10.1109/CVPRW53098.2021.00109.

[6] G. Leonardo, A. Irene, D. B. Alberto, and C. Roberto, "Optical Flow based CNN for detection of unlearnt deepfake manipulations | Elsevier Enhanced Reader," Apr. 08, 2021. https://reader.elsevier.com/reader/sd/pii/S0167865521000842?token=DE0CE36FB1491D3C8935866228D8C68A879D5156E78473D28759FA58B821208387A6D7CFE906783DFC02CC35CB7DBD05&originRegion=eu-west-1&originCreation=20221017163750 (accessed Oct. 17, 2022).

[7] A. Mitra, S. P. Mohanty, P. Corcoran, and E. Kougianos, "Detection of Deep-Morphed Deepfake Images to Make Robust Automatic Facial Recognition Systems," in 2021 19th OITS International Conference on Information Technology (OCIT), Dec. 2021, pp. 149–154. doi: 10.1109/OCIT53463.2021.00039.

[8] H. D. Doshi, A. N. Savla, S. S. Dholakia, H. D. Gandhi, S. Suratkar, and F. Kazi, "Improving Deepfake Video Detection Using Data Augmentation Techniques," In Review, preprint, Aug. 2022. doi: 10.21203/rs.3.rs-1844392/v1.

[9] S. Yang, W. Xiao, M. Zhang, S. Guo, J. Zhao, and F. Shen, "Image Data Augmentation for Deep Learning: A Survey." arXiv, Apr. 18, 2022. Accessed: Oct. 18, 2022. [Online]. Available: http://arxiv.org/abs/2204.08610

[10] F. Matern, C. Riess, and M. Stamminger, "Exploiting Visual Artifacts to Expose Deepfakes and Face Manipulations," in 2019 IEEE Winter Applications of Computer Vision Workshops (WACVW), Jan. 2019, pp. 83–92. doi: 10.1109/WACVW.2019.00020.

[11] H. Khalid, S. Tariq, M. Kim, and S. S. Woo, "FakeAVCeleb: A Novel Audio-Video Multimodal Deepfake Dataset." arXiv, Mar. 01, 2022. Accessed: Dec. 12, 2022. [Online]. Available: http://arxiv.org/abs/2108.05080

A.1 Tested Models

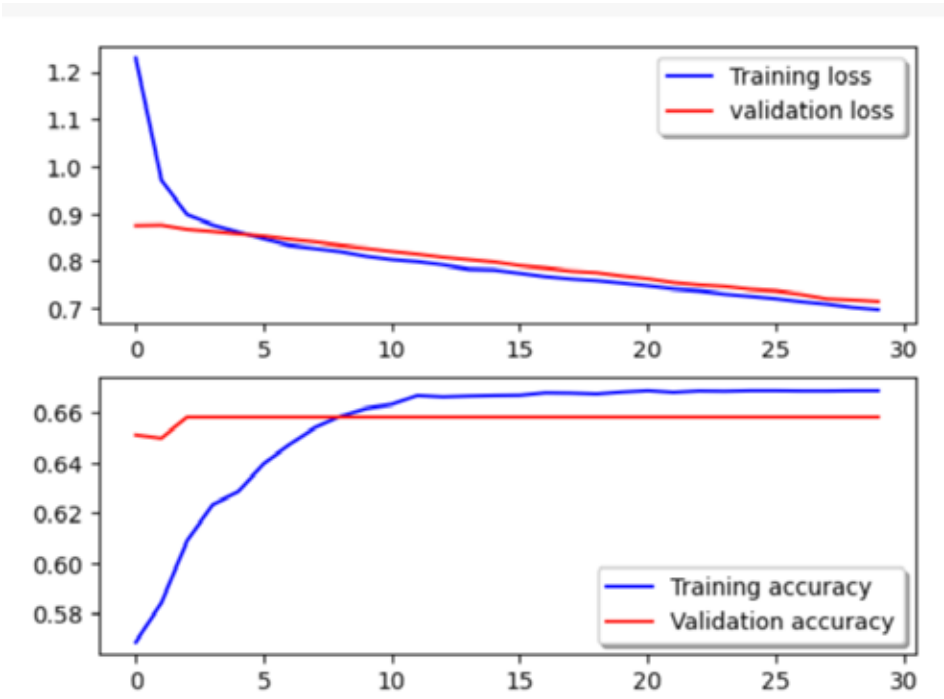


Figure 8 - EfficentNetB7 model

EfficentNetB7 Model

Resnet101 Model

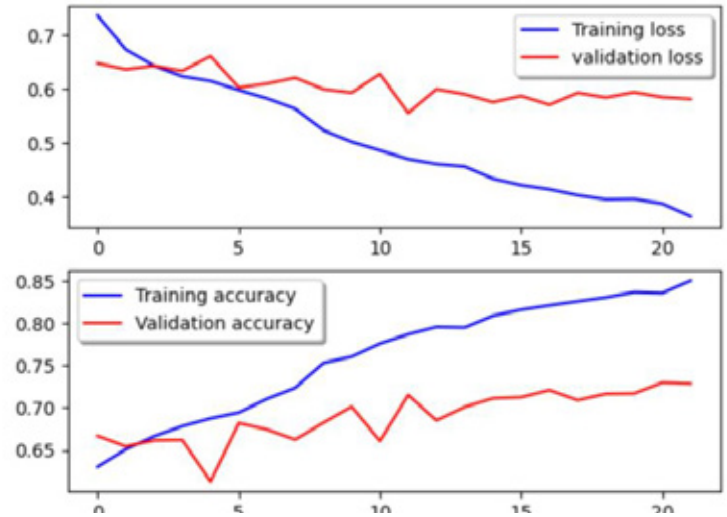


Figure 9 - Resnet101 model

Resnet50 Model

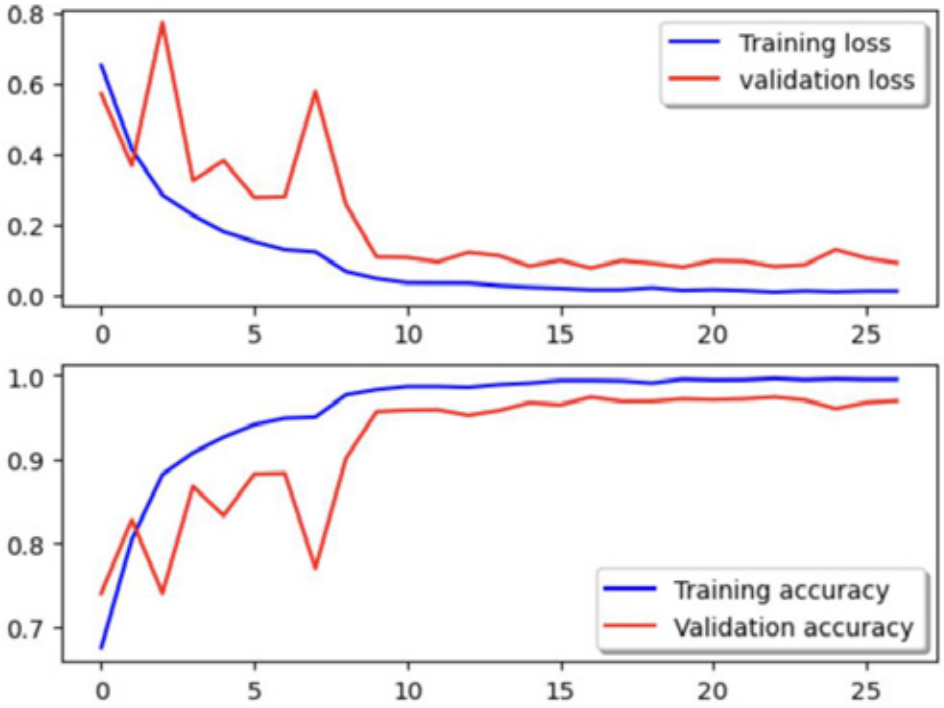


Figure 10 - resnet50 model

B - Architectural Design

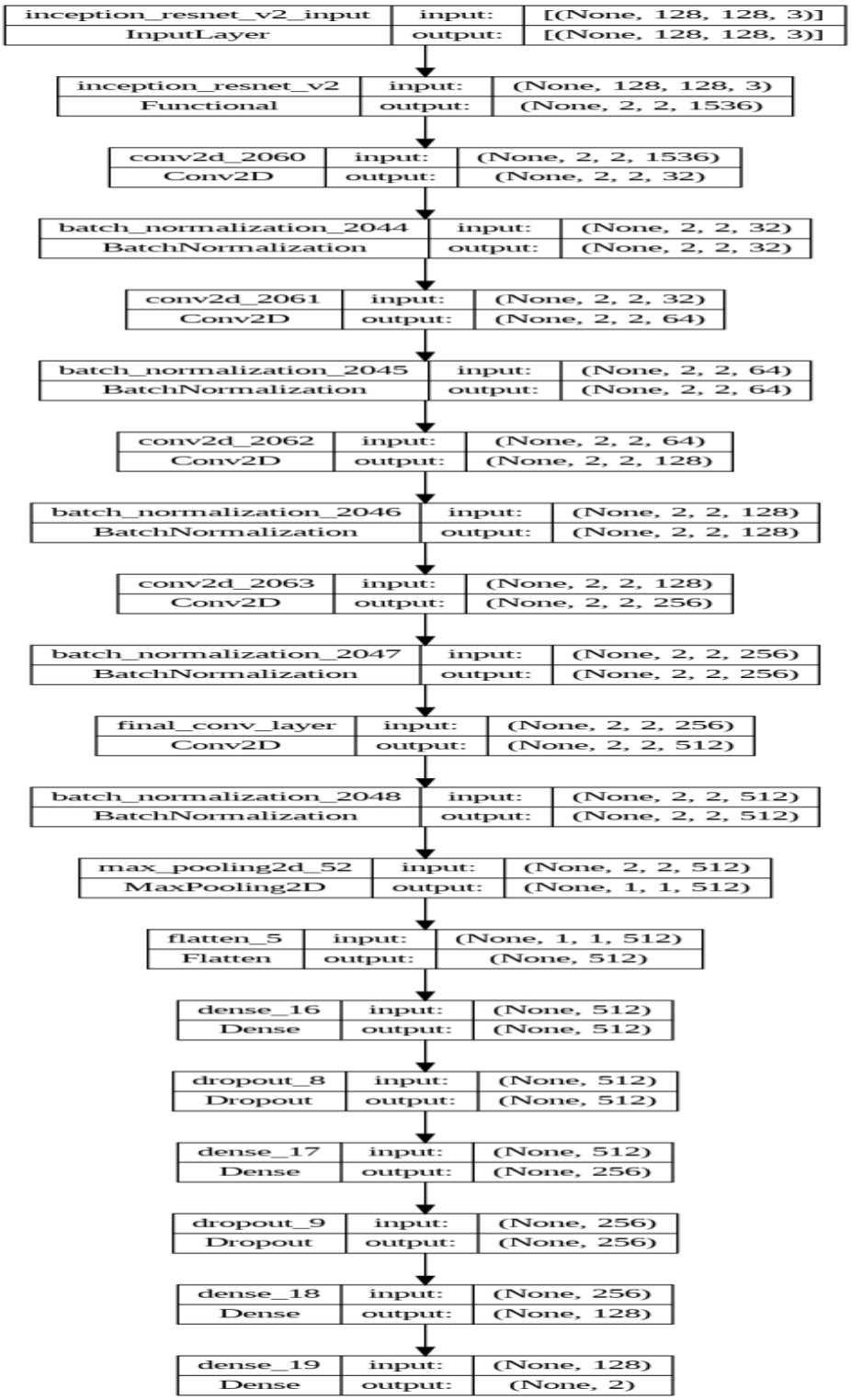


Figure 11 - CNN architecture

Sinhala Active Voice into Passive Voice Converter using Rule Based Approach with Grammar Error Correction

Oshini Ilukkumbura

School of Computing, Informatics Institute of Technology, Colombo 06, Western Province, Sri Lanka, aishcharya.2019757@iit.ac.lk

Sulochana Rupasinghe

School of Computing, Informatics Institute of Technology, Colombo 06, Western Province, Sri Lanka, sulochana.r@iit.ac.lk

ABSTRACT

In the world, more than 16 million people speak Sinhala, which is the official language of Sri Lanka. It is a morphologically complex language that evolved from Pali and Sanskrit. The Sinhala language is one of the few that has diglossia, where the spoken and written forms are different. Because it is governed by numerous grammar rules, Sinhala's written form is quite difficult. There are just a few implementations of Sinhala Grammar due to its intricacy and the limited availability of computerized resources.

Active to passive conversion in Sinhala is a crucial step in the processing of natural language since it shifts the subject and object of a phrase. However, it has been challenging to create an accurate and useful tool for Sinhala active to passive conversion due to a paucity of resources and study in this field.

The proposed system is capable of converting three-word-long sentences written in Sinhala active voice into Sinhala passive voice using a rule-based approach with an 85% accuracy rate, as that domain has received the most attention. The system develops as a learning application for students who are learning and practicing Sinhala active voice and passive voice related grammar rules, which are crucial to learn for the GCE O/L examination. Also, the proposed system is able to evaluate the grammatical correctness of the Sinhala sentences using a deep learning model with an overall 91.6% accuracy rate.

KEYWORDS

Sinhala Grammar, Sinhala Active voice into Passive Voice, Natural Language Processing, Deep Learning, Rule Based Approach

1 Introduction

Grammar is the basis for efficient communication. Grammatical errors can have a similar collision on the content and accessibility of communication as an inaccurately wired telephone wire can during a telephone conversation. Grammar norms are consequently crucial for the clarity of meaning and intention in daily life. Sinhala is widely spoken in Sri Lanka and serves as both the official and national language[4]. The language itself is complex, requiring strict adherence to multiple grammar rules for proper formal writing. Depending on the syntax and semantics employed, a word, phrase, sentence, or paragraph can convey vastly different meanings. Therefore, maintaining precision in the language is essential to ensure the intended meaning is accurately conveyed[1].

As a natural language, Sinhala is a complex & sophisticated language evolved from Pali & Sanskrit. The correctness of formal writing entirely rests on these clearly established norms when it comes to the written context, which is a complicated language with numerous spelling and grammatical restrictions. Depending on the syntax and semantics utilized, a word phrase, a sentence, or a paragraph may be understood to imply something different[1].

Due to complexity of the language there's only a few applications that can detect grammar errors on active voice sentences along with spelling error correction. The understanding of Sinhala active voice and Sinhala

passive voice among students who sat for GCE O/L exams is not satisfactory, according to the conclusions of the 2016, 2017, and 2018 G.C.E (O/L) Sinhala Language and Literature Evaluation Reports. It was advised to follow appropriate steps to improve knowledge. Currently, there is no way to convert a sentence written in Sinhala active voice into Sinhala passive voice sentence without manual processing.

This research implements a system that can be used to convert sentences written in Sinhala active voice into Sinhala passive voice sentences. By using this implementation students will be able to learn grammar rules and practice exercises to gain knowledge related to Sinhala passive voice. This could be achieved as a novel approach since there is no related implementation related to the conversion of Sinhala active voice into Sinhala passive voice. The prototype will be presented as a web application where the users can input a Sinhala active voice sentence which needs to be converted into Sinhala passive voice. This would be more beneficial for students to learn and practice Sinhala active voice into passive conversion-related grammar rules as a supportive tool for their education.

2 Background

2.1 Language Structures

Every language system's essential function is to connect meaning and expression in order to enable verbal expression for thought and feeling while also making that expression understandable to others[3]. Language structures deal with how words and sentences in a given language are arranged and put together to convey meaning. They include the patterns and laws governing word choice, sentence construction, and idea expression. Different languages have different language structures, which might include things like grammar, syntax, morphology, and semantics. According to the word order of the sentence English, Russian, Spanish and Chinese are SVO (Subject – Verb – Object) languages while Sinhala, Korean, Japanese, and Hindi are SOV (Subject – Object - Verb) sentences.

2.2 Sinhala Language

Sinhala is a contemporary Indo-Aryan language from a historical perspective[7]. As a result of its connections to Tamil, English, Portuguese, and Dutch, modern Sinhala has advanced[5]. Sinhala also includes an alphabetic writing system known as the abugida, which is a segmental writing system that records consonant-vowel sequences as single units[9].

2.3 Sinhala Language Structure

In Sinhala, which is a free order language, a sentence's primary components are its subject, verb, and object[8]. Figure 1 demonstrates all the possible English sentences with free word order “Mother hit the younger sister with a stick”.

I. අම්මා | නංගිට | කොට්ටයකින් | ගැහුවා ය.
Mother | to the younger sister | with a stick | hit

II. අම්මා | කොට්ටයකින් | නංගිට | ගැහුවාය.
Mother | with a stick | to the younger sister | hit

III. නංගිට | අම්මා | කොට්ටයකින් | ගැහුවාය.
To the younger sister | Mother | with a stick | hit

IV. නංගිට කොට්ටයකින් අම්මා ගැහුවාය.
To the younger sister | with a stick | Mother | hit

V. කොට්ටයකින් අම්මා නංගිට ගැහුවාය.
With a stick | Mother | To the younger sister | hit

VI. කොට්ටයකින් නංගිට අම්මා ගැහුවාය.
With a stick | To the younger sister | Mother | hit

Figure 1: Example for Structure of Sinhala Language

There are 25 distinct basic sentence constructions in Sinhala[9]. A sentence is often broken up into two parts: a noun phrase (NP) and a verb phrase (VP). The two components of a sentence in Sinhala grammar are Uktha

(Subject) and Akyatha (Predicate). In Sinhala sentences, the subject and predicate agree on the person, gender, and number[9].

2.3.1 Noun Phrase (NP)

A common noun, proper noun or a pronoun can all be used as the Noun Phrase, represented by NP. The Sinhala noun phrase also includes determiners and adjectival phrases along with head noun. The grammatical structure of Sinhala NP is extremely intricate. It may include a variety of clause types, including subordinate, adjectival and relative clauses[9]. Table 1 describes the breakdown of the noun phrase 'නොයෙක් කෙටි හරවත් දේශන / noyek keti harawath deeshana' to emphasize the structure of a Sinhala noun phrase.

Table 1: Example of Noun Phrase Structure

Noun Phrase 'නොයෙක් කෙටි හරවත් දේශන / noyek keti harawath deeshana' (some short meaningful lectures)		
Determiner 'නොයෙක් / noyek' (some)	Adjective Phrase 'කෙටි හරවත් / keti harawath' (short meaningful)	Noun 'දේශන/ deeshana' (lectures)
	Degree 'කෙටි / keti' (short)	Adjective 'හරවත්/ harawath' (meaningful)

2.3.2 Verb Phrase (VP)

A verb phrase (VP) is a group of words that begins with a verb and may include noun phrases and adverbial phrases (ADVP) in addition to the verb, according to generative grammar[9]. The verb base (krudantha) and the suffix make up the majority of the verb phrase's head. Different suffixes will result in verb forms that are morphologically distinct[6]. Table 2 describes the breakdown of the verb phrase 'ඉතා අලංකාරව නටයි/ itha alankarawa natai' to emphasize the structure of a Sinhala verb phrase. It is evident from this that adverbs in Sinhala are inserted before the verb. Additionally, a verb phrase may contain a noun phrase[9].

Table 2: Example of Verb Phrase Structure

Verb Phrase 'ඉතා අලංකාරව නටයි / itha alankarawa natai'' (dance very beautifully)		
Noun Phrase	Adverb Phrase 'ඉතා අලංකාරව / itha alankarawa' (very beautifully)	Verb 'නටයි / natai' (dance)
	Degree 'ඉතා / itha' (very)	Adverb 'අලංකාරව / alankarawa' (beautifully)

2.4 Sinhala Grammatical Voices

The active voice and the passive voice are the two most common types of sentences in the Sinhala language. Additionally, this can be viewed as a similarity between English and Sinhala. The traits of these two kinds are displayed below.

2.4.1 Sinhala Active Voice

Sentences using the active voice are by far the most common in the Sinhala language. Table 3 demonstrates a few examples to help explain this grammatical feature.

Table 3: Examples of Sinhala Active Voice Sentence

Sentence 1: ‘කොටියා සතුන් මරයි. / kotiya sathun marai.’ (Tiger kill animals.)		
Doer of the action ‘කොටියා / kotiya’ (Tiger)	Receiver of the action ‘සතුන් / sathun’ (animals)	Verb form: Present Tense, Active, Third person, Singular ‘මරයි / marai’ (kill)
Sentence 2: ‘මම ගෙදර ගියෙමි. / mama gedara giyemi’ (I went home.)		
Doer of the action ‘මම / mama’ (I)	Receiver of the action ‘ගෙදර / gedara’ (home)	Verb form: Present Tense, Active, First person, Singular ‘ගියෙමි / giyemi’ (went)

Regarding this type of voice, the following conclusions can be made.

- Doer of the action keeps always as active.
  In both examples, the doer in both sentences is in the active form of the noun. (‘කොටියා / kotiya’) & (‘මම / mama’)
- Receiver of the action keeps always as inactive.
  According to the given example, the receiver in both sentences is in the inactive form of the noun. (‘සතුන් / sathun’) & (‘ගෙදර / gedara’)
- The verb is inflected according to doer in a given sentence, taking on an active inflection form.
  In the first example the doer (‘කොටියා / kotiya’) is third person singular noun. Accordingly, the verb has included in third person, singular active form. (‘මරයි / marai’)
  The doer (‘මම /mama) in the second example is a first-person singular noun. The verb has been included in the first person, singular active form. (‘ගියෙමි /giyemi)
- The tense of a given sentence is chosen by the author and is independent of every other factor.

2.4.2 Sinhala Passive Voice

In contrast to active voice, passive voice is less common in Sinhala. The examples included in table 4 will provide more information on the grammatical characteristics of passive voice in Sinhala sentences.

Table 4: Examples of Sinhala Passive Voice Sentence

Sentence 1: ‘කොටියා විසින් සත්තු මරනු ලබති. / kotiya wisin saththu maranu labathi’ (Animals are killed by tigers.)		
Doer of the action ‘කොටියා / kotiya’ (Tiger)	Receiver of the action ‘සත්තු / saththu’ (animals)	Verb Form: Action, Present Tense, Third Peron, Plural ‘මරනු ලබති / maranu labathi’ (are killed)
Sentence 2: ‘ඔවුන් විසින් අපි අමතනු ලදිමු. / owun wisin api amathanu ladimu.’ (We were contacted by them.)		
Doer of the action ‘ඔවුන් / owun’ (them)	Receiver of the action ‘අපි / api’ (we)	Verb Form: Action, Past Tense, First Peron, Plural ‘අමතනු ලදිමු / amathanu ladimu ’ (were contacted)

Regarding this type of voice, the following conclusions can be made.

- Doer of the action keeps always as inactive.
  In both examples, the doer in both sentences is in the inactive form of the noun. (‘කොටියා / kotiya’) & (‘ඔවුන් / owun’)
- Receiver of the action keeps always as active.
  According to the given example, the receiver in both sentences is in the active form of the noun. (‘සත්තු / saththu’) & (‘අපි /api)
- The verb is inflected according to receiver in a given sentence, taking on a passive inflection form.
  In the first example the reciever(‘සත්තු /saththu) is third person plural noun. Accordingly, the verb is included in third person, plural passive form. (‘මරනු ලබති / maranu labathi’)
  At the second example the reciever (‘අපි /api) is first person plural. According to that the verb has included in first person, plural passive form. (‘අමතනු ලදිමු /amathanu ladimu)
- The tense of a given sentence is chosen by the author and is independent of every other factor.

3 Related Work

The elaborate design of multiple spelling and grammar rules complicates the Sinhala language[12]. Even though it is challenging to put all of these principles into practice in a certain order, follow all grammar rules in one application[9]. Despite the difficulties mentioned above, there have been only a few efforts in recent years to develop automated grammar checkers for Sinhala.

3.1 Sinhala Grammar Error Correction

The University of Moratuwa's " මහාරාවණා" project has produced a basic grammar checking scenario for a sentence of just three words[4]. Due to the extensive use of the Sinhala language and the fact that the system is ruled-based, it is exceedingly challenging to make future changes. For the time being, the Sinhala language lacks a proper lexical analyzer.

The presented Sinhala spell & grammar checking tool to ensure that a context written in the Sinhala language is grammatically and spelled correctly as the first application which can perform both functionalities in the same application. The spell error checking component adopted a data driven methodology to overcome the challenges encountered by current spell-checking software and uses a rule-based lexicon analysis methodology to develop a novel method for ensuring grammatical accuracy. Based on these phrase kinds, lexicons have been implemented. To construct the lexicon, six types of noun phrases, eleven types of verb phrases, four types of adjective phrases, and nine types of adverb phrases were identified. The grammar was created using ten basic sentence patterns that were chosen. As anticipated, the neural network mathematical model did not produce precise findings. The percentage value behaved inconsistently with the intended accuracy range. Only the numerical pattern of correct and incorrect sentences was taught to the model; it did not learn the Sinhala language. Since there is no clear end to the implementation's use of separate numbers for special Sinhala words, self-learning is impossible. As a result, this method's accuracy did not meet expectations[1].

The Grammatical error detection computational model to detect grammatical errors in simple active sentences in Sinhala text. The researchers analyzed the sentences word-wise and noun/verb-wise using a part-of-speech and morphological analyzer. They selected Sinhala sentences from online newspapers with no spelling errors and 3-5 words. The data was parsed to a decision tree trained with a dataset of 200 records and 22 predefined Sinhala grammar rules, achieving an accuracy score of 84.61%. If the grammar was incorrect, a neural network-based sentence correction algorithm was used to suggest a correction by adding a suitable verb suffix. This approach provides a hybrid method for pointing out grammatical errors & suggesting corrections in written Sinhala sentences which is a good solution to cover a considerable amount of scope in Sinhala grammar error correction with a good accuracy rate[10].

A Grammar error correction and detection module using rule-based approach as a component of a Sinhala language automated essay-checking system for students in primary schools. Particularly for free word order, fairly difficult grammatical rules, and novelties in Sinhala grammar, it detects grammar faults, identifies error types, and offers remedies with good accuracy rates. The module covers the use of the present and past tenses, singular and plural numerals, male and female pronouns, and the 1 st, 2 nd , and 3 rd person pronouns in sentences with the active voice and suggests grammar errors according to both subject and object. The grammar rules for vibakthi (case), however, and the context's significance are not covered. When the accuracy rates of the module are compared to those of other projects, it is discovered to have a high degree of accuracy[11].

The proposed automated system for Sinhala spelling and grammar checking is composed of four main components: preprocessing and analysis of spelling errors, detection and analysis of the subject and tense, analysis of subject characteristics in sentences, and grammar analysis based on rules. The system uses a customized dataset that includes the most frequently used Sinhala words since publicly available datasets for Sinhala language are limited. The system combines rule-based and machine learning algorithms to handle preprocessing, morphological and lexical analysis, subject analysis, and tense analysis for accurate results. The rule-based grammar analysis incorporates a novel hybrid approach that uses both rule based and ML algorithms. Overall, the proposed methodology aims to provide an accurate and efficient automated system for detecting & correcting spelling and grammatical errors in the Sinhala language but in the research paper they haven’t mentioned about the accuracy percentages and the details of relevant evaluation approaches to get a clear idea about the performance of the implementation[4].

These prior studies exhibit notable advancements in the realm of Sinhala language spell checking and grammar correction. While certain approaches emphasized data-driven and neural network methodologies, others adopted hybrid techniques that combined rule-based and machine learning strategies. Each research

endeavor has played a role in improving the precision and effectiveness of processing Sinhala text, although their scopes and constraints differ. Nevertheless, there are still untapped opportunities for further research and development to tackle the intricacies of the Sinhala language such as grammar error correction of Sinhala passive voice sentences with considerable accuracy conversion, including handling more intricate grammatical rules and broadening coverage to encompass a wider array of sentence structures and tenses.

3.2 Active Voice into Passive Voice Conversion of a Natural Language

That proposed system performs the basic conversion of active voice into passive conversion of English language using LL (1) parsing technology after identifying that the less number of works has done related to English active voice into passive conversion. To make sure that the input follows the required syntax, the system is built using predetermined rules that are conveyed to the user. The system validates the syntax against the specified rules once the user enters the text. If the input complies with the rules, the system checks for the verb in the sentence and transforms it into passive voice by experiencing the stated steps one at a time. Rule based approach is a great way to start a implementation that hasn't been done previously. It is so much better if they have mentioned the accuracy of the proposed system & the evaluation approaches that have been used to assess the system's performance[2].

While the proposed system introduces a novel approach to English active voice to passive voice conversion, its lack of information about accuracy and evaluation results is a notable limitation. On the other hand, the existing works in Sinhala language spell checking and grammar correction have made significant progress in improving the accuracy and efficiency of processing Sinhala text. To further advance the proposed system, future research should focus on providing a comprehensive evaluation of its performance and exploring the integration of other linguistic and machine learning techniques to enhance accuracy and coverage.

4 Research Methodology

This research aims to design, implement, and evaluate a system that is able to convert Sinhala active voice sentences into Sinhala passive voice sentences as an aid for students to learn and practice Sinhala passive voice related grammar rules along with Sinhala grammar error correction for both sentences written in Sinhala active voice and Sinhala passive voice.

The system contains 2 major components as follows:

- Sinhala grammar error correction
- Sinhala active voice into passive voice conversion

Moreover, figure 2 & figure 3 illustrate the architecture and system process flow of the proposed system.

4.1 Sinhala Grammar Error Correction

4.1.1 Dataset Selection

Due to the less availability of Sinhala grammar related datasets the author was able to find a dataset to implement the grammar error correction model from a past student at the university who also worked on Sinhala grammar error correction. That dataset consists of 154,014 sentence pairs as grammatically incorrect sentences along with grammatically correct sentences. This includes Sinhala active voice sentence pairs as well as Sinhala passive voice sentences.

4.1.2 Implementation of Grammar Error Correction Model

The deep learning model employs TensorFlow as the deep learning framework and Keras as a high-level API for building and training neural networks, ensuring efficient model development and seamless integration with other libraries. The use of LSTM (Long Short-Term Memory) addresses the vanishing gradient problem in sequential data processing, enabling the model to capture long-range dependencies in input sequences. The Time Distributed layer allows the Dense layer to be applied to each time step of the LSTM output, ensuring effective handling of sequential data. Additionally, the Embedding layer creates word embeddings, representing words in continuous vector spaces, which aids in capturing semantic relationships between words. The inclusion of Bidirectional LSTM enables the model to process input sequences in both forward

and backward directions, enhancing the ability to gather contextual information. Moreover, the model leverages the attention mechanism, allowing it to focus on relevant parts of the input sequence during translation, enhancing translation accuracy and context preservation.

4.2 Sinhala Active Voice into Passive Voice Conversion

4.2.1 Dataset Selection

To develop a data science-related project, it is pivotal to have a proper dataset. To implement the model that can convert Sinhala active voice sentences into passive voice there wasn't an available dataset. Therefore, the author has created a dataset by including Sinhala active voice sentence and the passive voice sentence of it as a pair. To create that dataset the author had to look deep into the grammar rules that are going to cover through this implementation and got the help of Sinhala teachers to create and evaluate the dataset.

4.2.2 Implementation of Sinhala Active Voice into Passive Voice Conversion Model

In this implementation, the author has used a library called "Sinling" to perform POS tagging and tokenization. Moreover, it has used a Morphological analyser tool implemented by the Natural Language Processing Center of the University of Moratuwa.

As the first step, the system reads the data columns separately from the dataset and preprocesses them before using them in the Sinhala active voice into passive voice conversion. The system has the capability to provide relevant outputs for a list of Sinhala active voice sentences within the predefined scope. Considering the process of an input sentence at a time, first, the system pos tags the tokenized sentence. Since the scope has narrowed down to consider a three-word sentence first, the system validates the size of the input sentence. To perform the active voice into passive conversion, the words of the input sentence must consist of SOV structure. There's a function that takes a part-of-speech tagged sentence as input and checks if the sentence conforms to a specific structure required for converting an active voice sentence to a passive voice sentence in the Sinhala language. After the validation process, the identifies subject, object, and verb of the sentence passes to the relevant functions for the conversion process.

4.2.2.1 Subject Conversion

If the identified subject of the input sentence is a noun or preposition, it passes to the function which performs the subject conversion. Since there's no dataset or an existing method to achieve the Sinhala *Uktha Anuktha* conversion of nouns and prepositions, the author has created a data dictionary to perform that conversion.

4.2.2.2 Object Conversion

If the identified object of the input sentence is a noun or preposition, it passes to the function which performs the object conversion. For this conversion, also make use of the earlier mentioned data dictionary.

4.2.2.3 Verb Conversion

If the input sentence's identified verb belongs to the verb category, it passes to the function which performs the verb conversion. Using the morphological analyser the system takes the morphological details of the input verb. The morphological information can identify the tense and have the verb's stem. The morphological analyser can only determine whether the input verb belongs to the past or non-past tense. Because of that, the author has defined rules to identify whether that verb belongs to the present or future tense. To perform verb conversion, it is necessary to identify the person, number, and gender of the object. To perform that there's a function to analyze the object with the help of the morphological analyser and also some algorithms built using rule-based approach. Using the identified verb stem, verb tense, object person, object number & object gender can take the relevant part to join with the verb stem. It also provides the details of used grammar rules with some examples to display on the user interface as additional details. By using the word joiner of "Sinling" library, the passive form of the verb is created by joining the identified verb stem and the relevant part to join with the verb stem.

After performing all three conversion processes, the system generates the relevant output.

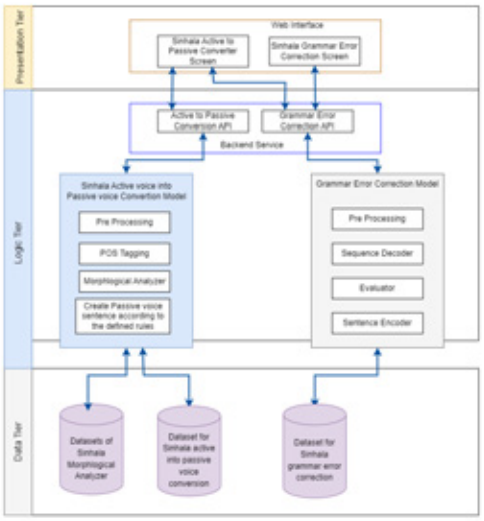


Figure 2: Three-Tiered Architecture of the proposed system



Figure 3: System Process Flow Chart

5 Results and Discussion

5.1 Sinhala Active voice into passive conversion model Testing

Since the model is based on a rule-based approach the testing was conducted using manual testing. A test data set with 300 pairs of Sinhala active voice and Sinhala passive voice sentences were used to perform testing phase.

Table 5 demonstrates the method that was used to check the accuracy of the model.

Table 5: Accuracy on Active to Passive Conversion Model

Number of Inputs	Number of Corrected Inputs	No of Incorrect Inputs	Percentage
300	225	45	85%

5.2 Sinhala Grammar Error Correction Model Testing

The above figure shows the overall accuracy of the deep learning model.

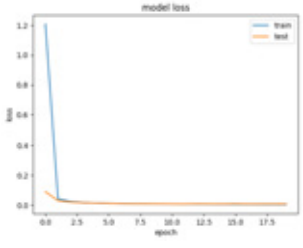


Figure 4: Grammar Error Correction Model Accuracy

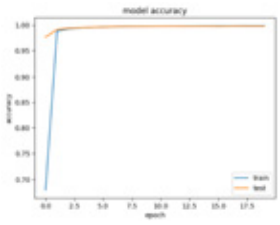


Figure 5: Grammar Error Correction Model Loss

Figure 4 & figure 5 present the accuracy and model loss of the grammar error correction model respectively. Furthermore, a collection of 250 randomly selected grammatically correct and incorrect sentences were

gathered, sourced from a Sinhala grammar book for grades 4 and 5. The focus of this process was on frequent language scenarios, and random sentences were chosen to test each case.
The grammar error correction accuracy of the system was evaluated by comparing its grammatical corrections of the sentences to the actual corrections, and then calculating the overall accuracy of the system. CP – Correct Predictions | TP – Total Predictions | TA – Total Accuracy

$$TA = (CP/ TP) * 100$$
$$= (216/250) * 100$$
$$= 86.40\%$$

6 Limitations

The identified limitations related to Sinhala active voice into passive voice conversion are as follows.

- The accuracy of the model was only tested using the created dataset by the author.
- Performs only for Sinhala active voice sentences with 3-word length.
- Since that conversion model is based on rule-based approach the scope that system can perform is narrowed down into a defined scope.
- The dataset used to test the model includes around 500 sentence pairs since the author had to filter the created dataset according to the considered scope.
- Since the author uses a morphological analyser tool for verb stemming the conversion of the verb is dependent on the performance and the scope of the morphological analyser.

And the identified limitations related to the Sinhala Grammar Error Correction deep learning model are mentioned below.

- The prototype was designed to handle grammatical structures that contained a maximum of 8 words. While there was an effort to expand this capability, it was not feasible to achieve within the given timeframe.
- The accuracy of the Sinhala passive voice grammar error correction is bit less than the accuracy of the Sinhala active voice grammar error correction since the number of Sinhala passive voice sentence pairs included in the used dataset is considerably low.
- Out of the complete vocabulary, only 187,879 words are included in this.

7 Conclusion and Future Work

Since the existing applications are limited to spelling and grammar error correction, the author has introduced a novel implementation to cover the existing research gap related to Sinhala active voice into passive conversion. Also, the author improved the grammar error correction module to identify grammatical errors in Sinhala passive voice sentences. And the author has created a dataset including around 1000 Sinhala active voice and passive voice sentence pairs. That would be further elaborated into a large dataset to experiment with other approaches to reach better performance. Furthermore, the author has maintained a data dictionary which also can improve to capture the uktha anuktha variation of Sinhala nouns, which is highly important in Sinhala active voice into passive conversion.
As future enhancements, the capability of Sinhala active voice into passive voice conversion model can enhance by creating a large dataset and experiment with other approaches. Also, it can improve the conversion model to provide outputs of more than three-word length sentences by capturing complex grammar rules. Improving the accuracy of Sinhala Passive voice grammar error correction and assistance in identifying the specific grammatical feature responsible for the grammatical error are some future works to include in the grammar error correction module. Improve the system to convert the user input, which is a sentence written in Sinhala passive voice, into Sinhala active voice (Reverse process of the proposed system), include spelling error checker and suggest the correct word to replace the incorrect word, take user inputs as images and give the output as another user input option, take user inputs as audio clips and give the output as another user input option and enhance the proposed system's coverage of basic grammar rules included in

the GCE O/L syllabus, it is recommended to integrate grammatical error checking for nonverbal sentences, which includes sentences with vibhakthi (case), apara kriya, and puurwa kriya. In addition to the spell-checking functionality, the system should be able to identify the type of error and provide more precise suggestions for both subject and verb directions.

ACKNOWLEDGMENTS

The author would like to express his sincere appreciation and recognition to the project supervisor, all the numerous language experts, and Sinhala teachers from various schools. Special gratitude is extended to the UOM NLP center for granted access to the "SinMorphy" morphological analyser included in the author's research project.

REFERENCES

[1] Abeyrathne, L. et al. 2018. Spell and Grammar Checking Tool for Sinhalese අකුරු ස ෝදුව-සියඹ සුඹ ක් කරනු රිසිසයනි. (2018).
[2] Fahad, M. and Beenish, H. 2020. An Approach towards Implementation of Active and Passive voice using LL(1) Parsing. 2020 International Conference on Computing and Information Technology, ICCIT 2020. (Sep. 2020). DOI:https://doi.org/10.1109/ICCIT-144147971.2020.9213802.
[3] Finegan, E. 2014. Language: Its Structure and Use - Edward Finegan - Google Books. Cengage Learning.
[4] Goonawardena, M. et al. 2022. Automated Spelling Checker And Grammatical Error Detection And Correction Model for Sinhala Language. 2022 International Research Conference on Smart Computing and Systems Engineering (SCSE). (Sep. 2022), 184–189. DOI:https://doi.org/10.1109/SCSE56529.2022.9905126.
[5] Gunasekara, A.M. 2008. A Comprehensive Grammar of the Sinhalese Language - Abraham Mendis Gunasekara - Google Books. Godage International Publishers (PVT) Ltd.
[6] Hettige, B. and Karunananda, A.S. 2011. Computational model of grammar for English to Sinhala Machine Translation. 2011 International Conference on Advances in ICT for Emerging Regions (ICTer). (Nov. 2011), 26–31. DOI:https://doi.org/10.1109/ICTER.2011.6075022.
[7] Jayawardana, T. 1989. The Surface Case System in Sinhala. Journal of Humanities & Social Sciences, University of Kelaniya. (1989), 264–277.
[8] Kariyakarawana, S.M. 1998. The syntax of focus and wh-questions in Sinhala. Karunaratne & Sons Limited.
[9] Liyanage, C. et al. 2012. A computational grammar of Sinhala. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 7181 LNCS, PART 1 (2012), 188–200. DOI:https://doi.org/10.1007/978-3-642-28604-9_16/COVER.
[10] Pabasara, H.M.U. and Jayalal, S. 2020. Computational model for detecting grammatical mistakes in sinhala text. researchgate.net. (2020).
[11] Sewwandi, U. et al. 2021. A Rule Based Approach for Detection and Correction of Grammar Errors in Written Active Voice Sinhala Sentences. 21st International Conference on Advances in ICT for Emerging Regions. (Dec. 2021), 159–164. DOI:https://doi.org/10.1109/ICTER53630.2021.9774800.
[12] Weerasinghe, R. et al. 2008. NLP Applications of Sinhala: TTS & OCR. ACL Anthology - Proceedings of the Third International Joint Conference on Natural Language Processing: Volume-II. (2008).

A APPENDICES

A.1 Output of Sinhala Active to Passive Voice Conversion



Figure 6: Output of Sinhala Active to Passive Voice Conversion

A.2 Output from Sinhala Grammar Error Correction Model

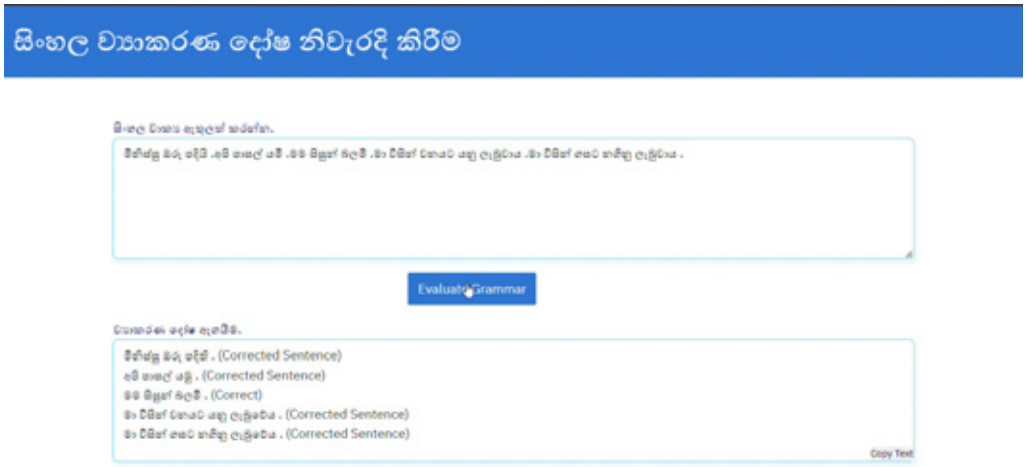


Figure 7: Output from Sinhala Grammar Eror Correction Model

HealthySmyl: Analysis and Recommendations Based Lifestyle Enhancing Application for HIV Patients

Selshia Sbanumi Regis Sagayaratnam
Business School, Informatics Institute of Technology, Colombo 06, Sri Lanka, selshia.2019727@iit.ac.lk

Kumudini Sarathchandra
Business School, Informatics Institute of Technology, Colombo 06, Sri Lanka, kumudini.s@iit.ac.lk

ABSTRACT

The Human Immunodeficiency Virus (HIV) is a highly contagious, fatal virus that causes Acquired Immuno Deficiency Syndrome (AIDS), a condition in which the body's immune system weakens. It remains a serious global health concern, with thousands of individuals dying from the disease every year. Maintaining a healthy lifestyle that includes a regular workout routine, nutritional food patterns, positive thinking, and constant encouragement aids in better management of the disease. It is challenging for anybody to succeed in maintaining a suitable lifestyle immediately after being diagnosed with a disease like HIV/AIDS. The mental impact of daily medication and lifestyle adjustments brought on by treatment can be difficult for those who have been diagnosed with the disease. This necessitates the development of a solution that can assist patients in improving their post-diagnosis lifestyle.

The purpose of this study is to discover the patients' pain points and develop a solution to address their challenges in following a healthy lifestyle post-diagnosis of HIV/AIDS. The research proposes 'HealthySmyl' - a mobile application-based solution that includes a nutrition analyzer based on image categorization, a workout recommendation system, a medication tracker, and other beneficial features. 'HealthySmyl' seeks to make it more convenient for newly diagnosed patients to adopt the suggested lifestyle, while also supporting those who are currently receiving medications.

KEYWORDS

HIV/AIDS, Image Classification, Image recognition, Exercise Recommendation, Food Recommendation, Lifestyle Management, Medication Tracker

1 Introduction

The Human Immunodeficiency Virus (HIV) continues to pose a significant threat to global public health, with millions of people living with the virus. People Living with HIV/AIDS (PLWHA) must continue to maintain a healthy lifestyle in order to manage the disease and improve their overall health, despite remarkable advances in HIV treatment and care. However, numerous obstacles hinder their ability to maintain a healthy lifestyle effectively.

In pursuit of a healthy lifestyle, PLWHA encounters a multitude of interdependent obstacles. These obstacles include following special dietary requirements and restrictions, understanding the importance of regular exercise and engaging in it, difficulties in maintaining good mental health, possessing a limited knowledge about disease management, difficulties in maintaining regular sleep patterns, difficulties in tracking complex medication schedules, and difficulties adjusting to the new lifestyle after diagnosis. Each of these obstacles has a substantial impact on the physical, mental, and emotional health of PLWH. To comprehensively resolve these complex challenges, a holistic and multidisciplinary approach is required. This research paper proposes a comprehensive solution for HIV patient care that emphasizes multiple aspects. It includes innovative

technological interventions such as nutrition tracking, medication adherence monitoring, personalized meal recommendations, exercise guidance, knowledge dissemination initiatives, and other supportive features intended to help PLWHA adjust to their new lifestyle, through the aid of an application. It is essential to collaborate with healthcare professionals, such as physicians, nurses, dietitians, nutritionists, and fitness experts, to develop and implement practical solutions tailored to the unique requirements of PLWHA. The proposed approach seeks to provide practical solutions and resources to empower PLWHA in their journey towards a healthier lifestyle by leveraging the expertise of these professionals.

The nutrition tracking component of the solution entails the creation of an application that enables PLWHA to trace their nutritional intake and make informed dietary decisions. Utilizing cutting-edge technologies such as image recognition, the application provides accurate and real-time analysis of food items, allowing PLWHA to ensure they are meeting their nutritional requirements and adhering to any dietary restrictions. In addition, the medication adherence monitoring feature assists PLWH in effectively managing their complex medication schedules. The application functions as a medication reminder system, sending notifications and providing alerts to ensure accurate and timely medication consumption. This feature also facilitates the monitoring of medication adherence over time, allowing healthcare providers to identify and address any adherence issues. The solution's individualized meal recommendation and exercise guidance components seek to assist PLWHA in adopting a healthy lifestyle. The application offers customized meal plans and exercise routines that promote optimal health and well-being by taking into account the user's dietary preferences, nutritional needs, and physical capabilities. In addition to promoting a healthy lifestyle, these characteristics contribute to the overall management of HIV and its associated complications. Apart from the above-mentioned components, the proposed solution emphasizes knowledge dissemination initiatives to improve PLWHA's understanding of their condition and equip them with accurate information about disease management, mental health, a balanced diet, and other crucial aspects of maintaining a healthy lifestyle.

By addressing the identified obstacles and implementing this multifaceted solution, it is anticipated that PLWHA will be better equipped to manage their condition, enhance their quality of life, and reduce the disease's associated risks. This comprehensive approach emphasizes the importance of collaboration between healthcare professionals, the effective use of innovative technologies, and the provision of ongoing support to empower PLWHA on their path to a healthier and more fulfilling life.

2 Background

2.1 An Overview of HIV in Sri Lanka

The Human Immunodeficiency Virus (HIV) saga has been illuminated by a few revolutionary scientific advances over the past four decades. HIV is a fatal, highly contagious virus that has claimed a large number of lives across the globe, with no known definitive cure as of yet. As of 2021, there were 38.4 million persons living with HIV/AIDS (PLWHA) worldwide, with 650,000 deaths [33]. HIV can be transmitted in two different ways. The first method is horizontal transmission, which occurs when a person comes into direct contact with the blood, sperm, rectal fluids, vaginal fluids, or breast milk of an HIV-positive patient with detectable viral load. In order for transmission to occur, the HIV in these fluids must enter the circulation of an HIV-negative individual via a mucous membrane, open cuts/sores, or direct injection from a needle or syringe. The second transmission mechanism is vertical. Vertical transmission of HIV is possible between infants born to HIV-positive mothers, who can become infected during pregnancy, in the birth canal, or during lactation [10,25].

Sri Lanka has a lower HIV prevalence as compared to global figures, where 1.5% of 2019 deaths were caused by the infection. 227 new cases of HIV infection were discovered in Sri Lanka that same year, and 0.17 per 100,000 persons died as a result of the virus [21]. This is lower than the majority of South Asian nations. It is estimated that 70% of individuals living with HIV in the country are aware of their HIV status [17].

In 2022, a sharp increase in the number of patients was observed [18,27]. During the first three quarters of 2022, the number of newly infected individuals has more than doubled, and the total number has increased by nearly 48% compared to the previous year. Concerns have increased regarding the number of HIV-positive cases among adolescents (18-30 years), as 50 new cases were identified in this age group during the [4].

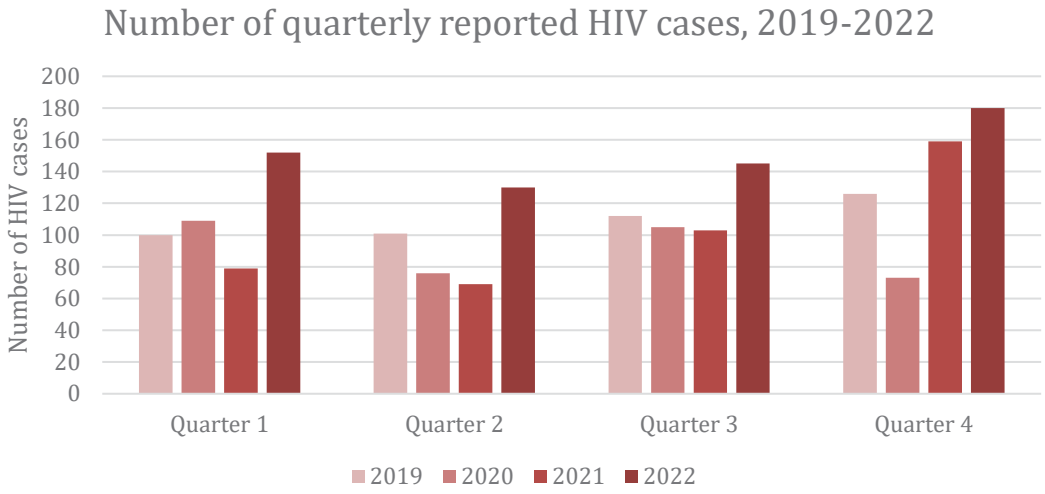


Figure 1: Quarterly Reported HIV cases in Sri Lanka from 2019 - 2022

The graph in Figure 1 depicts the number of quarterly HIV cases reported from 2019 to 2022, with an evident upward trend. Some Key Population (KP) groups continue to engage in hazardous behavior, according to the available evidence. Sri Lanka has the potential for such exposure due to the country's escalating internal and external migration, which has resulted in a flourishing tourism industry and a rise in the number of foreign tourists. And a rise in the use of substances like Crystal Methamphetamine, also known as ICE, has led to an increase in chemsex among adolescents [20,28].

2.2 Challenges of Living with HIV

Since the day of their diagnosis, PLWHA encounter a variety of obstacles daily. ART (Antiretroviral therapy) could significantly suppress the spread of HIV. Despite ART and ongoing virological suppression, however, many HIV-infected individuals still struggle to recover their immune systems.

Multiple physical obstacles exist for PLWHA, including poor concentration, fatigue, shortness of breath, muscle weakness, and joint and muscle pains. Due to the compromised immune system, opportunistic infections also pose a significant challenge. Moreover, wasting syndrome and unintentional weight loss are major concerns. In contradiction, obesity and overweight are currently medical concerns among PLWHA as well. This may contribute to other diseases such as cardiovascular disease, diabetes, and even certain types of cancers. Stress, depression, and anxiety are also experienced by these patients [1,16,29].

Patients also face obstacles related to social stigma and public acceptance. Numerous victims have reported incidents in which hospitals refused to treat HIV patients for other health concerns due to their HIV condition. Some patients have reported being mistreated in specialized facilities while receiving treatment for the disease itself [11,14,22]. In Sri Lanka, there have been instances in which neighbors have stoned or set fire to the homes of HIV patients and discriminated against their children by refusing to admit them to schools. In addition, it is extremely difficult for these patients to obtain suitable employment, resulting in severe financial difficulties. The public's lack of knowledge and awareness is a significant obstacle for PLWHA to live in harmony with society [12,30]. In addition to it the lack of knowledge found among PLWHA itself is another major challenge.

3 Methods for Better Management of HIV

3.1 Dietary Requirements and Restrictions

Immunodeficiency and adverse health outcomes result from a vicious cycle involving HIV infection and inadequate dietary consumption. Healthy diet lays the framework for people to live well with HIV infection, boosting immune function, antiretroviral medication efficacy, and general quality of life. Those with HIV who are well-nourished are more resistant to illness and more capable of performing daily tasks than those who are malnourished. For the medication to function properly, a robust nutritional status is required. In a limited scientific study, it was observed that daily micronutrient consumption increased body weight and mass, decreased HIV RNA levels, improved CD4 counts, and decreased the incidence of opportunistic infections.

The energy requirements of HIV-positive individuals are 20%–30% greater than those of their age and gender-matched unaffected counterparts. A man with HIV should strive for 100-150 grammes of protein per day, while a woman with HIV should aim for 80-100 grammes per day. However, the ratio of protein to total caloric should not exceed 15 to 20 percent. This is crucial for preventing nutritional recovery syndrome. Sodium(salt) and sugar consumption should also be limited [5]. A patient's intake of fat and oil should also be kept to a minimum. The patient should consume 30% of their daily calories from fat in order to ingest sufficient amounts of the appropriate fat. 10% or more of daily calories should come from monounsaturated fat-rich foods, such as almonds, seeds, avocado, fish, canola and olive oils. Less than 10% of daily calories should come from polyunsaturated lipids. Examples include fish, hazelnuts, flaxseed, corn, sunflower, soybean, and safflower oil. The daily caloric intake of saturated lipids such as fatty meat, poultry with skin, butter, whole-milk dairy foods, and coconut and palm oils should be less than 7% [7,26,31].

Vitamins A (found in dark green, yellow, orange, or red vegetables and fruit, liver, whole eggs, and milk), B (found in fish, chicken, nuts, grains, avocados, broccoli, and green leafy vegetables), C (found in citrus fruits), D (found in cod liver oil, beef liver, tuna, and salmon), and E (found in green leafy vegetables, vegetable oils, and peanuts), regulate the pathogenesis of HIV infection. In addition, anemia and iron deficiency are common among PLWH. However, both low and high iron reserves have been associated with a higher risk of disease progression and other adverse outcomes. To reach a definitive conclusion, additional research is necessary. However, randomized trials have shown that iron supplementation in conjunction with ART has a positive effect on the progression of HIV disease and related health outcomes. [2,9,19].

Consumption of water is another crucial aspect of healthy existence, particularly for PLWHA. Women should consume approximately 2 liters/8 cups of water per day, while males should consume approximately 2.6 liters/10 cups. This is necessary for proper digestion, preventing dehydration, and delivering oxygen and nutrients to cells. and to attain numerous additional health benefits [6]. HIV patients are also advised to avoid the consumption of raw eggs, foods that contain raw eggs (such as homemade cookie dough), raw or undercooked poultry, meat, and seafood, and unpasteurized milk or dairy products to avoid food-borne illnesses that may cause an adverse effect on the overall health [23].

3.2 Exercise and Physical Activities

With optimal ART use, the lifespan of PLWHA nearly tripled. Likewise, their physical changes are varying as well. As a result of the toxic side effects of the ART medication, the consequences of the viral load in the body, and the effects due to the combination of both, PLWHA have begun exhibiting abnormalities such as elevated blood lipids, central fat accumulation, lipodystrophy, and diminished aerobic capacity. All of these are known risk factors for malignancy, cardiovascular disease, and other life-threatening conditions. There are numerous factors to investigate as potential causes of this elevated chronic disease risk. The majority of causes are of a metabolic nature. This category includes poor dietary practices, an inactive or sedentary lifestyle, stress, and other poor lifestyle choices. Another factor is the affordable cost of energy-dense, sugary, and fatty foods. By engaging in resistance training, aerobic exercise, or a combination of the two, healthy and other therapeutic groups frequently alter these risk factors [15]. It has been demonstrated that PLWHA may react similarly, which could increase the lifespan and quality of life.

As a person's immune system deteriorates as a result of the illness, it opens the door to numerous opportunistic illnesses and neurological problems, such as sleeping disorders, which may be caused by a lack of dopamine release (which the adrenal glands produce to regulate sleep) and may contribute to the deterioration of the patient's health. In this regard, physical exercise appears to be a non-pharmacological treatment that seeks to reduce the medication's side effects and opportunistic illnesses, offering PLWHA several benefits.

3.3 Lifestyle Management

After diagnosis, maintaining a healthy lifestyle is difficult. Effective management and the adoption of a healthy lifestyle improve a patient's response to medication. Lifestyle management entails the consumption of recommended, nutrient-rich foods, timely medication administration, sufficient rest, regular exercise, and effective mental health care. Even with ART medications, the prevalence of HIV-associated neurocognitive disorders (HAND) has not decreased. Consequently, PLWHA has difficulties with attention, concentration, and memory. Thus, the victims may experience forgetfulness regarding the punctual consumption of medications [8].

In addition to improving lifestyle, mental health should also be a priority. A study found that social support was viewed as the most important factor in sustaining a healthy lifestyle, whereas socioeconomic or financial factors were viewed as the most significant obstacles [3]. When victims are subjected to social pressure or tension, they are severely affected. Social stigma regarding the subject causes PLWHA to be despondent and anxious. Family acceptance, one-on-one counselling sessions, and support groups help PLWHA to live a life of higher quality. Self-motivational concern, desire, and purpose are crucial in inducing behavior change and therapy adherence among PLWHA [11,13,32].

4 Materials and Methods

The study in the discussion was mixed-method research to design and develop a mobile-based lifestyle-enhancing application for PLWHA. The research was performed following 5 major phases, which are discussed below.

4.1 Phase 01: Requirement Assembling

After ideating and identifying a research scope, primary and secondary research methods were used in the requirement-gathering phase. As for secondary research, background reading, collecting government statistics, and collecting global statistics from renowned organizations were performed to gain existing data relevant to the study. 3 Primary research methods, Questionnaires, Interviews, and focus group discussions were used during this phase to identify requirements.

As of the questionnaire, a sample size calculation was conducted to ensure that the anticipated number of observations will offer a realistic possibility of producing a definitive response, and a snowball sampling technique was used to select the sample of respondents (PLWHA) for this purpose. The sample was diverse based on various factors.

In order to acquire a better understanding of the recommended lifestyle for PLWHA, interviews were conducted with medical professionals such as virologists, general consultants, and medical officers. The interviews enabled the author to obtain a better understanding of the target audience’s pain points from the healthcare providers’ perspective.

The focus group discussion was conducted with the complete consent of randomly selected PLWHA in a private setting at the Medical Centre for Family Care – Gothatuwa, with no other individuals present. Through this discussion, first-hand experience of those infected by the disease, their pain points, and their expectations were discussed bringing lighter to how the solution should be designed.

4.2 Phase 02: Analysis

The responses of all the primary research methodologies were thoroughly analyzed by the author. A list of requirements was derived based on the analysis. Competitor Research was a secondary research method that was utilized in this phase to identify the gaps present that need to be addressed based on the primary research conducted. In addition, design thinking methods such as mind mapping, brainstorming and creating fish bone diagram.

The scope of the solution was defined in this stage and the list of functional and non-functional requirements was finalized. UML diagrams such as Use case diagram, Activity diagrams, and Class diagram were created to visualize the project and to define the scope more accurately. The following features were defined in the proposed solution at this phase: Nutrition tracker based on image processing, Meal recommendation based on machine learning technology, medication consumption and medication refill tracker, exercise tracker, exercise recommendation-based machine learning technology, BMI tracker, an Information hub and gamification components to motivate the user continuously.

4.3 Phase 03: Design

The design of the solution was performed in this phase. Initially the color palette was selected with the Primary color as blue. The color blue has been used primarily in the design to represent credibility, trust, and calmness. The color red, which is often associated with the disease, has been avoided in the design to ensure the comfort of using the application for patients. The design also considered emotions, the privacy and mental health concerns of users by avoiding frequent visual reminder of their health concerns whenever they use the application [24].

The application's flow of functions is maintained in a simple manner, prioritizing the user experience. Navigation is simplified by including navigation icons for the application's major features at the bottom of the screen in the navigation bar. Visual consistency is maintained by keeping all buttons and typefaces consistent throughout the application.

The following high level architecture diagram (Figure 2) clearly illustrates the design and architecture of the proposed solution.

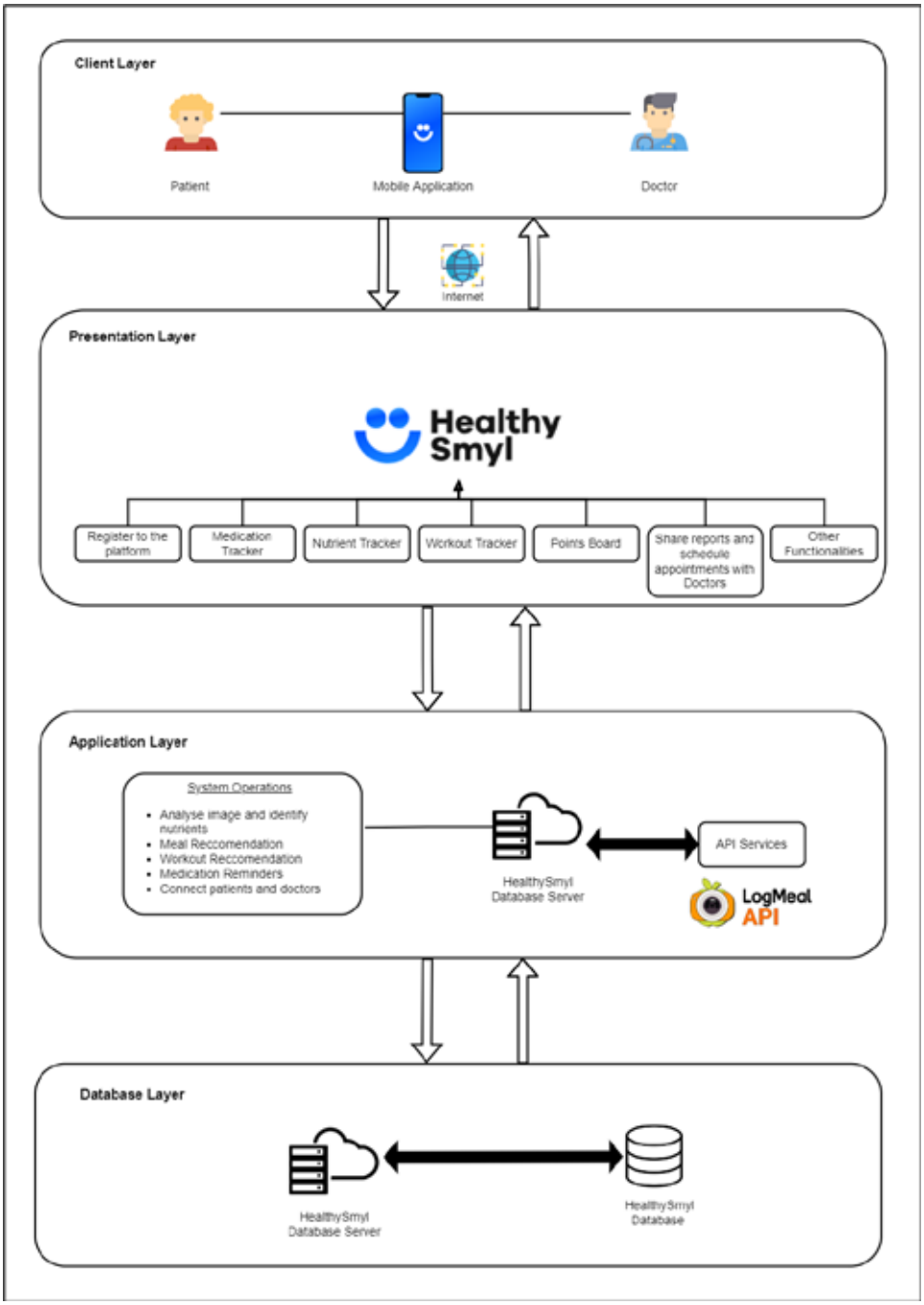


Figure 2: High Level Architecture of the Proposed Solution

4.4 Phase 04: Development

The solution was developed using Visual Studio Code, as a progressive web application using Python Flask, React JS, and SCSS. A Firebase database was used to store data, and APIs such as EDAMAM and LogMeal API were used for the nutrition tracker and recommendation features.

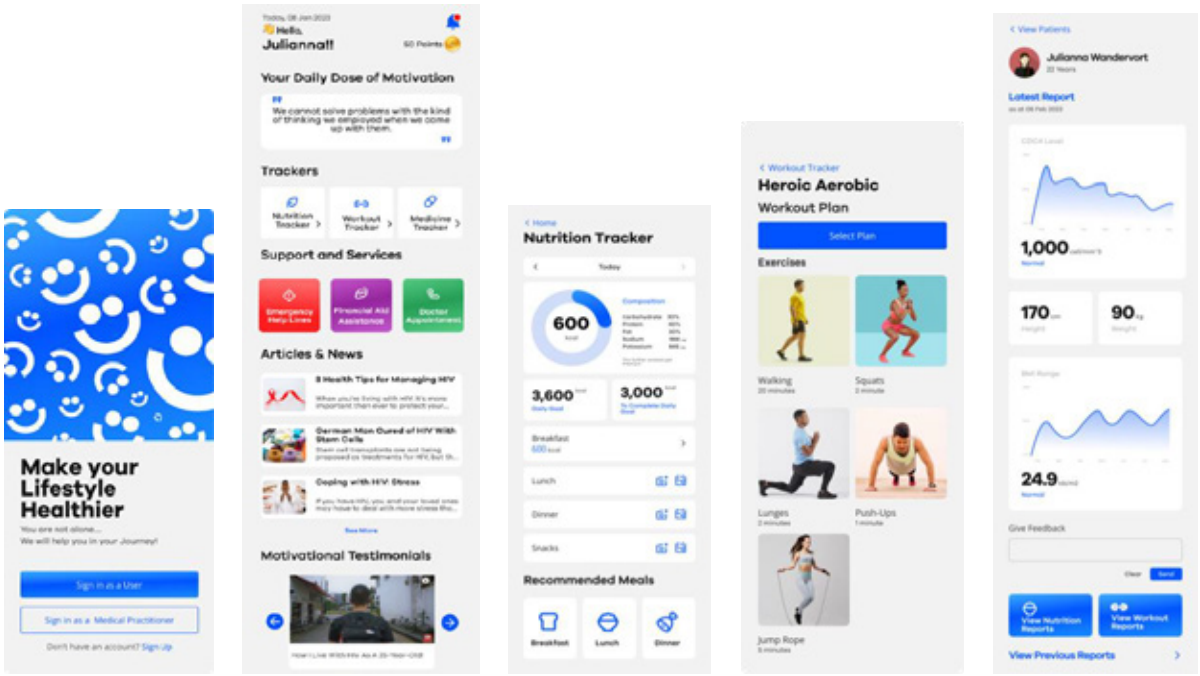


Figure 3: Screenshots of the Application UI

4.5 Phase 05: Testing

Functional and User Testings were performed for the developed solution to ensure accuracy and expected performance. Functional testing was conducted using the black-box technique for the implemented key functionalities of the developed solution. The testing was successful, and all test cases provided the expected output. Especially in regard to the image processing component for nutrition analysis, 20 images from the data set were added and tested, of which 17 were identified correctly. This portrays an 85% accuracy of the system. The following equation was used to calculate the accuracy.

$$Accuracy\ Level\ = \frac{Number\ of\ correctly\ recognized\ images}{Number\ of\ totally\ tested\ images} \times 100$$

To validate the effectiveness of the implemented solution, a prototype was subjected to evaluation by domain experts, technical experts, and end-users operating within the industry. The evaluation was conducted based on two key criteria, namely concept-based evaluation, and design-based evaluation. A set of pre-identified questions were asked from these testers to receive the needful feedback.

5 Recommendations and Future Work

5.1 Recommendations

The following recommendations can be made to improve the proposed solution, HealthySmyl, based on the research findings. These recommendations concentrate on various aspects of the application, including localization of the language, integration with health platforms, recommendation systems, practitioner engagement, image processing, nutrition recognition, and security measures.

It is recommended to implement a language localization feature in HealthySmyl. This function should allow users to select their preferable language, such as English, Sinhala, or Tamil. The application can serve a broader user base and provide a more inclusive experience if it supports multiple language preferences. It would ensure that users can access and comprehend the content in their preferred language, thereby enhancing user satisfaction and accessibility.

HealthySmyl should be integrated with popular health platforms such as Fitbit, Apple Health, and Google Health in order to increase its efficacy and breadth. This integration would permit users to synchronize their health data from these platforms with HealthySmyl. By aggregating health data from various sources, users are able to access and analyze their data in a centralized location. It would provide a comprehensive view of their health and wellness, allowing for improved monitoring and administration of their overall health.

To enhance HealthySmyl's recommendation systems, the application should employ machine learning techniques. The application can generate personalized and accurate recommendations by training a machine learning model with user data, preferences, and behavior. These recommendations would be tailored to the individual health objectives of each user, taking into account dietary preferences, exercise routines, and medical conditions. By providing more pertinent and personalized recommendations, the application can effectively steer users toward healthier decisions and behaviors.

Important for engaging healthcare professionals is the development of a practitioner portal within HealthySmyl. This portal should provide practitioners with secure access to their patients' data, enabling them to monitor progress, analyze trends, and provide individualized advice. Appointment scheduling, patient communication, and data visualization would improve the efficacy of healthcare delivery. HealthySmyl can become a valuable instrument for both patients and healthcare professionals by incorporating practitioners into the application.

Improving HealthySmyl's image processing model is recommended to enhance its ability to accurately recognize local cuisines. By training the model with a greater variety of local culinary images, it is able to provide users with more accurate nutritional data. This enhancement would assist users in making dietary decisions that are in line with their cultural or regional preferences. It would enhance the overall utility and relevance of the application for users in various contexts.

To further enhance HealthySmyl's nutrition recognition model, it is essential to consider meal quantity. Taking into consideration portion sizes and quantities would allow the application to provide more accurate nutritional information to its users. Accounting for the fact that nutritional values differ based on the quantity consumed would enhance the application's capacity to assist users in effectively managing their dietary goals. Providing precise nutritional information based on meal size would increase user contentment and confidence in the application.

Implementing Multi-Factor Authentication (MFA) in HealthySmyl is highly recommended in order to increase security measures. By requiring multiple authentication factors, such as passwords and one-of-a-kind verification codes, the application can strengthen account security and secure sensitive health data. MFA adds an additional layer of security, decreasing the likelihood of unauthorized access and data intrusions. Ensuring the privacy and security of user information is crucial for the application's success and credibility.

5.2 Future Work

Future implementations of HealthySmyl could involve expanding the initiative to include minors affected by the disease along with the adult population. This expansion would ensure that the application provides support and resources to a broader range of HIV/AIDS-affected individuals, addressing their unique requirements and difficulties.

In an environment where the government is digitizing medical records, integrating the application with Electronic Health Records (EHRs) would be beneficial. Both practitioners and patients can benefit from the

integration of HealthySmyl with EHR systems. Practitioners would have seamless access to their patients' medical records, enabling them to make more informed decisions and provide more personalized care. The application would provide patients with a more comprehensive view of their health history and treatments, fostering continuity of care.

Incorporating an e-bill analysis system into the medication refill monitoring feature can improve user experience and convenience. This system would analyze electronic invoices or digital prescription data automatically and update medication refills without requiring user input. By automating the process, users can easily monitor and manage their medication refills, thereby reducing the likelihood of errors and omissions.

These future implementations would expand the scope and functionality of HealthySmyl, allowing it to serve a larger population, integrate seamlessly with existing healthcare systems, and offer more effective and user-friendly medication management features.

6 References

[1] Gift Treighcy Banda, Grace Mwale, Michael Chimwala, Linda Malimusi, and Enock Chisati. 2019. Common impairments and functional limitations of HIV sequelae that require physiotherapy rehabilitation in the medical wards at Queen Elizabeth Central Hospital, Malawi: A cross sectional study. *Malawi Med J* 31, 3 (September 2019), 171–176. DOI:https://doi.org/10.4314/mmj.v31i3.2

[2] Grant R. Campbell and Stephen A. Spector. 2012. Toll-Like Receptor 8 Ligands Activate a Vitamin D Mediated Autophagic Response that Inhibits Human Immunodeficiency Virus Type 1. *PLoS Pathog* 8, 11 (November 2012), e1003017. DOI:https://doi.org/10.1371/journal.ppat.1003017

[3] Bernadette Capili, Joyce K. Anastasi, Michelle Chang, and Olugbenga Ogedegbe. 2014. Barriers and facilitators to engagement in lifestyle interventions among individuals with HIV. *J Assoc Nurses AIDS Care* 25, 5 (2014), 450–457. DOI:https://doi.org/10.1016/j.jana.2014.01.003

[4] Ceylon Today. 2022. Have we forgotten HIV/AIDS? *Ceylon Today*. Retrieved November 17, 2022 from https://ceylontoday.lk/2022/10/28/have-we-forgotten-hiv-aids/

[5] Pauline Cheung, Monica Gonzalez, Sonia M. Grant, Esther M. Trepal, and Lisa M. Zullig. 2010. *Eating Tips: A Nutrition Guide for People Living with HIV/AIDS*. God's Love We Deliver, USA.

[6] Department of Health of Australia. 2021. Water – A vital nutrient. *Better Health Channel*. Retrieved December 12, 2022 from https://www.betterhealth.vic.gov.au/health/healthyliving/water-a-vital-nutrient

[7] Nancy Dumais (Ed.). 2020. *Nutrition and HIV/AIDS Implication for Treatment, Prevention and Cure* (First ed.). INTECHOPEN LIMITED, London. Retrieved from https://mts.intechopen.com/storage/books/7856/authors_book/authors_book.pdf

[8] Christian Eggers, Gabriele Arendt, Katrin Hahn, Ingo W. Husstedt, Matthias Maschke, Eva Neuen-Jacob, Mark Obermann, Thorsten Rosenkranz, Eva Schielke, Elmar Straube, and For the German Association of Neuro-AIDS und Neuro-Infectiology (DGNANI). 2017. HIV-1-associated neurocognitive disorder: epidemiology, pathogenesis, diagnosis, and treatment. *J Neurol* 264, 8 (August 2017), 1715–1727. DOI:https://doi.org/10.1007/s00415-017-8503-2

[9] Ezinna E. Enwereji, Martina C. Ezeama, Prince E. N. Onyemachi, Ezinna E. Enwereji, Martina C. Ezeama, and Prince E. N. Onyemachi. 2019. *Basic Principles of Nutrition, HIV and AIDS: Making Improvements in Diet to Enhance Health*. IntechOpen. DOI:https://doi.org/10.5772/intechopen.84719

[10] Anthony S. Fauci and H. Clifford Lane. 2022. Four Decades of HIV/AIDS — Much Accomplished, Much to Do. *N Engl J Med* (February 2022). DOI:https://doi.org/10.1056/NEJMp1916753

[11] Nelsensius Klau Fauk, Karen Hawke, Lillian Mwanri, and Paul Russell Ward. 2021. Stigma and Discrimination towards People Living with HIV in the Context of Families, Communities, and Healthcare Settings: A Qualitative Study in Indonesia. *International Journal of Environmental Research and Public Health* 18, 10 (January 2021), 5424. DOI:https://doi.org/10.3390/ijerph18105424

[12] Nadia Fazlulhaq and Isuri Kaviratna. 2007. HIV/AIDS-affected women take woes before "court." Retrieved December 16, 2022 from https://www.sundaytimes.lk/070819/News/nws18.html

[13] Perry N. Halkitis, Kristen D. Krause, and Dorice L. Vieira. 2017. Mental Health, Psychosocial Challenges and Resilience in Older Adults Living with HIV. In *Interdisciplinary Topics in Gerontology and Geriatrics*, M. Brennan-Ing and R.F. DeMarco (eds.). S. Karger AG, 187–203. DOI:https://doi.org/10.1159/000448564

[14] Behzad Imani, Shirdel Zandi, Salman khazaei, and Mohamad Mirzaei. 2021. The lived experience of HIV-infected patients in the face of a positive diagnosis of the disease: a phenomenological study. *AIDS Res Ther* 18, (December 2021), 95. DOI:https://doi.org/10.1186/s12981-021-00421-4

[15] Jason R. Jaggers and Gregory A. Hand. 2016. Health Benefits of Exercise for People Living With HIV: A Review of the Literature. *American Journal of Lifestyle Medicine* 10, 3 (May 2016), 184–192. DOI:https://doi.org/10.1177/1559827614538750

[16] John R. Koethe, Cathy A. Jenkins, Bryan Lau, Bryan E. Shepherd, Amy C. Justice, Janet P. Tate, Kate Buchacz, Sonia Napravnik, Angel M. Mayor, Michael A. Horberg, Aaron J. Blashill, Amanda Willig, C. William Wester, Michael J. Silverberg, John Gill, Jennifer E. Thorne, Marina Klein, Joseph J. Eron, Mari M. Kitahata, Timothy R. Sterling, and Richard D. Moore. 2016. Rising Obesity Prevalence and Weight Gain Among Adults Starting Antiretroviral Therapy in the United States and Canada. *AIDS Res Hum Retroviruses* 32, 1 (January 2016), 50–58. DOI:https://doi.org/10.1089/aid.2015.0147

[17] Kelsi A. Kriitmaa. 2018. *Behavioural Surveillance (IBBS) Survey Among Key Populations at Higher Risk of HIV in Sri Lanka*. National STD / AIDS Control Programme, Sri Lanka.

[18] Ariyaratne Manathunge, Jelena Barbaric, Tomislav Mestrovic, Sriyakanthi Beneragama, and Ivana Bozicevic. 2020. HIV prevalence, sexual risk behaviours and HIV testing among female sex workers in three cities in Sri Lanka: Findings from respondent-driven sampling surveys. *PLoS One* 15, 10 (October 2020), e0239951. DOI:https://doi.org/10.1371/journal.pone.0239951

[19] Saurabh Mehta and Julia L. Finkelstein (Eds.). 2018. *Nutrition and HIV: Epidemiological Evidence to Public Health*. CRC Press, New York (NY). Retrieved October 4, 2022 from http://www.ncbi.nlm.nih.gov/books/NBK572225/

[20] Ministry of Health. 2017. *National HIV Monitoring & Evaluation Plans Sri Lanka, 2017-2022*. National STD/AIDS Control Programme, Ministry of Health, Sri Lanka, Sri Lanka.

[21] Ministry of Health. 2021. *Annual Report 2020*. National STD/AIDS Control Programme, Ministry of Health, Sri Lanka.

[22] G. Nanayakkara and Karawita DA. 2018. Knowledge on HIV transmission and attitude towards providing care for people living with HIV (PLWHA) and adherence to univer sal precautions (UP) among attendants in the National Hospital of Sri Lanka (NHSL). *Sri Lanka Journal of Sexual Health and HIV Medicine* 4, (December 2018), 34. DOI:https://doi.org/10.4038/joshhm.v4i0.74

[23] National Institute of Health of USA. 2021. HIV and Nutrition and Food Safety | NIH. *Living with HIV*. Retrieved December 12, 2022 from https://hivinfo.nih.gov/understanding-hiv/fact-sheets/hiv-and-nutrition-and-food-safety

[24] Don Norman. 2007. *Emotional Design: Why We Love (or Hate) Everyday Things*. Basic Books.

[25] Bharat S. Parekh, Chin-Yih Ou, Peter N. Fonjungo, Mireille B. Kalou, Erin Rottinghaus, Adrian Puren, Heather Alexander, Mackenzie Hurlston Cox, and John N. Nkengasong. 2018. Diagnosis of Human Immunodeficiency Virus Infection. *Clinical Microbiology Reviews* 32, 1 (November 2018), e00064-18. DOI:https://doi.org/10.1128/CMR.00064-18

[26] Jennifer Robinson. 2021. Nutrition and HIV/AIDS. *WebMD*. Retrieved December 12, 2022 from https://www.webmd.com/hiv-aids/guide/nutrition-hiv-aids-enhancing-quality-life

[27] Max Roser and Hannah Ritchie. 2019. HIV / AIDS. *Our World in Data* (2019). Retrieved October 5, 2022 from https://ourworldindata.org/hiv-aids

[28] Chaturanga Samarawickrama. 2022. Chemsex trending in Sri Lanka. *Daily Mirror*. Retrieved November 17, 2022 from https://www.dailymirror.lk/breaking_news/Chemsex-trending-in-Sri-Lanka/108-247982

[29] Maryline Sicotte, Étienne V. Langlois, Joséphine Aho, Daniela Ziegler, and Maria Victoria Zunzunegui. 2014. Association between nutritional status and the immune response in HIV + patients under HAART: protocol for a systematic review. *Systematic Reviews* 3, 1 (February 2014), 9. DOI:https://doi.org/10.1186/2046-4053-3-9

[30] UNFPA. 2022. Fighting the stigma of living with HIV in Sri Lanka. Retrieved November 17, 2022 from https://reliefweb.int/report/sri-lanka/fighting-stigma-living-hiv-sri-lanka

[31] Vanaja. P, Vasanthi. R, Sasikala. S, and Sivakumar. T. 2021. Relationship Between HIV and Nutrition. *HIV Nursing* 21, 02 (December 2021), 68–72. DOI:https://doi.org/10.31838/hiv21.02.09

[32] Rachel C. Vreeman, Brittany M. McCoy, and Sonia Lee. 2017. Mental health challenges among adolescents living with HIV. *Wiley Online Library* (2017). Retrieved October 30, 2022 from https://onlinelibrary.wiley.com/doi/full/10.7448/IAS.20.4.21497

[33] World Health Organization. 2022. HIV. Retrieved November 17, 2022 from https://www.who.int/data/gho/data/themes/hiv-aids

iCIIT 2023

THE 4TH INTERNATIONAL CONFERENCE ON INNOVATIONS IN INFO-BUSINESS & TECHNOLOGY

Conference Track

COMPUTING



Developing An Automated System to Analyzing the Software Developers Interest Using Twitter Data

K. N. M. S. Nandasiri †
Department of Information Technology, University of Moratuwa, Sri Lanka, maleesh.18@itfac.mrt.ac.lk

K. G. L. S. Ananda
Department of Information Technology, University of Moratuwa, Sri Lanka, lakshitha.18@itfac.mrt.ac.lk

D. V. S. Maduranga
Department of Information Technology, University of Moratuwa, Sri Lanka, sachith.18@itfac.mrt.ac.lk

K. D. H. S. Bandara
Department of Information Technology, University of Moratuwa, Sri Lanka, hansaka.18@itfac.mrt.ac.lk

C. P. Wijesiriwardena
Department of Information Technology, University of Moratuwa, Sri Lanka, chaman@uom.lk

I. T. S. Piyatilake
Department of Computational Mathematics, University of Moratuwa, Sri Lanka, thilinisp@uom.lk

† Corresponding Author

ABSTRACT

The information technology sector is a quickly expanding, well-liked, successful, and significant enterprise. In every IT organization, the software development process involves team effort. Therefore, choosing the best developers for the team to develop a software project is the first activity that could affect the project's outcome. Some project managers choose the team based on their preferences without considering developers willingness and knowledge which might pose problems for the project. It is crucial to assess developers' interests and skills to determine whether they are a good fit for the project. In this paper, we present a system that utilizes Twitter data to analyze and predict the interests of developers using a Sentimental analyzer based on the Naive Bayes algorithm and Vader sentiment analyzing technique. The proposed system leverages the Twitter REST API to extract tweets from relevant user accounts and applies a modified sentiment analyzer to analyze and measure the level of interest of software developers regrading a particular software development technology. The system has potential applications in various fields, such as targeted marketing, content curation, and personalized recommendations. We provide a detailed description of the system architecture, implementation, and evaluation, demonstrating the system's effectiveness in accurately predicting developer interests based on Twitter data.

Keywords: Software Development, Twitter, Naive Bayes Classifier, Sentimental Analyzer, Vader Analyzer

I INTRODUCTION

Software developers are a crucial demographic in the information technology industry, and their interests and preferences are of great significance to software companies, startups, and other organizations [1]. Choosing the best team of software developers to work on a project is one of the biggest problems in software development. It's crucial to pick developers that are really interested in the project in addition to having the required knowledge and abilities. This is so that the project's developers can provide better results since motivated and engaged developers are more likely to be involved in the project. By considering their level of interest, we want to discover the most interested developers for a particular project in this research study. Twitter is a popular social media platform that enables users to express their opinions, interests, and thoughts

in real time through short messages known as tweets. The vast amount of data generated by Twitter users presents an opportunity for researchers to analyze this data to gain insights into users' interests, behaviors, and preferences.

In this paper, we aim to develop a system that can accurately predict the interests of developers using Twitter data and sentiment analysis techniques. This system extracts tweets from Twitter users, preprocesses tweet contents, and classifies the tweets using a Sentiment Analyzer based on Naive Bayes classifier algorithm. Vader Sentiment analyzing technique is used to calculate the interest levels of the users and rank them according to their interest levels. The effectiveness of the proposed system is evaluated for accurately predicting developer interests using Twitter data. Our proposed approach can be useful for various applications such as targeted marketing, content curation, and personalized recommendations. This paper provides a detailed description of the system architecture, implementation, and evaluation. The results of our experiments demonstrate the effectiveness of our system in accurately predicting developer interests based on Twitter data.

The rest of this paper is organized as follows: Section II includes the literature review. Proposed approach, design and implantations are presented in Section III. Main results of this research are presented in Section IV. Section V presents some concluding remarks and future research directions.

II LITERATURE REVIEW

Bougie et al. [3] conducted to understand the usage of twitter in software engineering. They have chosen tweeters to represent three communities which are Eclipse, Linux, and MXUnit. The researchers have categorized these tweets into Software engineering-related (often also work-related) topics (e.g., projects being worked on, seeking or providing technical help), Gadgets and technological topics (e.g., iPhone 4, product news from Microsoft), Current events outside of technical topics (e.g., FIFA World Cup, Ottawa earthquake in summer of 2010) and Daily chatter (e.g., family, weekend activities) [2]. Although they found four categories in this study, they have narrowed down their focus on the first category to analyze their technological interests related to Software Engineering work or activities. The subtopics that were discovered within this category are problem Solving, Self-promotion, complaints, and the use of a specific tool for work.

As per another study conducted by David Jaderberg on the topic Sentiment and topic classification of messages on Twitter [3]. It has used natural language processing to classify messages posted on twitter based on the sentiment and the topic of the tweets. The sources of data are a network of users who tweets science-related content. The system is programmed using Python language as it provides most of the libraries used in this study which are Natural Language Toolkit (NLTK) for NLP tasks and Python Twitter Tools for interacting with the Twitter API. When the application is running, it automatically downloads and evaluates new tweets from a small subset of Twitter users, which includes primarily scientists and some news outlets covering science. This task has been accomplished using the Python Twitter Tools. Then the tweets are preprocessed using the metadata of the Twitter posts. Then Naïve Bayes classifier was used for the classification tasks. When categorizing a tweet's content and sentiment using this method, respectively, the accuracy for the earlier was measured at 29.9% and 33.5%. As the response generation significantly relies on the classifications of tweets, the program's performance is constrained by the trained Naïve multinomial Bayes classifiers' poor classification accuracy. This study has suggested a few techniques to improve accuracy such as increasing the size of the training dataset.

Agarwal et al. provides a comprehensive overview of sentiment analysis on Twitter data [2]. The authors first describe the importance and challenges of analyzing sentiment on social media platforms. They then present various approaches and techniques for sentiment analysis on Twitter data, including lexicon-based approaches, machine learning approaches, and deep learning approaches. The authors also discuss the limitations and future directions of sentiment analysis on Twitter data, such as the need for more robust methods for handling sarcasm and irony. The paper concludes by highlighting the potential applications of

sentiment analysis on Twitter data, such as brand monitoring, customer feedback analysis, and political opinion mining.

In the vast landscape of Twitter data analysis, previous research has primarily centered around categorizing posts within software engineering groups, yet the exploration of developers' specific interest levels in these technologies remains relatively uncharted. While interest level analyzers have been employed in product reviews, there is a noticeable lack of studies that delve into the analysis of tweets concerning software engineering topics. Additionally, existing interest level analyzers have been predominantly designed to measure the overall sentiment of tweets, disregarding the crucial context of technology-specific content. As a result, our research takes a novel approach, seeking to develop an interest level analyzer for the essence of tweets pertaining to technologies. By this we aim to enhance the accuracy and precision of our analyzer, providing valuable insights into developers' genuine interests within the software engineering domain. This endeavor holds the potential to not only benefit software companies, startups, and organizations in selecting the most passionate developers for projects but also open avenues for targeted marketing, content curation, and personalized recommendations in the software development landscape.

III METHODOLOGY

A . Proposed Approach

For this module we need to retrieve tweets and the associated data from relevant Twitter user accounts. The user's interest will then be predicted using the processed and analyzed data that has been obtained. Using Twitter APIs, we must first collect tweets from Twitter user profiles regarding the given technological keywords. The tweets will be filtered and extracted via APIs. We must connect to Twitter's REST API to perform this extraction. So, to access the API and retrieve tweets and other information from it, we must create a Python client. Then we aim to compile all the tweets regarding given technology into a single data frame for straightforward analysis. Additionally, a few other actions will be carried out to finish the preprocessing because the main goal is to make the text of the tweets available for the stage of analysis. In the analyzing stage, the preprocessed data will be analyzed using an unsupervised learning algorithm called VADER to analyze the sentiment value of the twitter posts in related technology. These posts will be then used to train a Naïve Bayes algorithm based sentimental analyzer to have a more accurate prediction over the given software developer’s twitter posts filtered and extracted using the Twitter APIs. The users will then have a sentiment score regarding the related technology and then they will be ranked according to the users’ interest level in this module.

B. Analysis and Design

Tweets are extracted using Twitter APIs from relevant Twitter user accounts. A Python client is constructed in this procedure to access the API and pull tweets and other information from it. Prior to being inserted into the main processing flow, raw data must undergo preprocessing. The following actions are considered to prepare the texts. Tokenizing: Change the sentence into a collection of a single word, Stemming: Returns a word into basic form (root word) by eliminating existing additive (Porter Stemmer algorithm is used for the English language), and Filtering: Eliminating stop words. Stop word is a common word that has little or no meaning but required in the structure of grammatical language, Weighting: Calculate Tf-Idf for each word. TF-IDF score represents the relative importance of a term in the document and the entire corpus. TF-IDF score is composed of two terms: the first computes the normalized Term Frequency (TF), the second term is the Inverse Document Frequency (IDF), computed as the logarithm of the number of documents in the corpus divided by the number of documents where the specific term appears.

For the classification Naive Bayes classification is used. It is a classification algorithm that uses the Bayes theorem in its application. A Naive Bayes variant called Multinomial Naive Bayes (MNB) was created to address the classification of text content. Multinomial distribution in (1) is used by MNB as a classification characteristic together with a word's weight or frequency.

$$P(X|c) = log\frac{N_c}{N} + \sum_{i=1}^{n} log\frac{t_i+\propto}{\sum_{i=1}^{n} t+\propto}. \quad (1)$$

Here $P(X|c)$ is the probability document *X* in class *c,* N_c is the total documents in class *c,* N is the total documents, t_i is the weight term *t,* $\sum_{i=1}^{n} t$ is the total weight term *t* in class *c* and α is the smoothing parameter.

C. High Level Design and Architecture

Figure 1 shows the high-level design used to retrieve a collection of tweets from users. The text is preprocessed into vector data. The classification process will classify the user's text into a labelled dataset. The results are the interest level of a user for a particular technology.

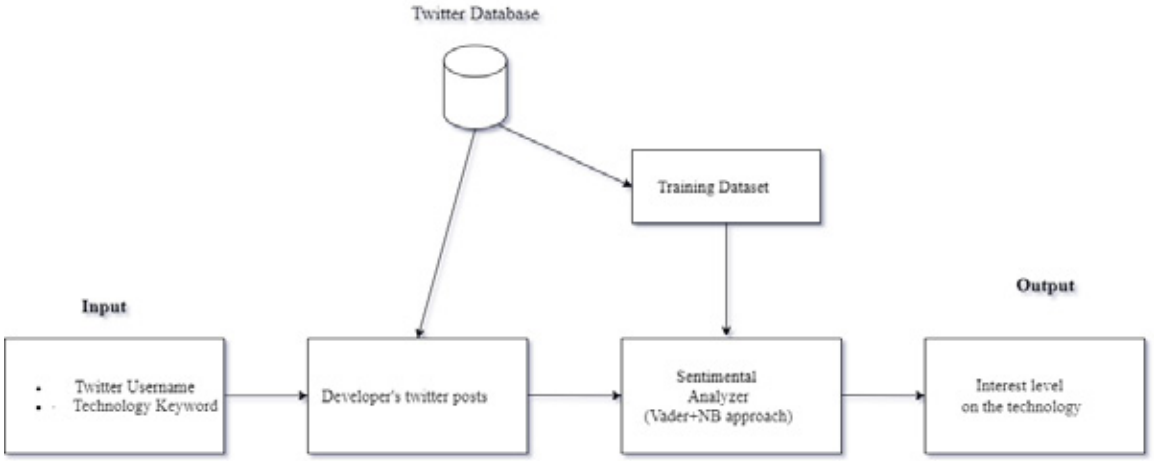


Figure 1: Interest analyzing using Twitter content.

D. IMPLEMENTATION

This component is responsible for extracting tweets from Twitter users, preprocess tweet contents, analyze & classify twitter contents to calculate interest levels of the users with a novel sentimental analyzer to analyze the text content related to Software Engineering topics. Before accomplishing our main target, we must pass several stages. The first one is to extract data from Twitter using Twitter APIs. There are two types of APIs available related to this, namely Twitter REST API and Twitter Streaming API. In our work REST API access is the most suitable one since we don't need real-time tweet. After taking the developer authentication for REST API access we proceed with implementation. Using these authentication details, we generate a python script to extract tweets. It acts as a client which extracts tweets via Twitter API. We can further extract the twitter posts based on the twitter user and the keywords used in the tweets by applying search queries when extracting data using the Twitter API. The proposing sentimental analyzer analyzes the sentimental value of the twitter posts regarding a technology. This sentimental analyzer is based on a Valence Aware Dictionary for Sentiment Reasoning (VADER) and Naïve Bayers classification algorithm [5]. Our goal is to achieve higher accuracy when analyzing sentimental value on the posts related to a specific technology. Figure 2 shows the detailed design of the system architecture.

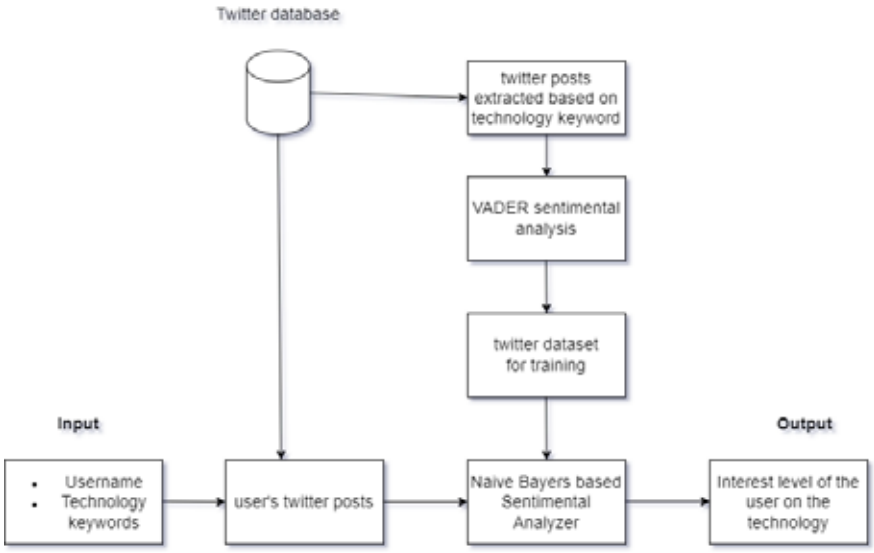


Figure 2: detailed architecture of the proposed sentiment analyzer

To compare the accuracy of the novel sentimental analyzer the accuracy results will be taken from

1. A direct approach using VADER sentimental analysis[4]. Figure 3 shows the design used to analyze the twitter posts sentiment using VADER sentimental analysis approach.

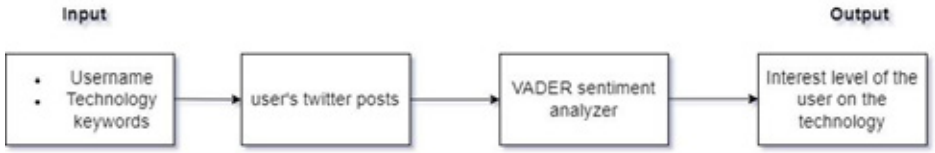


Figure 3: VADER sentimental analysis

2. A direct approach using sentimental analysis based on Naïve Bayers algorithm[5]. Figure 4 shows the design used to analyze the twitter posts sentiment using Naïve Bayers sentimental analysis approach.

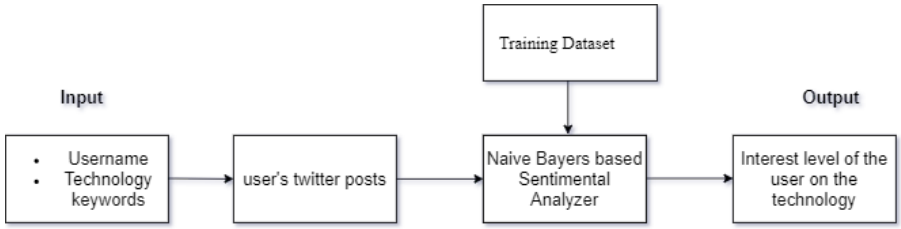


Figure 4: Naïve Bayers sentimental analysis

The comparison will be carried out using a twitter posts dataset which consists of technology related posts. According to this comparison the proposed system's effectiveness is being tested.

IV Results and Discussion

The effectiveness of the proposed system is tested by the accuracies of the proposed sentiment analyzer with VADER sentiment approach and Naïve Bayers sentiment approach. The Sentiment140 dataset [3] which consist of 1.6 million tweets was used to extract software development technology related tweets for testing

accuracy. The following Table 1 indicates the accuracy of each approach in predicting sentiment for each technology based on the given number of tweets.

Table 1: Number of tweets extracted for testing accuracy for each technology with accuracy scores.

Technology	Number of Tweets	VADER Approach	Naïve Bayers Approach	Proposed Approach
JavaScript	91	63.04%	50%	59.78%
MySQL	96	59.37%	56.25%	67.71%
ASP.NET	41	51.21%	54.54%	56.09%

The provided results present the performance of three different approaches (VADER, Naïve Bayes, and Proposed Approach) in predicting the interest level of developers in three different technologies (JavaScript, MySQL, and ASP.NET) based on the analysis of a certain number of tweets for each technology.

VADER Approach:

JavaScript: The VADER approach achieved a sentiment accuracy of 63.04% for JavaScript tweets. While this percentage might seem reasonable, it's essential to note that the VADER approach is primarily designed for general sentiment analysis and may not be tailored to the specific nuances and technical context of software engineering tweets.

MySQL: For MySQL tweets, the VADER approach achieved a sentiment accuracy of 59.37%. Similar to the JavaScript case, this score might appear acceptable, but it might not fully capture the nuances of developers' interest in MySQL.

ASP.NET: The VADER approach obtained a sentiment accuracy of 51.21% for ASP.NET tweets. This percentage suggests that the VADER approach struggled more in identifying developers' interest in ASP.NET, potentially due to the technical nature of the topic.

Naïve Bayes Approach:

JavaScript: The Naïve Bayes approach scored 50% accuracy for JavaScript tweets. This score is not much better than a random guess, indicating that the Naïve Bayes model might not be well-suited for analyzing the specific context and interest levels in JavaScript tweets.

MySQL: For MySQL tweets, the Naïve Bayes approach achieved a sentiment accuracy of 56.25%. While better than chance, this score still suggests that the model may not fully capture the nuances of developers' interest in MySQL.

ASP.NET: The Naïve Bayes approach obtained a sentiment accuracy of 54.54% for ASP.NET tweets. Again, this score is not significantly better than random guessing, highlighting potential limitations in analyzing the specific interests of developers in ASP.NET.

Proposed Approach:

JavaScript: The proposed approach achieved a sentiment accuracy of 59.78% for JavaScript tweets. While this percentage is better than the previous two approaches, it still indicates that there is room for improvement in accurately predicting developers' interest in JavaScript.

MySQL: For MySQL tweets, the proposed approach achieved a sentiment accuracy of 67.71%, the highest among the three approaches. This result suggests that the proposed approach might have captured some of the nuances and specific interest levels of developers in MySQL more effectively.

ASP.NET: The proposed approach obtained a sentiment accuracy of 56.09% for ASP.NET tweets. Although slightly better than the previous two approaches, this score still indicates that there might be challenges in precisely predicting developers' interest in ASP.NET.

In a critical evaluation, it's important to consider the limitations and potential biases of the approaches and the dataset used for evaluation. The sample size of tweets for each technology might be relatively small, which can impact the generalizability of the results. Additionally, the selection of specific sentiment analysis techniques and machine learning models might influence the accuracy of interest level predictions. Further research and experimentation with larger datasets and diverse machine learning models could lead to more robust and reliable results.

V Conclusion

This research paper addresses the underlying problem of selecting the best team of software developers for a project. It emphasizes that the interests and preferences of developers are crucial factors in achieving better project outcomes. Motivated and engaged developers who are genuinely interested in a project are more likely to be actively involved and deliver better results. Using the Naive Bayes classifier and sentimental analysis, we were able to accurately classify and measure the level of interest of developers based on their tweets. This information could be valuable for targeted marketing, content curation, and personalized recommendations. Our study also contributes to the growing body of literature on the use of social media data for understanding the activities and interests of developers. While previous studies have focused on categorizing tweets into different topics or analyzing the online behavior of developers, our approach focuses on predicting their interests based on the content of their tweets. One of the challenges we faced during this research is gathering enough data regarding a given technology as some development technologies might not be popular among developers as the rest. The change of policies regarding twitter APIs was an issue we had to face as it provided some new restrictions when extracting tweets using Twitter APIs. Overall, we believe that our system has the potential to provide valuable insights into the interests and activities of developers, and we hope that our work will inspire further research in this area.

References

[1] Ansari, A. R., Khan, M. A., & Khan, R. Z. (2017). Predicting developer interests using Twitter data. *ACM Transactions on Software Engineering and Methodology*, 26(3), 1-33.

[2] Apoorv Agarwal, Boyi Xie, Ilia Vovsha, Owen Rambow, and Rebecca Passonneau. 2011. Sentiment Analysis of Twitter Data. In *Proceedings of the Workshop on Language in Social Media (LSM 2011)*, Association for Computational Linguistics, Portland, Oregon, 30–38. Retrieved June 9, 2023 from https://aclanthology.org/W11-0705

[3] Jäderberg David. 2016. *Sentiment and topic classification of messages on Twitter : and using the results to interact with Twitter users*. Retrieved June 9, 2023 from https://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-294364

[4] Alec Go, Richa Bhayani, and Lei Huang. Twitter Sentiment Classification using Distant Supervision.

[5] Toni Pano and Rasha Kashef. 2020. A Complete VADER-Based Sentiment Analysis of Bitcoin (BTC) Tweets during the Era of COVID-19. *Big Data Cogn. Comput.* 4, 4 (November 2020), 33. DOI:https://doi.org/10.3390/bdcc4040033

[6] Huma Parveen and Shikha Pandey. 2016. Sentiment analysis on Twitter Data-set using Naive Bayes algorithm. In *2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)*, IEEE, Bangalore, India, 416–419. DOI:https://doi.org/10.1109/ICATCCT.2016.7912034

Predicting the Possibility of Chronic Pancreatitis

Bawantha Wickramarachchi

Department of Computing, Informatics Institution of Technology, Colombo 06, Sri Lanka, bawantha.2018624@iit.ac.lk

Iresh Bandara

Department of Computing, Informatics Institution of Technology, Colombo 06, Sri Lanka, iresh.b@iit.ac.lk

ABSTRACT

This research explores the possibilities of developing a system for predicting the possibility of Chronic Pancreatitis (CP) in any given abdominal Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) image using deep learning techniques. When evaluating the previous research carried out in similar domains, it has been identified that the supervised learning approach will be more efficient in outcoming more accurate prediction. The final implementation of the model proves the necessity of using a deep learning model to satisfy the necessary requirements of the research carried out.

KEYWORDS

Chronic Pancreatitis, Deep Learning, Medical Imaging

1 Introduction

Chronic Pancreatitis presents a notable challenge in the field of medical diagnosis, particularly during its early stages due to its elusive nature. This often manifests without distinct symptoms, posing difficulties for medical practitioners in timely identification and treatment. Consequently, delayed intervention may occur, potentially leading to disease progression and reduced chances of favourable patient outcomes. Because of that reason, this research aims to address the challenging issue of early diagnosis for Chronic Pancreatitis by leveraging the potential of a deep learning approach. As none of the studies carried out in the computer science domain focused on CP, this study will be a valid contribution to the domain. Through the analysis of MRI or CT scans of the abdomen, the developed model aims to provide an output indicating the possibility of a person being diagnosed with CP, differentiating between healthy and damaged states of the pancreas. The primary objective of this research is to increase medical experts' attention to potential cases of CP at an early stage, where timely intervention can significantly impact patient outcomes.

1.1 Chronic Pancreatitis

The pancreas is located in the abdominal cavity[1] behind the stomach and has dual functions in the exocrine[2] and endocrine[3] systems. It can secrete juice containing several digestive enzymes and hormones [1]. These enzymes help in exocrine functionality by breaking down foods in the stomach during digestion, whereas hormones help regulate blood sugar [2]. If these digestive enzymes get activated inside the pancreas, it can lead to irritating the cells of the pancreas. Because of this, the pancreas becomes inflamed. When this

[1] The abdominal cavity is the anatomical space located within the abdomen, encompassing the area below the chest and above the hips. It serves as a vital region that houses various organs responsible for essential bodily functions.

[2] The exocrine system consists of glands that produce substances like sweat, tears, saliva, digestive juices, and milk. These substances are released through ducts or openings to reach specific body surfaces or target areas, aiding in processes such as digestion, lubrication, protection, and temperature regulation

[3] The endocrine system is a network of glands and organs that produce hormones and release them directly into the bloodstream. These hormones act as chemical messengers, traveling throughout the body to regulate various functions such as growth, metabolism, reproduction, mood, and stress responses.

inflammation comes and lasts for a shorter period, it is known as Acute Pancreatitis (AP). If this condition persists for longer, it will be considered chronic pancreatitis. This damage to the pancreas is irreversible in chronic pancreatitis, so there is no permanent cure for this disease. It can be controlled up to a considerable level with early detection.

1.2 Etiology and Pathophysiology of Chronic Pancreatitis

Long-term alcohol abuse is the most frequent cause of CP compared to other etiologies. In roughly 70% of cases, alcohol addiction is a contributing factor [3]. The ability of the pancreas to harm itself generally occurs when the digestive enzymes are triggered before they are released into the small intestine. The process is called autodigestion [4]. It is the reason why excessive alcohol consumption results in chronic inflammation. Alcohol can activate digestive enzymes, according to research that has already been conducted. Inflammation results from the early activation of gland-resident enzymes [5]. The pancreatic duct is blocked by gallstones, genetics, cystic fibrosis, a narrow pancreatic duct, and high levels of calcium (hypercalcemia) and triglyceride fats (hypertriglyceridemia) predominate in the blood will all contribute to pancreatitis in addition to alcohol consumption [3].

When gallstones[4] are pushed out of the gallbladder and trapped in the pancreatic duct, gallstone pancreatitis develops. This obstruction in the secretory pathway causes the pancreatic enzymes to become saturated. Since such enzymes have nowhere to go, they will activate inside the pancreas, resulting in inflammation and, eventually, CP [6]. Like gallstones, hypercalcemia and hypertriglyceridemia can result in pancreatitis by a similar mechanism. Higher blood levels of calcium and triglyceride fat generate calcium deposits in the pancreatic duct and fat deposits in the pancreas tissue [7]. Hereditary pancreatitis (HP) is a type of pancreatitis brought on by genetic causes. Acute pancreatitis symptoms in hereditary pancreatitis patients often appear in childhood. Cancer.Net shows that patients with HP have a higher risk of developing pancreatic cancer [8].

1.3 Treatments for Chronic Pancreatitis

The effectiveness of the various treatments for chronic pancreatitis depends on the extent of the condition and the patient's response to treatment. Pancreatic enzyme replacement therapy (PERT), which involves administering pancreatic enzymes to the body as a capsule to aid food digestion and nutrient absorption, is one of the most popular therapies for chronic pancreatitis. PERT is beneficial in significantly lowering pain, enhancing nutritional status, and improving quality of life in CP patients [9]. PERT is not always successful in all patients, and long-term treatment can have negative consequences such as diarrhoea, bloating, and abdominal pain [10].

Endoscopic therapy, which uses specialized endoscopic methods to clear congestion in the pancreatic duct and enhance pancreatic drainage, is another option for treating CP [11]. Endoscopic therapeutic techniques include sphincterotomy, pancreatic duct stenting, and endoscopic retrograde cholangiopancreatography (ERCP). These operations improve its function by clearing any obstructions in the tubes that link the pancreas to other body regions. Endoscopic therapy has been established in numerous trials to effectively treat chronic pancreatitis patients' pain and enhance their quality of life [12]. The advantages of this treatment, nevertheless, might not last forever.

1.4 Disease Prediction Using Machine Learning

Machine learning (ML) is a technique that industry experts are increasingly using to solve problems that arise in daily life because of technological advancements. Nowadays, machine learning contributes to every sector to advance innovation and increase process efficiency [13]. The health sector is one of the main domains

[4] Gallstones are hard deposits that form in the gallbladder, an organ under the liver. The gallbladder and pancreas are connected to the small intestine through a shared duct system. The gallbladder stores bile, which helps digest fats, and releases it into the bile duct, which combines with the pancreatic duct to form the common bile duct, leading to the small intestine.

among them. There are several types of machine-learning-based research carried out for disease prediction. As a result of machine learning techniques, it is possible to comprehend the symptoms of a specific disease and make accurate and early predictions [14]. These systems create predictions about fresh data using statistical approaches to identify patterns and relationships in the data [30]. In order to develop prediction models for a variety of purposes in healthcare, including disease prediction, diagnosis, and treatment planning, ML models have been extensively used.

Several machine learning techniques can be used to design and develop such prediction models. There are four categories of machine learning algorithms such as supervised, semi-supervised, unsupervised, and reinforcement learning. Supervised machine learning uses labelled datasets to train algorithms for accurate data classification and outcome prediction. Unsupervised machine learning analyses and clusters of unlabelled datasets using machine learning techniques. These algorithms can identify hidden patterns or data clusters without human intervention. Semi-supervised learning follows a medium approach between supervised and unsupervised learning. When considering reinforcement learning, that technique can perceive and understand its environment, act, and learn through mistakes [15]. In this research, the author tries to develop a prediction model using supervised machine learning. If the supervised methodology does not perform as expected, other machine learning approaches will be trialled, and the suitable technique with the maximum accuracy will be determined.

There are several studies carried out to compare the effectiveness of ML algorithms in the medical domain. One such study compared the performance of LightGBM and XGBoost as ML algorithms for disease prediction utilizing laboratory tests [31]. To validate the models, the study used stratified five-fold cross-validation. The outcomes showed that the LightGBM and XGBoost models predicted the five most likely diseases with great accuracy, precision, and recall. The study also showed how important it is to choose ML algorithms that can deal with missing values in the dataset and underlined the potential of ML models in disease prediction. The study adds to the expanding body of research demonstrating the effectiveness of ML algorithms in creating precise and reliable disease predictions, which have the potential to have a big impact on healthcare decision-making and enhance patient outcomes. Furthermore, a systematic review of ML prediction of chronic diseases was carried out to validate the performance of these ML-based methodologies [32]. During the study, it was found that ML models including K-nearest neighbors, Naive Bayes, Deep Neural Networks, and random forest, frequently achieved superior performance in predicting chronic diseases. It also underscored the originality and relevance of machine learning in enhancing clinical decision-making and healthcare organization, presenting a compelling case for the integration of ML models in disease prediction and prevention strategies.

1.5 Importance of Early Detecting Chronic Pancreatitis

One of the primary benefits of early CP detection is the ability to stop disease progression and the emergence of comorbidities [16]. Early symptom relief improved pancreatic function, and decreased risk of complications like diabetes and pancreatic cancer can all be achieved with the proper medical and surgical treatments. Additionally, early diagnosis of CP might assist in identifying people who can benefit from lifestyle changes, such as quitting smoking and alcohol use, which are established risk factors for the condition. Early identification of chronic pancreatitis can also aid in managing related disorders such as pancreatic pseudocysts and pancreatic duct strictures, which, if left untreated, can cause serious morbidity.

The Marseille, Rosemont, and Asia-Pacific criteria are only a few diagnostic standards created to help identify CP early [17]. Chronic pancreatitis is diagnosed using a combination of clinical, imaging, and laboratory factors. In order to help with the diagnosis and staging of chronic pancreatitis, imaging techniques, including endoscopic ultrasound and magnetic resonance cholangiopancreatography, have also been employed. Overall, improving patient outcomes and lowering the disease burden depends significantly on the early identification of chronic pancreatitis.

2 Existing Works

2.1 Traditional Diagnostic Methodologies in the Medical Domain

The diagnosis of chronic pancreatitis has been made by physicians using various methods. When determining the disease's diagnosis, criteria like the patient's symptoms, a history of acute pancreatitis that has occurred repeatedly, alcohol use, and a family history of cystic fibrosis or pancreatitis will be considered first. Other medical imaging procedures will be used to confirm the illness, including magnetic resonance cholangiopancreatography, endoscopic ultrasound, abdominal ultrasound, and endoscopic retrograde cholangiopancreatography (ERCP). Blood tests will also be considered to reach the final diagnosis of the disease in cases of hypercalcemia and hypertriglyceridemia. The ability of the medical expert to diagnose CP early ultimately depends on his or her knowledge of this practice. There is a chance of a later diagnosis because this disease has no distinctive symptoms compared to other diseases. This could result in the patient's health getting worse. An automated system that can detect the early signs of chronic pancreatitis helps overcome these shortcomings and fill the hole left by the conventional diagnosis approach.

2.2 Pancreatitis Detection Systems

Not many studies have been done to create a diagnosis system for pancreatitis. One such study presented a Relevance Vector Machine (RVM)-based intelligent pancreatitis detection method [18]. It makes use of the patients' fundamental data, including their age, gender, liver and kidney function tests, as well as other electrolytes collected 48 hours after the patient was admitted to the hospital. This system can diagnose the existence of AP using the information provided. A dataset containing information from several pancreatitis cases was used to achieve this. This dataset included age, gender, discharge diagnosis, type of illness severity, aetiology, patient medical history, complaints of symptoms, abdominal pain time, 48 hours after admission, and after discharge of blood electrolytes and liver and kidney function test data. In that study, the author created a reasonably accurate system, demonstrating that RVM has the capability needed for disease diagnosis. The conclusion states that additional study is still required to improve diagnosis accuracy.

Another study was conducted to create a classification algorithm using a similar diagnosis strategy as the previous research. The author employed the XGBoost classification method in that system [19]. The constructed classification algorithm was trained using a pancreatitis dataset that includes parameters including ID, gender, smoking, drinking, blood pressure, serum calcium, lipid profile, haemoglobin, age, and kind of disease. In addition to the finished XGBoost algorithm, the author assessed the performances of K-Nearest Neighbour, Random Forest, Decision Tree, and Support Vector Machine on the same pancreatitis dataset.

Both pancreatitis detection systems mentioned above were created using a dataset in tabular format. Another author has adopted an innovative strategy to create a system for detecting pancreatitis by examining CT scan images. In order to identify pancreatitis on abdominal CT scan images by capturing the crucial elements of the local lesion, that study presented a Region-Manipulated Fusion Networks (RMFN) architecture [20]. The created model was trained using individually acquired abdominal CT scans from patients with pancreatitis and healthy individuals. According to the author, three qualified doctors examine each of these individually acquired CT images and divide them into two categories: normal abdominal CT images and abdominal CT images of people with pancreatitis. Before concluding with the implementation of RMFN, this study compares the performance of these picture data with several other models, including SVM, VGG11, VGG16, and AlexNet.

2.3 Multi-disease Prediction Systems

Numerous studies on multi-disease prediction systems have been conducted utilizing various technologies. In many cases, having a general understanding of the ailment will be helpful. Electronic medical records (EMRs) have grown in popularity as a method of patient data tracking due to technological evaluation.

Zhenchao claims that many algorithms based on machine learning have been implemented to categorize and map input predictions, such as symptoms and patient demographics. However, because those methods rely so significantly on the manually labelled EMR training data, the predictions for rare diseases for insufficient labelled data are less accurate. A novel disease prediction model built on Graph Neural Networks (GNNs) has been constructed to fill this gap [21]. This system uses the patient record graph and the medical concept graph as inputs, embedding each node in each graph to aggregate the data of the sample neighbours. Merging the embeddings of symptoms offered in the EMRs for a specific patient aids the system in performing a vector representation. Measuring the closeness between the patient's embeddings and any disease helps estimate the likelihood of the condition. Due to the lack of data on these ailments, it has also used a similar process to detect rare diseases. The study's findings generally indicate that the established model may more precisely identify the various types of uncommon diseases by detecting the underlying relationship between symptoms and disease. Because of a lack of data, forecasts concerning uncommon diseases' overall performance are erroneous. As a result, the model is unable to detect the entire pattern.

Using the lion and butterfly optimization method, a hybrid learning system that has been tuned and can predict multiple diseases was created [22]. The prediction made in this study is only as accurate as the available datasets. This study mostly focuses on disorders like Alzheimer's, dermatological, heart, Parkinson's, thyroid disease, and breast and lung malignancies. This system collects information on numerous diseases from trustworthy sources and uses the lion-based butterfly optimization method to find the best features for each ailment. The original datasets are combined with the attributes chosen most effectively. The lion-based butterfly optimization technique optimized the number of hidden neurons after the feature sets were combined and used to deep belief and neural networks. It improved the system's ability to forecast outcomes accurately. This approach cannot be used to forecast pancreas-relevant disease because it is dependent on the dataset, which is not available. This methodology cannot be used to forecast an illness like chronic pancreatitis because of the similarity of its symptoms with other diseases, even though this system has a dataset and superior prediction accuracy.

2.4 Pancreatitis Related Anomaly Detection Systems

Anomalies connected to pancreatitis have only been the subject of relatively few investigations. In one study, machine learning approaches were used to increase the accuracy of mortality level predictions for acute pancreatitis [23]. According to the author, researchers are employing cutting-edge machine learning and data mining algorithms to identify patients with high mortality risks early instead of depending on manual grading systems in this era of data and technology. Because of this, the author created a new implementation to close the current gaps. This study also drew attention to several flaws in all previous studies on predicting mortality levels in AP. This study clarified that one of the main issues is the dearth of adequate public datasets for AP, which are necessary for proper training in ML models. The present datasets' substantial class imbalance and several missing data points are also mentioned. The author has joined three open datasets, MIMIC-III, MIMIC-IV, and eICU, to create a larger dataset to get around the problem. Additionally, it was found that classifiers created using this pooled dataset outperformed those trained on smaller datasets by a wide margin. In addition, the author employed state-of-the-art innovative approaches to address fundamental data issues. The results showed that iterative imputer is the best way to add missing values to AP data. It performs better than both traditional methods and KNN-based imputation.

Another study looks at the usefulness of the radiomics approach for differentiating Pancreatic Ductal Adenocarcinoma (PDAC) from Autoimmune Pancreatitis (AIP) without the need for invasive treatments [20]. In order to identify the most useful feature set, the study extracted and analyzed multi-domain characteristics. The optimal differentiation strategy was then determined by contrasting several feature selection tactics and machine learning classifiers. The best Radi-M approach was developed due to the investigation, and it was compared to clinical factor-based prediction models and medical professionals. Investigated was the multi-domain features' capacity for discrimination. The findings of this study show that

the non-invasive distinction of AIP and PDAC using PET/CT scans benefits from the integration of multi-domain characteristics.

Acute Necrotizing Pancreatitis (ANP) is a serious type of pancreatitis that develops when the pancreatic tissues die. A study was conducted to create an ANP detection system with artificial intelligence (AI). It is also recognized as the pioneering work that predicts necrosis development in AP using AI. Other research until that point lacked the requisite predictive potential that the created model had. The study also produced a user-friendly web application that can be updated and maintained regularly and predicted necrosis at admission. The prediction interpretation supports the application, making understanding how the algorithm works easier. Overall, this approach might improve the ability to anticipate AP necrosis early, leading to better patient outcomes.

3 Proposed Approach

The author has determined that the supervised deep-learning approach is more appropriate and accurate for systems that produce predictions based on medical images. After comparatively studying supervised deep learning algorithms like Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN), the author decided to move forward with the CNN approach. It is because CNN excels at feature extraction and quickly captures statistical data combined with parameters [24]. Convolution, pooling, activation functions, regularization methods, and optimization algorithms are just a few elements that must be thoroughly understood to design an effective CNN architecture.

Convolutional layers serve as the fundamental unit of the CNN models that extract characteristics from the input images [25]. This layer of the CNN model is thought to be the most important for that reason. The author has effectively extracted features from given MRI and CT images using five convolutional layers with different filter sizes for the CNN architecture created for this work. Which filter sizes to utilize depends on how complex the feature extract is. Because of this, local characteristics are recorded using smaller filters, while global characteristics are collected using distinct sets of huge filters. Batch normalization is a method used to raise the CNN model's training level during the development process. As a result of normalizing the activations of the preceding layer, the network becomes stable throughout training [26].

This strategy improved the generalization and convergence rates of deep neural networks. Thus, the author included a batch normalization layer following each convolutional layer to speed up training and improve the performance of the entire model. Activation functions give the neural network non-linearity, allowing it to recognize complex relationships and patterns in the image collection. Among all the available activations, Rectified Linear Units (ReLU) can be regarded as one of the most common functions. Despite being widely used, this activation function is flawed in treating zero as the gradient for negative inputs, a problem known as the dying ReLU problem [27]. The created CNN model uses the Leaky ReLU activation function to overcome this problem, preventing the gradient from approaching zero by providing a negative slope for negative inputs.

The model then utilizes Max Pooling layers to reduce the dimensionality of the created feature maps. Max Pooling is a downsampling method employed by CNN that reduces the computational cost of the network while also strengthening the model's resistance to minute changes in the input images [28]. Two different 2D max pooling layers were added to the resulting model. The first set of convolutional layers follows the first max pooling layer, while the second set follows the second max pooling layer. Then the maximum value is picked from each non-overlapping subregion of the input feature map to carry out the pooling operation. After the 2D max pooling layer, a dropout layer was applied to combat model overfitting. A regularization technique called a dropout layer prevents complex co-adaptations on training data, which lowers the likelihood that the model would become overfit. Dropouts avoid overfitting by erratically switching some neurons' output to zero during training. The rate parameter determines how likely it is to turn a neuron's output to zero [29]. By doing this, the model is forced to acquire other robust features rather than being overly reliant on one specific set of properties.

The feature maps are then flattened into a single long vector and fed into fully connected layers to categorize the features obtained from the convolutional layers. These are intended to categorize the features taken from the convolutional layers. This architecture uses two entirely interconnected layers. A batch normalization layer follows the first completely linked layer. The batch normalization layer helps to improve the model's speed, performance, and stability by normalizing the output of each layer. A dropout layer is introduced after the first fully linked layer to stop further overfitting. The output of the model is produced by the final fully connected layer, which has two nodes with sigmoid activation functions. Since this is a binary classification problem, the output has been translated into the range of 0 and 1, which can be thought of as a probability, using the sigmoid activation function.

After compiling the designed layers of the CNN model, it was trained using the healthy and pancreatitis MRI and CT image datasets. In order to train the binary classification model, four datasets were used in total. Those four datasets can be classified as healthy abdominal MRI and CT image datasets and Pancreatitis abdominal MRI and CT image datasets. All these medical images were taken from publically available datasets such as radiology and CHAOS medical image datasets. This study aims to train two models for MRI and CT images. Since the contrast between these images is quite different, the author has to move forward with hyperparameter tuning for specified parameters to get the maximum accuracy for each model.

To get the prediction from the trained model, desired abdominal MRI/CT image should be passed into the relevant model in the Numpy array format, and it will output predicted possibilities to the defined healthy and pancreatitis classes. If the healthy class percentage value is quite higher than the pancreatitis value, the medical image more likely belongs to a healthy patient, but if the model output of similar values or values with less difference to both classes, the victim might have a possibility to infect CP in the future. Likewise, the CP possibility can be determined based on those values. The system will show those predicted outcomes and possibilities detected for each class in the application interface.

4 Evaluation

The author evaluated the model to find the most effective parameter value combination by altering potential hyperparameters such as batch size, epoch count, and learning rate on each dataset.

4.1 Batch Size

As the first phase in the testing process, batch size testing was carried out. By testing with different batch sizes and evaluating the effects on test data accuracy and model training speed, it is possible to discover the best batch size. The starting epoch and batch size settings specified by the author are 6 and 11, and the learning rate for the CT image dataset is set at 0.001. The initial epoch and batch size for the MRI dataset was set to 12 and 11, with a learning rate of 0.001. Once the model displayed overfitting behavior, the batch size gradually increased by adding the current value plus four each time. The tables below demonstrate how accuracy and other parameters improved in CT and MRI datasets.

Table 1: Batch size evaluation on CT image dataset

Batch Size	Accuracy	F1-Score	Recall	Precision
11	85.00%	85.00%	85.00%	85.09%
16	89.55%	89.43%	89.55%	91.35%
21	97.27%	97.27%	97.27%	97.34%
26	80.91%	80.19%	80.91%	86.18%
31	94.55%	94.53%	94.55%	95.08%
36	99.09%	99.09%	99.09%	99.11%

Table 2: Batch size evaluation on MRI dataset

Batch Size	Accuracy	F1-Score	Recall	Precision
11	85.68%	85.56%	85.68%	87.27%
16	88.18%	88.06%	88.18%	89.77%
21	99.09%	99.09%	99.09%	99.09%
26	92.27%	92.23%	92.27%	93.31%
31	88.41%	88.20%	88.41%	90.75%
36	84.55%	84.17%	84.55%	88.19%

4.2 Epoch Value

The CT image dataset exhibits high accuracy for batch size 16 without exhibiting any overfitting or underfitting behaviors, according to the evaluation of batch size. In order to further assess the model's accuracy by altering the epoch value, the author has chosen batch value 16 as the batch size. It was done by keeping the same 0.001 learning rate.

Table 3: Epoch count evaluation on CT image dataset

Epoch Count	Accuracy	F1-Score	Recall	Precision
6	89.55%	89.43%	89.55%	91.35%
8	93.12%	92.96%	93.12%	94.87%
9	97.27%	97.27%	97.27%	97.41%
10	98.75%	98.74%	98.75%	98.77%
12	99.09%	99.09%	99.09%	99.11%
14	100%	100%	100%	100%

The MRI dataset also showed a good accuracy level for batch size 16. Therefore, the author tested the accuracy level again by changing the epoch values of the model by keeping the constant batch size at 16.

Table 4: Epoch count evaluation on MRI dataset

Epoch Count	Accuracy	F1-Score	Recall	Precision
6	60.00%	52.38%	60.00%	77.78%
8	66.14%	61.17%	66.14%	79.97%
9	72.27%	69.96%	72.27%	82.16%
10	80.23%	79.01%	80.23%	85.97%
12	88.18%	88.06%	88.18%	89.77%
14	100%	100%	100%	100%

4.3 Learning Rate

After evaluating epoch values and batch sizes, it was found that the CT image dataset showed promising accuracy for an epoch value of 6 and a batch size of 16, while the MRI dataset showed promising accuracy for an epoch value of 12 and a batch size of 16. It can be a result of the fact that the MRI dataset size is marginally smaller than the CT image dataset. The author will assess the model performance for various learning rates by maintaining the promising epoch and batch sizes as a constant.

Table 5: Learning-rate evaluation on CT image dataset

Learning Rate	Accuracy	F1-Score	Recall	Precision
0.001	89.55%	89.43%	89.55%	91.35%
0.0028	93.64%	93.61%	93.64%	94.35%
0.0046	86.36%	86.11%	86.36%	89.29%
0.0064	69.09%	61.68%	69.09%	82.37%
0.0082	51.82%	37.25%	51.82%	75.46%
0.01	50.00%	33.33%	50.00%	25.00%

Table 6: Learning-rate evaluation on MRI dataset

Learning Rate	Accuracy	F1-Score	Recall	Precision
0.001	88.18%	88.06%	88.18%	89.77%
0.0028	92.95%	92.89%	92.95%	93.77%
0.0046	97.73%	97.73%	97.73%	97.76%
0.0064	75.45%	73.88%	75.45%	83.54%
0.0082	85.23%	84.44%	85.23%	89.31%
0.01	95.00%	95.00%	95.00%	95.09%

5 Conclusion

After the comprehensive analysis, it was identified that both datasets produce more evenly distributed accuracy levels for the 0.001 learning rate. Like the learning rate, batch size 16 performs well on both models without indicating overfitting behaviors, but following the analysis of the epoch count, it was discovered that the CT model performs well at epoch value 6, whereas the MRI model performs best at epoch value 12. These hyperparameter values were selected based on the absence of any discernible indications of overfitting within the model. Based on that, the finalized CT model has an accuracy level of 89.55%, while the MRI model has an accuracy level of 88.18%.

6 Future Enhancements

Several future improvements can be considered to raise the prediction accuracy of each model. First, it will be advantageous to increase the dataset by gathering a bigger collection of abdominal MRIs or CT images from pancreatitis victims and healthy people. On this expanded dataset, a CNN binary classification model with additional layers can be trained to determine the accuracy of the prediction. Other than that, it is also possible to create and train an unsupervised ACGAN model to recognize images of pancreatitis from images of the healthy pancreas using a large number of MRI or CT scans. Ideally, getting an abdominal dataset using MRI, CT, or both modalities with more than 2500 images in healthy and pancreatitis victim categories would be preferable.

Acknowledgment

The author would like to express sincere gratitude to the guided experts and evaluators for their valuable feedback and insights throughout the research project. That mentorship and assistance have been crucial in ensuring the successful completion of this study.

References

[1] Dr. Radhika A, “Pancreatitis - Know More Of It!,” 2019. https://www.lybrate.com/topic/pancreatitis-know-more-of-it/afae025fb5837fb096d2f0da1ffb1b57 (accessed Oct. 18, 2022).

[2] Steven D Freedman, “Patient education: Chronic pancreatitis (Beyond the Basics),” 2022. https://www.uptodate.com/contents/chronic-pancreatitis-beyond-the-basics/print (accessed Oct. 18, 2022).

[3] Judith Marcin, “Chronic Pancreatitis: Causes, Symptoms, & Treatment,” 2018. https://www.healthline.com/health/chronic-pancreatitis (accessed Oct. 18, 2022).

[4] M. Ann, “Pancreatitis And Alcohol: Long-Term Effects Of Ethanol On Pancreas,” AlcoRehab.org, 2019. https://alcorehab.org/the-effects-of-alcohol/pancreatitis/ (accessed Jan. 21, 2023).

[5] K. Lavarone, “Can alcohol cause pancreatic cancer? Link and more,” Nov. 24, 2022. https://www.medicalnewstoday.com/articles/what-causes-pancreatic-cancer-alcohol (accessed Jan. 21, 2023).

[6] R. Tee-Melegrito, “Gallstone pancreatitis: Causes, symptoms and treatments,” Mar. 18, 2022. https://www.medicalnewstoday.com/articles/gallstone-pancreatitis (accessed Jan. 21, 2023).

[7] T. W. Frick, “The role of calcium in acute pancreatitis,” Surgery, vol. 152, no. 3 Suppl 1, pp. S157-163, Sep. 2012, doi: 10.1016/j.surg.2012.05.013.
[8] “Hereditary Pancreatitis,” Cancer.Net, Jun. 25, 2012. https://www.cancer.net/cancer-types/hereditary-pancreatitis (accessed Apr. 22, 2023).
[9] I. Trestini et al., “Pancreatic Enzyme Replacement Therapy in Patients Undergoing First-Line Gemcitabine Plus nab-paclitaxel for Advanced Pancreatic Adenocarcinoma,” Frontiers in Oncology, vol. 11, 2021, Accessed: Jan. 21, 2023. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fonc.2021.688889
[10] D. de la Iglesia-García et al., “Efficacy of pancreatic enzyme replacement therapy in chronic pancreatitis: systematic review and meta-analysis,” Gut, vol. 66, no. 8, pp. 1354–1355, Aug. 2017, doi: 10.1136/gutjnl-2016-312529.
[11] J.-M. Dumonceau, “Endoscopic therapy for chronic pancreatitis,” Gastrointest Endosc Clin N Am, vol. 23, no. 4, pp. 821–832, Oct. 2013, doi: 10.1016/j.giec.2013.06.004.
[12] A. Seicean and S. Vultur, “Endoscopic therapy in chronic pancreatitis: current perspectives,” Clin Exp Gastroenterol, vol. 8, pp. 1–11, Dec. 2014, doi: 10.2147/CEG.S43096.
[13] Jay Selig, “What Is the Definition of Machine Learning?,” Mar. 14, 2022. https://www.expert.ai/blog/machine-learning-definition/ (accessed Nov. 07, 2022).
[14] D. K. Rathore and P. K. Mannepalli, “Diseases prediction and classification using machine learning techniques,” AIP Conference Proceedings, vol. 2424, no. 1, p. 070001, Mar. 2022, doi: 10.1063/5.0076768.
[15] Katrina Wakefield, “A guide to the types of machine learning algorithms | SAS UK,” 2022. https://www.sas.com/en_gb/insights/articles/analytics/machine-learning-algorithms.html (accessed Nov. 07, 2022).
[16] M. Jayaraj, “Evaluation and Management of Suspected Early Chronic Pancreatitis (ECP) - PubMed,” 2020. https://pubmed.ncbi.nlm.nih.gov/32592065/ (accessed Jan. 21, 2023).
[17] T. B. Gardner, D. G. Adler, C. E. Forsmark, B. G. Sauer, J. R. Taylor, and D. C. Whitcomb, “ACG Clinical Guideline: Chronic Pancreatitis,” Official journal of the American College of Gastroenterology | ACG, vol. 115, no. 3, p. 322, Mar. 2020, doi: 10.14309/ajg.0000000000000535.
[18] S. Li, S. He, J. Yang, Y. Sun, D. Cheng, and au Shi, “Intelligent pancreatitis diagnosis-based on relevance vector machine,” in 2014 International Conference on Machine Learning and Cybernetics, Jul. 2014, pp. 601–606. doi: 10.1109/ICMLC.2014.7009676.
[19] G. Karteek and # K Yasudha, “For Acute Pancreatitis using Supervised Machine Learning Algorithms,” International Journal for Research in Engineering Application & Management, 2020, doi: 10.35291/2454-9150.2020.0290.
[20] Y. Zhang et al., “Radiomics analysis for the differentiation of autoimmune pancreatitis and pancreatic ductal adenocarcinoma in 18 F-FDG PET/CT,” Medical physics, vol. 46, no. 10, pp. 4520–4530, Oct. 2019, doi: 10.1002/MP.13733.
[21] Z. Sun, H. Yin, H. Chen, T. Chen, L. Cui, and F. Yang, “Disease Prediction via Graph Neural Networks,” IEEE Journal of Biomedical and Health Informatics, vol. 25, no. 3, pp. 818–826, Mar. 2021, doi: 10.1109/JBHI.2020.3004143.
[22] A. Kumar, “Optimized hybrid learning for multi disease prediction enabled by lion with butterfly optimization algorithm | SpringerLink,” Jan. 09, 2021. https://link.springer.com/article/10.1007/s12046-021-01574-8 (accessed Oct. 05, 2022).
[23] M. A. B. Hameed and Z. Alamgir, “Improving mortality prediction in Acute Pancreatitis by machine learning and data augmentation,” Computers in Biology and Medicine, vol. 150, p. 106077, Nov. 2022, doi: 10.1016/J.COMPBIOMED.2022.106077.
[24] S. S. Yadav and S. M. Jadhav, “Deep convolutional neural network based medical image classification for disease diagnosis,” Journal of Big Data, vol. 6, no. 1, p. 113, Dec. 2019, doi: 10.1186/s40537-019-0276-2.

[25] J. Brownlee, “What is Deep Learning?,” MachineLearningMastery.com, Aug. 15, 2019. https://machinelearningmastery.com/what-is-deep-learning/ (accessed Jan. 22, 2023).
[26] S. Ioffe and C. Szegedy, “Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift.” arXiv, Mar. 02, 2015. doi: 10.48550/arXiv.1502.03167.
[27] K. Leung, “The Dying ReLU Problem, Clearly Explained,” Medium, Sep. 23, 2021. https://towardsdatascience.com/the-dying-relu-problem-clearly-explained-42d0c54e0d24 (accessed Jan. 28, 2023).
[28] J. Brownlee, “A Gentle Introduction to Pooling Layers for Convolutional Neural Networks,” MachineLearningMastery.com, Apr. 21, 2019. https://machinelearningmastery.com/pooling-layers-for-convolutional-neural-networks/ (accessed Jan. 28, 2023).
[29] J. Brownlee, “A Gentle Introduction to Dropout for Regularizing Deep Neural Networks,” MachineLearningMastery.com, Dec. 02, 2018. https://machinelearningmastery.com/dropout-for-regularizing-deep-neural-networks/ (accessed Jan. 28, 2023).
[30] T. Papapanagiotou, “Towards MLOps: Technical capabilities of a Machine Learning platform,” *Prosus AI Tech Blog*, Apr. 30, 2021. https://medium.com/prosus-ai-tech-blog/towards-mlops-technical-capabilities-of-a-machine-learning-platform-61f504e3e281 (accessed Jan. 22, 2023).
[31] D. J. Park, M. W. Park, H. Lee, Y.-J. Kim, Y. Kim, and Y. H. Park, “Development of machine learning model for diagnostic disease prediction based on laboratory tests,” Sci Rep, vol. 11, no. 1, Art. no. 1, Apr. 2021, doi: 10.1038/s41598-021-87171-5 (accessed Oct. 04, 2022).
[32] F. M. Delpino, Â. K. Costa, S. R. Farias, A. D. P. Chiavegatto Filho, R. A. Arcêncio, and B. P. Nunes, “Machine learning for predicting chronic diseases: a systematic review,” Public Health, vol. 205, pp. 14–25, Apr. 2022, doi: 10.1016/j.puhe.2022.01.007 (accessed Oct. 05, 2022).

Optimization of Climatic Parameters of the Greenhouse to Maximize Chilli Production

A.P. Wijayawantha

Department of Instrumentation and Automation Technology, Faculty of Technology, University of Colombo, Sri Lanka, pabodiniwijayawantha@gmail.com

M.T.U.D. Mallikarathne

Department of Instrumentation and Automation Technology, Faculty of Technology, University of Colombo, Sri Lanka, mathushani97@gmail.com

A.M.H.M. Abeysinghe

Department of Instrumentation and Automation Technology, Faculty of Technology, University of Colombo, Sri Lanka, hashanmadumal786@gmail.com

C.S.Silva

Department of Instrumentation and Automation Technology, Faculty of Technology, University of Colombo, Sri Lanka, chathurika@iat.cmb.ac.lk

ABSTRACT

Chilli plants are prone to fluctuations in climatic parameters such as temperature, humidity, soil moisture, CO_2 concentration and light intensity. Therefore it is necessary to generate optimum climatic parameters to meet growing demands as a controlled environment provides several advantages. Greenhouse farming is a method where farmers grow crops in controlled environments where all environmental factors are modified according to the type of crop being grown. This research aims to identify the most suitable greenhouse climatic parameters to upgrade the Chilli plant state leading to optimal production. Study indicates that regression and machine learning models can predict corrective climatic conditions leading to improved conditions of Chilli. The novel contribution of this research is to find valid and reliable models that assist growers in predicting the best greenhouse micro-environment.

KEYWORDS

Greenhouse, regression analysis, artificial neural network, temperature, humidity, soil moisture, air quality, mobile application

1 Introduction

It is essential to provide a proper growing environment to keep the plants healthy and prosperous. However, it is challenging to maintain a proper environment for cultivation due to the adverse unpredictable climate changes. The greenhouse offers ideal environmental settings for the growth of crops by dynamically regulating the indoor environment which could improve crop yields and agricultural productivity. The greenhouse measurement and control system is complex, it needs various parameters in greenhouse automatic monitoring, information processing, real-time control and online optimization.

Temperature, humidity, air quality, soil moisture, etc., are the most common greenhouse climate variables considered in most of the research in the literature [1]. Any significant changes in one climate parameter could have an adverse effect on another climate parameter as well as the development process of the plants. Therefore, continuous monitoring and control of these climate factors will allow for maximum crop yield. The temperature should be maintained at an optimum level whenever possible since it influences most plant development processes including photosynthesis, transpiration, absorption, respiration and flowering. The incorrect temperature can compromise the humidity levels, cause the plants to stretch, or even affects the plants' ability to properly absorb CO_2. Humidity controls the moisture loss from the plant. High humidity cause to suffer nutrition deficiencies while low temperature adversely affects transpiration rates. CO_2 is one of the essential factors for plant respiration and microbial activities. The health of crops relies upon an adequate supply of moisture and soil nutrients, among other things. Therefore, it is crucial monitoring these greenhouse parameters continuously for maximum crop yield. Therefore, there is a necessity to implement an approach which could make informed decisions about the climatic environment of the greenhouse based on the real-time monitoring of various environmental factors within the greenhouse, such as temperature, humidity, light intensity, and soil moisture.

The goal of this research study is to predict the most suitable greenhouse climatic parameters to upgrade the Chilli plant state leading to optimal production. The main contribution of this research is to find a regression model to predict the linear dependencies and a machine learning model to investigate the nonlinear dependencies that will help growers in predicting the best greenhouse microenvironment.

2 Background

Crop growth is the result of the comprehensive impact of several environmental factors. However, most of the greenhouse climate control systems in the literature generally focus on controlling a single environmental factor at a time. Therefore, controlling a single parameter usually has difficulty in meeting the requirements for optimal crop growth. Monitoring and controlling multiple microclimate parameters is essential to overcome this disadvantage. However, multi-parameter control is always challenging in agricultural engineering due to its inherent complexity [2].

It is essential to monitor the microclimate parameters inside the greenhouse to make sure that the environment inside the greenhouse is suitable and most efficient for plant growth. The microclimate parameters inside the greenhouse have very complex dynamic dependencies on each other. Every change in one of them can change the others. As an example, humidity CO_2 and soil moisture affect the temperature in a greenhouse. When temperature increases relative humidity decreases and vice versa means there is an indirect proportion between the temperature and the relative humidity [3]. There is a direct correlation between leaf temperature and the amount of carbon dioxide the plants can absorb. Moreover, there is a direct proportion between the CO_2 and the temperature such that when the temperature decreased the CO_2 concentration decreases. Moreover, there is also a direct proportion between CO_2 and humidity such that when humidity decreased the CO_2 concentration decreases [4]. Increased CO_2 concentration increases temperature, which increases transpiration and hence depletes soil moisture [5]. Surface soil Moisture increases with relative Humidity and temperature is affected by soil moisture [6].

There are both linear and non-linear dependencies between microclimate parameters. Therefore it is essential to use appropriate methods which can monitor these linear and non-linear dependencies between them. Regression is a method which can identify linear dependencies between the climate parameters. Limited research has been reported in the literature on using regression analysis for greenhouse monitoring. They are also limited to analyzing one or two climatic parameters such as temperature and CO_2 [7]. Machine learning methods such as ANN (Artificial Neural network) have been used in the literature to investigate the non-linear dependencies between the climatic parameters of the greenhouse [8].

The greenhouse climatic factors are hard to control manually inside a greenhouse and then there is a need for automated design arises. When considering previous approaches to greenhouse automated monitoring, some are based on Arduino [9] [1] [10] some are based on Raspberry Pi [11] [12] and there are approaches based on PID controllers [2]. Temperature, humidity and soil moisture are the commonest parameters that are monitored in these works [1] [9] [10] [2]. In addition to the above parameters, some approaches have also monitored the air quality and pH value of soil [12] [11] [13]. In [9] [1] [10] [12] LDR has been used to monitor light intensity and DHT22, and DHT 11 has been used to monitor humidity and temperature. A lux meter and a digital thermometer are also used in the approach [9]. In [11] a CO_2 gas sensor has been used. The pH sensors and LM35 have been used to monitor the pH value and temperature of the soil [13]. In some approaches, collected data from the sensors are sent to the LCD screen for use [9] [11] [12]. Research and implementation of WSN (wireless sensor networks) in greenhouse environment monitoring and controlling is carried out popularly over the last years. Compared to wired systems, the installation of WSN is fast, cheap and easy. The work mentioned in [14] has been focusing on developing a cloud-based platform for greenhouse monitoring. Methods to control parameters such as temperature, light and soil moisture have also been suggested in some approaches [9] [1] [11]. There are a variety of mathematical models that have been developed to simulate the climate conditions within a greenhouse. These models could take into account climatic factors to predict how these conditions will change over time. These models can be used to optimize the greenhouse environment for crop growth and can help to identify ways to improve energy efficiency. A mathematical model for greenhouse monitoring and control can be developed using a variety of techniques, such as physics-based modelling, data-driven modelling, and model predictive control. However, modelling of the effect of multiple climatic parameters on the growth of the plants grown in the greenhouse is very limited.

3 Methodology

The methodology used in this research is constructive research and it allows to expand scientific understanding to solve practical problems. The primary data collection method [15] is used and the data is directly collected by taking readings from the attached sensors to the greenhouse prototype.

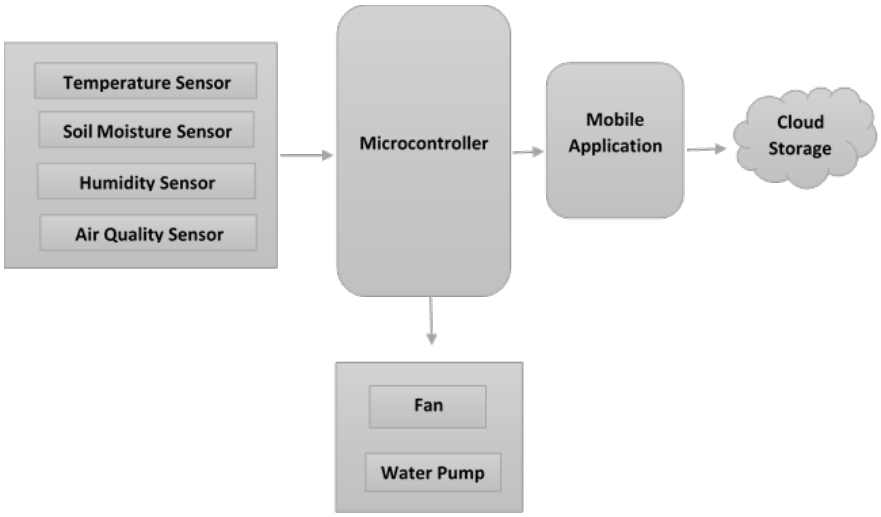


Figure 1: Hierarchical architecture of the approach

The proof of concept prototype consists of the following units as shown in Figure1. Evaluation of the proposed approach based on simulation and a physical implementation of the greenhouse prototype in the real-world environment.

4 Implementation

The simulation model of the computerized monitoring and control system for greenhouse microclimate parameters has been developed and tested in the environment of Proteus 8 Professional. Figure 2 introduces the schematic Proteus simulation model.

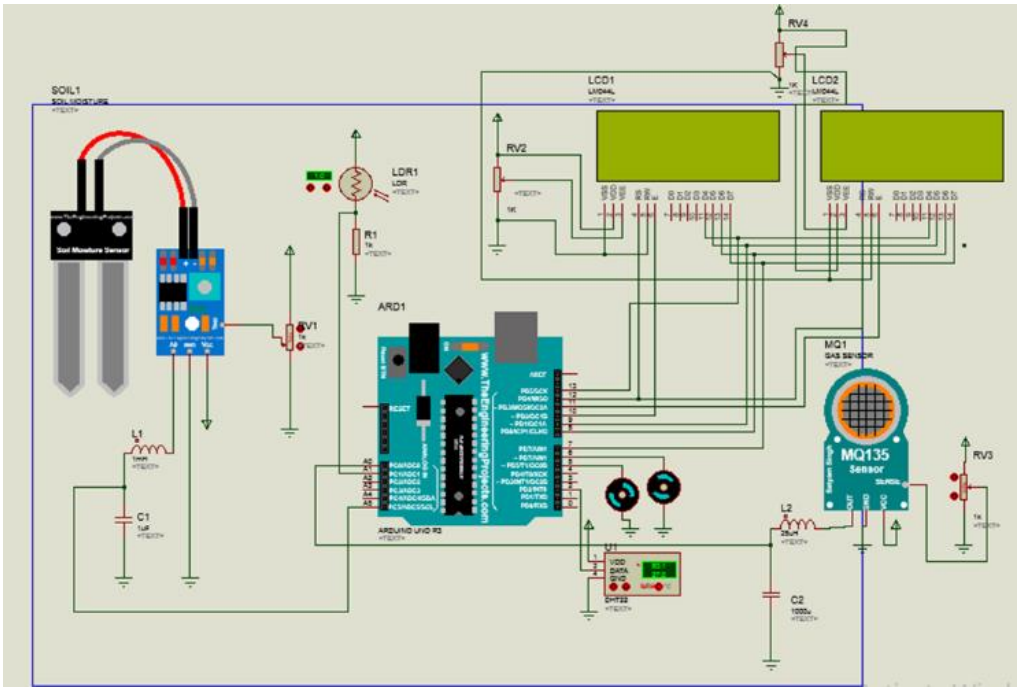


Figure 2: The schematic diagram of the monitoring system

The Chilli plant variant called Kochchi (Capsicum frutescens) is selected to be grown within the greenhouse prototype. A prototype of a greenhouse with a width of 1 meter and a length of 1.5 meters and a height of 1.5 meters is installed. For the irrigation system, a drip irrigation system installed in the water tank is used. One 220V ventilation fan is used. Due to the rise of the hot air, the fan is placed on the top of the greenhouse and used to transfer air from inside to outside. The 220V-5V relay is used to control the fan.

The target temperature for the Chilli plant inside the greenhouse is fixed at 70-80°F. The inner greenhouse's relative humidity is automatically adjusted to 60%-80%. The target CO_2 concentration within the greenhouse prototype is 340-1000 ppm. All the above-mentioned microclimate parameters of the inner greenhouse are automatically adjusted by controlling the cooling fan and

water pump as follows. The water pump is controlled by the average soil moisture. The water pump opens when the average soil moisture falls below the ideal soil moisture at intervals of 2 minutes to prevent excessive watering. The fan will be on when the air temperature or air humidity exceeds the ideal value. The fan is closed when any of the air humidity and air temperature reaches the ideal value. In addition, air temperature information is collected every 30 minutes to prevent sudden temperature fluctuations during the day and night. The ambient light sensor in the system measures the amount of light the plant receives during the day where an ideal threshold value is set for opening and closing the light.

Sensor placement is also very important, and the optimal placement of sensors throughout the greenhouse operation to meet the monitoring goals. In this work, four different sensors a temperature sensor, humidity sensor, air quality and soil moisture sensor have been used for collecting data in real-time. Each sensor is designated to measure a specific parameter of the greenhouse microclimate environment. The sensors are reporting their collected data to the Arduino Uno R3 microcontroller in a real-time manner. The DHT11 digital temperature and humidity sensor measures both the humidity as well as the air temperature. The combination of humidity and temperature sensing in one device saves physical space on the sensor station and also limits the number of input pins being used on the microcontroller. This sensor is operated within a 3.3V-5V power supply. The DHT11 sensor is equipped with a calibrated digital signal output. The MQ-135 gas sensor is used to sense CO_2. The operating voltage of this gas sensor is from 2.5V to 5.0V. An FC-28 soil moisture sensor has been used to determine the moisture of the soil. The value of moisture in the soil can be measured by placing two probes of soil moisture sensors inside the soil.

Remote monitoring is a practical solution to constantly monitor the greenhouse and equipment status. Remote monitoring systems are operating 24/7 and alert on environmental changes that threaten the plants. The sooner discovers a problem with the plants, the more inventory can be saved by taking preventive action. Therefore it is important to ensure continuous monitoring of environmental conditions with a remote monitoring system.

A smartphone application is created as shown in Figure 3 using the MIT app inventor and Firebase is selected as the database for smartphone applications created. MIT App Inventor is a web-based tool for creating mobile applications for Android devices. Google Firebase is a Google-backed application development software that enables the real-time database to let the developers store and sync data in real time.

A cloud system used to provide real-time status of all monitored conditions. With a cloud-based system, the device stores all of the data in the cloud. The users of the greenhouse have access to the data from any internet-connected device, such as a tablet or phone, and they can check the greenhouse status at any time. An alert will send to the owner when monitoring systems identify an abnormal change in status. The alerts are provided via a text message as an extra assurance. Arduino board and an extern WiFi module are used as a slave to the cloud computing algorithm.

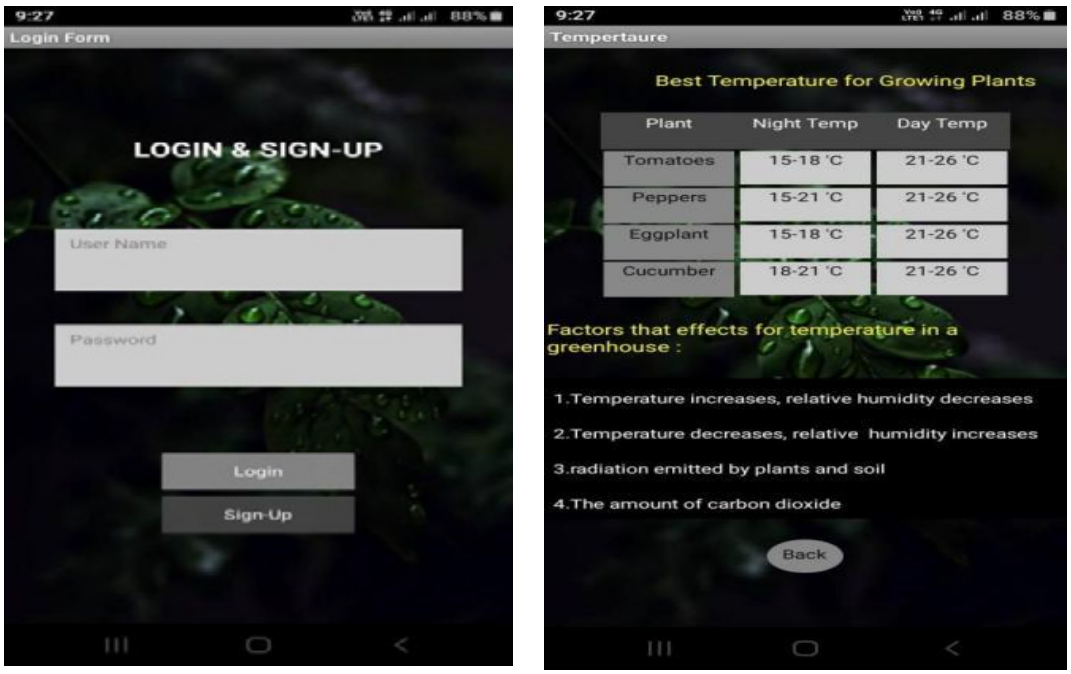


Figure 3: UIs of the smartphone application

Data history is valuable in identifying patterns and trends in environmental conditions. However, manually monitoring and recording environmental parameters takes a significant amount of personnel time. In this work Cloud-based logging is used to provide an unlimited number of records to view, graph, print and export data trends.

5 Evaluation

A. Regression model

In this research there are two main intentions for performing regression analysis:

- Perform linear regression that will predict a dependent variable using one independent variable. Both the correlation coefficient and coefficient of determination are estimated.

- Perform a multiple regression model with the use of multiple dependent variables. The coefficient of determination is estimated to identify the strength of the model apart from the estimation of the correlation coefficient between the variables.

In this study t-test, correlation coefficient and coefficient of determination are selected to evaluate the linear relationship between variables/attributes and the forecast accuracy of the proposed regression model.

t-test- The significance can be tested by performing a t-test for the regression slope. The t statistic is the coefficient divided by its standard error. The standard error is an estimate of the standard deviation of the coefficient, the amount it varies across cases. Pr(>|t|) gives the p-value for that t-test. The p-value describes the amount of confidence that each variable has some correlation with the dependent variable.

Correlation coefficient-These indicates the strength of the linear relationship between two different variables. Correlations are good for identifying patterns in data. A positive correlation signifies that both variables move in the same direction. A negative (inverse) correlation is an indication that both variables move in the opposite direction. When the correlation coefficient is near 1 or −1, the linear relationship is strong; when it is near 0, the linear relationship is weak.

Coefficient of determination (R^2)- The strength of the model can be identified using the coefficient of determination. Positive R^2 indicates that the predictions match the observations perfectly, 0 indicates that the predictions are as good as random guesses around the mean of the observed values and negative R^2 indicates that the predictions are worse than random.

T-test, coefficient of determination (R2) and correlation coefficient are selected to the forecast accuracy of the regression model.

The coefficient of determination is estimated to identify the strength of the model apart from the estimation of the correlation coefficient between the variables.

B. ***ANN Model***

ANN model is used to evaluate the nonlinear relationship between the climatic parameters. In this work, an artificial neural network with gradient descent and momentum back-propagation learning algorithm is developed. The dataset is composed of 270 instances acquired via sensors that measure light intensity, soil moisture, carbon dioxide, temperature, and humidity. The dataset aims to classify the corrective measures needed into 2 classes (1: environment correct – with 120 instances, 2: environment incorrect – with 150 instances (9%) using 5 numerical attributes namely: relative humidity, light intensity, soil moisture, temperature, and air quality. The network configuration consists of three inputs, two outputs and one hidden layer with ten neurons. Out of the total data 70% is allocated for training, 15% for testing and 15% for validation is allocated in ANN. The normalized sensor measurements between [0 1] which are used as training data are fed as input to the ANN. The accuracy of a learning algorithm is improved by tuning for multiple training iterations of ANN.

6 Results and Discussion

A. Regression Model

The following regression results are obtained using R Studio software. R Studio is a popular integrated development environment (IDE) for the R programming language, which is commonly used for statistical computing and data analysis.

A. Construction of Linear Regression Model

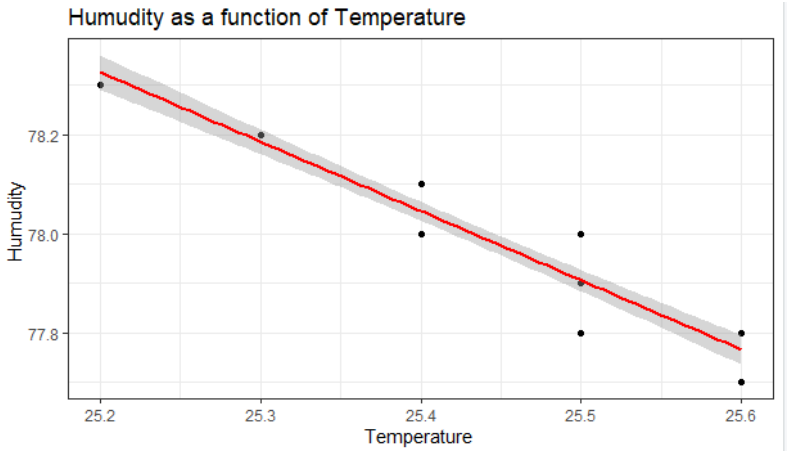


Figure 4: Regression plot between humidity and temperature

Figure 4 shows the linear regression between humidity and temperature. Equation (1) shows the correlation coefficient between temperature and humidity. A negative (inverse) correlation (-1.4) occurs and this is an indication that both variables move in the opposite direction. This proves the fact that humidity and temperature have an indirect proportion between them.

$$y = 110 - 1.4\,x \qquad (1)$$

Figure5 shows the linear regression between soil moisture and temperature.

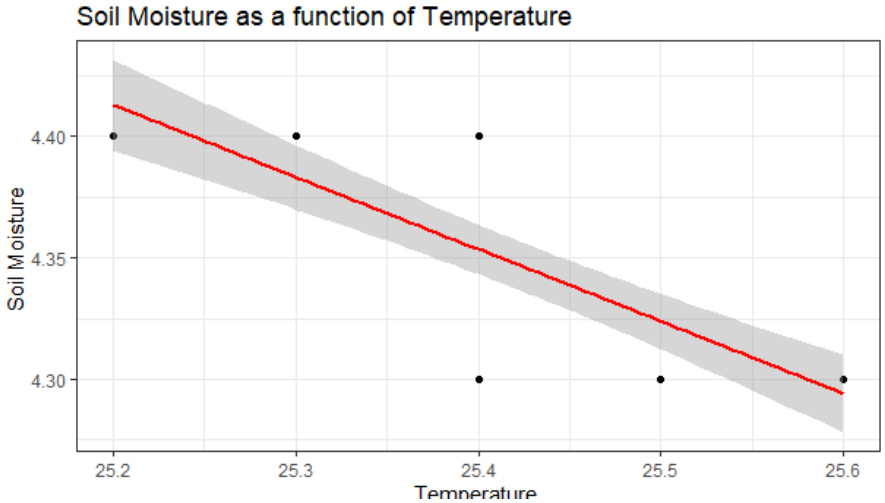


Figure 5: Regression plot between soil moisture and temperature

Equation (2) shows the quantitative linear relationship between them. A negative (inverse) correlation (-0.3) occurs and this is an indication that both variables move in the opposite direction. This proves the fact that soil moisture and temperature have an indirect proportion between them.

$$y = 12 - 0.3\,x \qquad (2)$$

B. Construction of Multiple Linear Regression Model

A linear model is constructed by naming temperature as the dependent variable and humidity, soil moisture, and air quality as the independent variables. Following (3) represents the linear model equation between the dependent and independent variables.

Temperature = Intercept + Humidity (x_1) + Soil Moisture (x_2) + Air.Quality (x_3)

Table 1: Statistical analysis of the independent variables

	Estimate
Intercept	53.55205
Humidity	-0.17666
Soil Moisture	-0.95837
Air Quality	0.17854

$$Y = 53.552 - 0.1766\,x1 - 0.9583\,x2 + 0.1785\,x3 \quad (3)$$

According to Table 1, the value of the intercept is 53.55 which is the value which predicts the value of the dependent variable (temperature) when all the independent variables are equal to zero.

Table 2: Parameters of multiple linear regression model

	Std.Error	t value	Pr(>|t|)
Intercept	10.67337	5.017	3.21e-05
Humidity	0.15536	-1.137	0.266
Soil Moisture	0.91161	-1.051	0.303
Air Quality	0.03625	-4.925	4.09e-05
Residual Standard Error	0.1226 on 26 degrees of freedom		
Multiple R-squared	0.914		
Adjusted R-squared	0.9041		
p-value	5.643e-14		

The final four rows in Table 2 are model diagnostics – the most important thing to note is the p-value which indicates whether the model fits the data well or not. The null hypothesis can be rejected as the p-value (here it is 5.643e-14) is less than 0.001 and conclude that there is a statistically significant relationship between the dependent variable and the independent variables. R-squared explains to what extent the variance of the dependent variable explains the variance of the independent variable. This model gives 0.9041 for multiple R-squared and 0.9041 for adjusted R-squared, which is approximately equal to one means that the observed variation is fully explained by the model's inputs.

The R-value on the regression plot means the coefficient of correlation. It shows how well the predicted outputs are matching with real outputs. According to the regression plot shown in Figure 6 of the neural network, it is clear that the R which represents all (training - 0.99994, validation - 0.99996 and testing – 0.99986) is 0.99993 which is very close to 1. This shows there is proper matching between predicted outputs and real outputs.

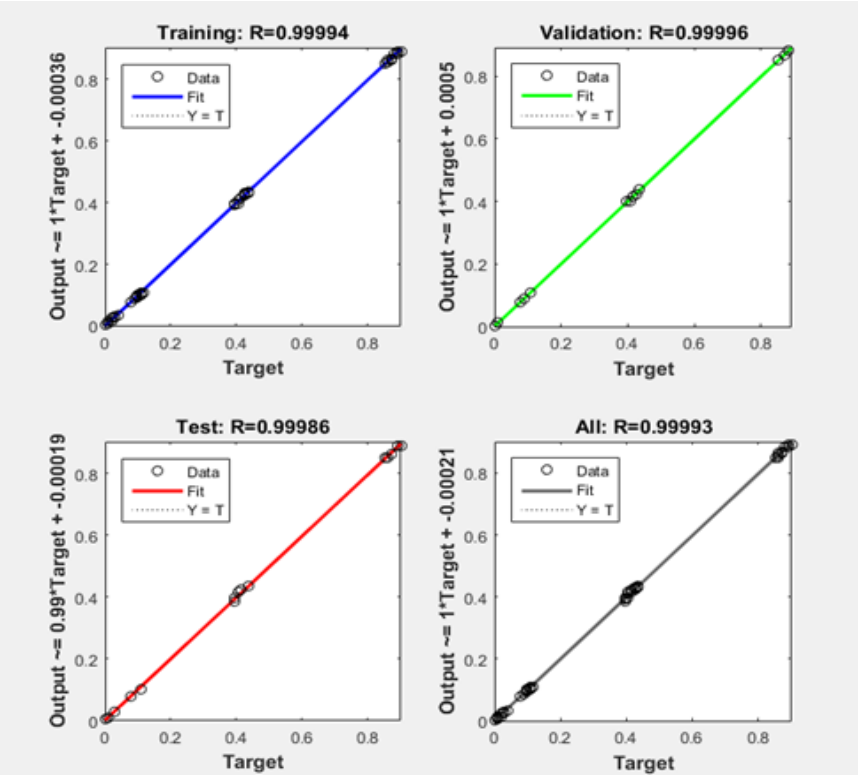


Figure 6: R-value on the Regression plot of ANN

7 Conclusion

The goal of this research study is to predict the most suitable greenhouse climatic parameters to upgrade the Chilli plant state leading to optimal production. A proof of concept prototype of a greenhouse is implemented with sensors and controllers. The paper presents the ANN model and regression model developed based on the monitoring of the climatic parameters of a greenhouse. The regression model shows the linear relationship between the climate parameters. The machine learning model shows the nonlinear dependency between the parameters. Moreover, the ANN model shows the accuracy between the real and the predicted climate parameters. The novel contribution of this research is to develop a linear and non linear models that assist growers in predicting the best greenhouse micro-environment.

8 Future Enhancements

In future, the same study will be repeated with other Chilli verities in Sri Lanka and will compare and contrast the current results with them. Gradually the feasibility of generalizing this research to other crops will be evaluated. It is expected to incorporate some sensors and controllers to improve the physical implementation of the greenhouse such as soil pH sensors, pressure regulators to control water pressure, digital timers to control water dripping, and digital flow meters to observe the water flow. Vision sensors will be incorporated to monitor the status of the plant conditions in the greenhouse. Moreover, sensor fusion can improve the overall performance of the system by reducing the limitations of each sensor and providing a more complete picture of the environment. This can lead to more efficient and productive greenhouses, with better-quality crops and lower costs.

9 References

[1] M. Anisha, U. Arsad, J. Starly, K. Dhanalakshmi, S. Anitha, M. Benisha, P. Chezhiyan and C. Ellio, "Arduino based Low-Cost Greenhouse Monitoring System for Small Scale Farmers," in *Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*, Tirunelveli, India, 2021.

[2] L. Wang and B. Wang, "Greenhouse microclimate environment adaptive control based on a wireless sensor network," *International Journal of Agricultural and Biological Engineering,* vol. 13, no. 3, p. 64–69., 2020.

[3] California Lightworks, "How Do You Maintain Greenhouse Temperature? " 28 June 2020. [Online]. Available: https://californialightworks.com/blog/how-do-you-maintain-greenhouse-temperature/. [Accessed 22 August 2022].

[4] L. YuLiang, N. Xiang, W. YuanXi and Z. Wei, "Study on the effect of temperature and humidity on the CO2 Concentration Measurement," in *IOP Conference Series: Earth and Environmental Science*, 2017.

[5] A. LEIPPRAND and D. GERTEN, "Global effects of doubled atmospheric CO2 content on evapotranspiration, soil moisture and runoff under potential natural vegetation," *Hydrological Sciences Journal,* vol. 51, no. 1, pp. 171-185, 2006.

[6] S. Ravi, P. D'Odorico, T. Over and T. Zobeck, "On the effect of air humidity on soil susceptibility to wind erosion: The case of air-dry soils," *Geophysical Research Letters,* vol. 31, no. 9, 2004.

[7] H. Krug and H. Liebig, "Static Regression Models for Planning Greenhouse Production," in *ISHS Acta Horticulturae 230: Symposium on High Technology in Protected Cultivation*, 1988.

[8] N. Kitpo, Y. Kugai, M. Inoue, T. Yokemura and S. Satomura, "Internet of Things for Greenhouse Monitoring System Using Deep Learning and Bot Notification Services," in *2019 IEEE International Conference on Consumer Electronics (ICCE)*, Las Vegas, NV, USA, 2019.

[9] M. H. Jewel, M. N. Mostakim and N. Bhuiyan, "Construction and Development of an Automated Greenhouse System Using Arduino Uno," *International Journal of Information Engineering and Electronic Business,* vol. 9, no. 3, pp. 1-8, 2017.

[10] Chwana, I. Nasution, S. Sundari and N. Rifky, "Data Acquisition of Multiple Sensor in Greenhouse Using Arduino Platform," in *International Conference of Sustainability Agriculture and Biosystem*, West Sumatera Province, Indonesia, 2020.

[11] J. Arshad, R. Tariq, S. Saleem, A. Rehman, H. Munir, N. Golilarz and A. Saleem, "Intelligent greenhouse monitoring and control scheme: An arrangement of Sensors, Raspberry Pi based Embedded System and IoT platform," *Indian Journal of Science and Technology,* vol. 13, no. 27, pp. 2811-2822, 2020.

[12] H. Riskiawan, S. Anwar, S. Kautsar, D. Setyohadi and S. Arifin, "On-line monitoring system in greenhouse area for chrysanthemum cultivation based on raspberry pi and IoT," in *The 3rd International Conference on Food and Agriculture*, Jember, East Java, Indonesia, 2021.

[13] M. Lavanaya and R. Parameswari, "Soil Nutrients Monitoring For Greenhouse Yield Enhancements Using Ph value with Wireless Sensor Network," in *2018 Second International Conference on Green Computing and Internet of Things*, Bangalore, India, 2018.

[14] M. Xiang, "Design of Intelligent Agricultural Greenhouse Monitoring System Based on 5G network cloud Platform," in *International Conference on Electronic, Electrical and Computer Applications*, Osaka, Japan, 2020.

[15] K. Lukka, "The Constructive Research Approach," in *Case study research in logistics*, Publications of the Turku School of Economics and Business Administration, 2003.

Machine Learning Approach to Predict Next Transaction Day with Customer Lifetime Value

Ms. Punsisi Kaludewa, Department of Computing, Informatics Institute of Technology, Colombo, Sri Lanka, punsisi.2018861@iit.ac.lk

Ms. Rukshala Weerasingha,, Department of Computing, Informatics Institute of Technology, Colombo, Sri Lanka, rukshala.w@iit.ac.lk

ABSTRACT

Analyzing consumer transactional data is crucial for businesses to strengthen customer relationships and attain sustainability in a volatile market. Customer transaction predictions currently heavily depend on human intuition and are considered an intellectual task, formulating and implementing this intuition as a data-driven approach can be valuable for companies in improving customer satisfaction and retaining their customer base. Although there are techniques available to analyze transaction trends, predicting the next transaction day has not been achieved to a satisfactory level. In this study, we propose a novel machine learning-based algorithm to predict the next transaction day of customers based on their customer lifetime value (CLV). The proposed algorithm aims to fill this gap by exploiting historical transactional data to its highest potential, resulting in an accurate prediction mechanism.

Keywords: CLV, Machine Learning, Predictive Modeling, Data Science.

1 Introduction:

In today's data-driven era, consumer-facing companies are realizing the power of consumer data exploitation. Targeted marketing has become increasingly popular among consumer-facing companies as they recognize the potential of leveraging consumer data. Companies try to obtain a competitive edge by anticipating when a consumer might need their goods or service. A survey by Epsilon [1] revealed that 80% of consumers are inclined towards doing business with companies that provide them with personalized experiences. One crucial aspect of this data exploitation is analyzing consumer transactional data, which can help businesses make decisions to improve customer relationships. Predicting the next transaction day of customers is especially critical for businesses to optimize inventory management, plan marketing campaigns, and personalize the customer experience [2]. In fact, Amazon's anticipatory shipping [3] is an excellent example of how next transaction day prediction can help businesses achieve a competitive advantage.

1.1 Target Marketing

Target marketing involves identifying and directing marketing efforts toward a specific group of consumers who are likely to be interested in a particular product or service. This approach relies on data on consumer demographics, psychographics, and purchasing behavior to create a profile of the ideal customer for a product. Targeted marketing offers several advantages over traditional mass marketing, such as saving time and resources by focusing efforts on consumers who are most likely to be interested in a product. Additionally, targeted marketing campaigns can be tailored to the specific needs and preferences of individual consumers, making them more effective at driving sales [4].

1.2 Customer Lifetime Value(CLV)

Customer segmentation based on CLV is an effective way to forecast the next customer purchase/transaction day, as it allows marketers to focus their efforts on the customers who are most valuable to the business. CLV measures the entire value that a customer will bring to a business over the course of a customer. By segmenting customers based on their CLV, marketers can prioritize their marketing efforts on those consumers who are likely to benefit the business, both in terms of revenue and long-term loyalty [5].

Furthermore, CLV-based customer segmentation can help businesses identify and address issues with customer retention. By identifying customers who are at risk of churning, or leaving the business, marketers can develop targeted retention campaigns that address the specific needs and concerns of those customers, increasing the likelihood that they will stay with the business and continue to make purchases [6]. Targeting these valuable customers not only helps to increase revenue but also promotes long-term customer loyalty.

1.3 Next Transaction Day Prediction

The COVID-19 pandemic has led to an increase in online transactions and e-commerce. As a result, companies are using analytics and business intelligence to gain insight into customer needs and offer better services [7]. Targeted marketing has become popular, where predictive modeling is used to forecast when a customer might require a product. For example, Amazon's anticipatory shipping system uses predictive modeling to forecast which products customers are likely to order in the future. This strategy reduces shipping times, optimizes inventory management, and reduces delivery costs for Amazon [3]. Similarly, next transaction day prediction systems can help businesses optimize inventory management by forecasting demand, reduce delivery times by anticipating orders, and improve customer experience by personalizing recommendations and marketing campaigns [8]. These systems can also identify high-value customers and increase customer lifetime value.

2 Existing Work

2.1 CLV Prediction

Several studies [2][9][10] have focused on predicting CLV using various machine learning techniques. Researchers have used a statistical model called “buy-till-you-die” with a neural network model. The statistical model uses past purchase summary statistics to predict customer behavior while the neural network model uses summarized data such as lifetime duration, number of historical purchases, average gaps between purchases, etc. The paper provides a quantitative and qualitative analysis of those two methods as well as recommendations on how to proceed in different cases and opportunities for future research [9].

Similarly, a study[10] proposes a machine learning approach for predicting the repurchase time of customers in e-commerce where the authors create features from customer recency, frequency and monitory(RFM) features from the dataset to train and test different machine learning models. The aim is to reduce the absolute error between the predicted and actual repurchase time of customers. The results show that the Random Forest regression model performs the best among the models tested, and that the RFM features are effective

in improving the prediction performance. The authors conclude that their approach can be useful for e-commerce businesses in predicting and optimizing customer repurchase time.

RFM[10] is a customer segmentation framework that considers three main factors,

- Recency- how recently a customer has made a purchase
- Frequency- how often a customer makes purchases
- Monitory- how much a customer spends on purchases

In a recent research [2], a combined approach of two prediction models, RFM and machine learning-based regression models, was used to estimate the customer lifetime value. The results showed that the combined approach outperformed the individual models in identifying profitable customers. Based on the findings, the authors suggest that using a combined strategy can enhance customer targeting and increase profitability for retailers.

These approaches incorporated various customer attributes and transactional data to develop accurate predictive models, enabling businesses to identify high-value customers and personalize marketing campaigns to maximize their revenue and profitability.

2.2 Transaction Day Prediction

There are several ongoing research studies where machine learning algorithms have been used to predict the next transaction date for customers based on their transaction history and other demographic information. The studies have used different techniques to develop the Next Purchase Date (NPD) predictor systems [11]. These techniques are Recurrent Neural Networks (RNN), Linear Regression (LR), Artificial Neural Networks (ANN), and Extreme Gradient Boosting (XGB). The authors have concluded that the NPD predictor system can predict when a customer will buy fast-moving retail products with a high degree of accuracy. The system can be used to personalise discount offers to customers when they are about to purchase items. Such offers will be delivered on the mobile devices of participating customers and ultimately physical, general paper-based marketing will be reduced.

In another study,[12] used a neural network algorithm to predict the next transaction date for customers of a mobile phone service provider based on their transaction history. Neural networks are a powerful class of machine learning algorithms inspired by the structure and function of the human brain.

There are studies [7] [13] which used a random forest algorithm to predict the next purchase date for customers on their transaction history and other customer attributes. Random forests are an ensemble learning method that combines multiple decision trees to improve prediction accuracy. The studies mentioned above primarily focus on predicting whether a customer will make a purchase within a specific range of dates, rather than predicting the date or date-range of the customer's purchase. This gap will be addressed in the proposed algorithm.

3 PROPOSED ARCHITECTURE

Currently, customer transaction predictions heavily rely on human intuition and are considered purely intellectual. However, if this intuition can be formulated and implemented, it can aid companies in achieving customer satisfaction and maintaining customer retention. Previous research in this field has focused primarily on predicting whether a customer will perform a transaction on a specific day, rather than predicting the exact next transaction day. To address this gap, a proposed architecture has been developed.

Before building the prototype, the author conducted a requirement survey and identified the need for customer segmentation to predict the next transaction date. To address this, the author proposed a novel architecture that combines unsupervised and supervised learning. The first step involves clustering customers using unsupervised learning [14]. The resulting clusters will then be used to train a supervised learning model to make the final predictions. By using unsupervised learning to group customers based on common characteristics, the proposed architecture enables more accurate predictions for the next transaction date. Supervised learning then leverages this customer segmentation to develop a tailored model that is trained on labeled data.

3.1 Data Modelling

	Invoice	StockCode	Description	Quantity	InvoiceDate	Price	Customer ID	Country
0	489434	85048	15CM CHRISTMAS GLASS BALL 20 LIGHTS	12	2009-12-01 07:45:00	6.95	13085.0	United Kingdom
1	489434	79323P	PINK CHERRY LIGHTS	12	2009-12-01 07:45:00	6.75	13085.0	United Kingdom
2	489434	79323W	WHITE CHERRY LIGHTS	12	2009-12-01 07:45:00	6.75	13085.0	United Kingdom
3	489434	22041	RECORD FRAME 7" SINGLE SIZE	48	2009-12-01 07:45:00	2.10	13085.0	United Kingdom
4	489434	21232	STRAWBERRY CERAMIC TRINKET BOX	24	2009-12-01 07:45:00	1.25	13085.0	United Kingdom

Figure 1: Sample of unprocessed dataset

The dataset utilized in this paper was obtained from a UK-based registered online retailer[15]. The dataset consists of 8 variables (attributes) which are detailed in Figure 1. The data includes all transactions that took place in the years 2010 and 2011, with a focus on transactions that occurred between December 1st, 2010, and November 30th, 2011. During this time period, there were a total of 18,291 valid transactions. It is worth noting that obtaining financial data requires authorization from centralized financial bodies, which can be challenging. Hence, this dataset was selected primarily to demonstrate the concept rather than representing actual financial data.

The data will be parsed to extract “CustomerID”, “InvoiceDate”, “Quantity”, “UnitPrice”, and “InvoiceNo” fields. The total transaction amount for each invoice will be computed. Then, for each customer and invoice date combination, the sum of all invoice amounts will be calculated, resulting in the total amount spent by the customer on a given transaction day.

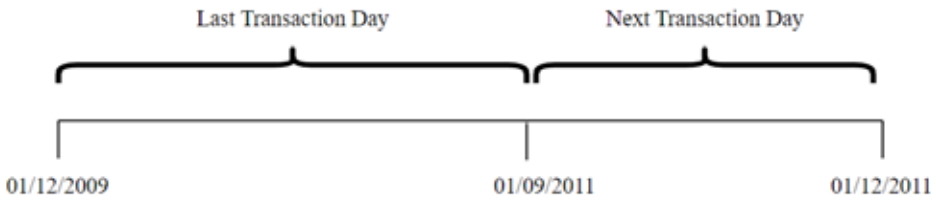


Figure 2: Data division based on day range

The dataset is divided into two parts to create the independent variable for the model. The first part involves obtaining the last transaction date of each customer, while the second part involves obtaining the next transaction date of each customer. This variable will serve as the dependent variable for the model and will be used to predict future customer behavior.

3.2 Feature Engineering

In addition to selecting relevant features from the dataset, new features were also created to gain a deeper understanding of customer behavior using the RFM model. The RFM model is a behavior-based approach used to analyze customer behavior and make predictions based on the data in the dataset[16] which will be used to segment customers and measure the CLV of each customer to check the customer loyalty.

The RFM model has a long history of being used practically in a variety of industries, especially direct marketing. Decision-makers can effectively identify valuable consumers and create efficient marketing plans to optimize client lifetime value by employing the RFM model [17].

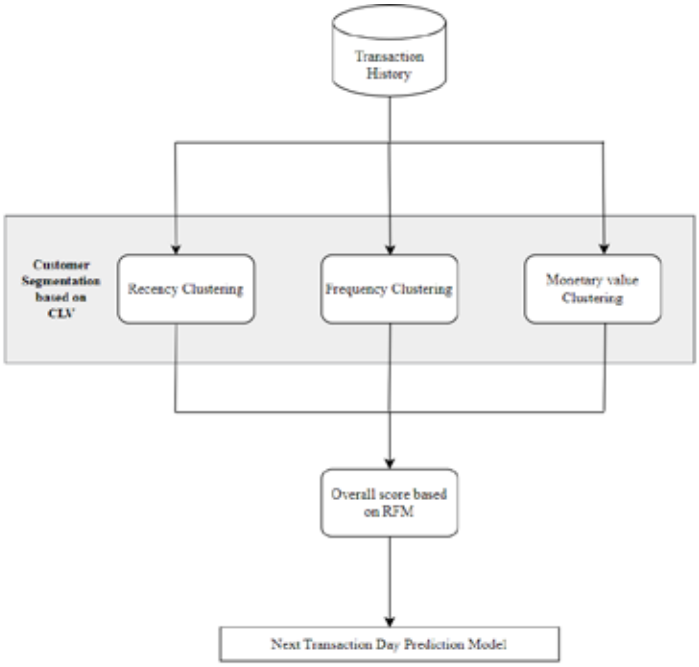


Figure 3: Proposed Next Transaction Day Prediction Framework

Figure 3 shows the novel architecture proposed by the author to predict the next transaction date from the dataset. The below steps were followed in order to create the new features to train the machine learning model,

- Using the K-means clustering method, the customers are segmented. The elbow method is used to decide the number of clusters.
- Getting the overall score by adding the values obtained by RFM clusters of each customer. Using this score, customer loyalty can be determined as shown in the below figure.

Overall Score	Segment/Loyalty
3-4	Low
5-6	Mid
7-8	High

Figure 4: Overall Score Segments

The overall score, which was computed based on the RFM clusters, has been divided into three categories to categorize the loyalty of customers as shown in Figure 4. Customers who have an overall score of 3-4 are classified as low loyal customers, while those with a score of 5-6 are classified as moderately loyal customers. Customers with a score of 7-8 are classified as highly loyal customers.

	MinCorrelationCoeff	MaxCorrelationCoeff
CustomerID	-0.042810	0.041278
NextTXNDay	-0.530915	0.451357
Recency	-0.964459	0.857385
RecencyCluster	-0.964459	0.968221
Frequency	-0.704095	0.476721
FrequencyCluster	-0.704095	0.271931
Amount	-0.377883	0.799774
AmountCluster	-0.360798	0.799774
OverallScore	-0.932676	0.968221
Segment_High	-0.081302	0.650685
Segment_Low	-0.990396	0.857385
Segment_Mid	-0.990396	0.879445
NextTXNDayRange	-0.530915	0.484624

Figure 5: Final Features with Correlations

In order to predict the specific next transaction date, a new feature was created to represent the range of the next transaction day. This approach differs from existing studies [11][12] that only focused on predicting whether a customer would perform a transaction on a given day. The next transaction days were categorized into ranges and assigned to a column called "NextTXNDayRange" as shown in Figure 5, which provides the next transaction day range instead of just predicting a binary value of whether the customer performs a transaction or not.

For this study, the transaction day has been categorized into three ranges. This allows for a more accurate prediction of the next transaction day and provides valuable insights for businesses to better plan their inventory and reduce logistics costs.

3.3 Machine Learning Approach

After deriving all the features shown in Figure 5, several machine learning models which are suitable for multiclass classification were trained and tested with the objective of minimizing the absolute error between the predicted and actual repurchase dates. This section will briefly review the techniques used in the study.

- Random Forest: It is a type of ensemble learning method that combines several decision trees to make predictions. It is known for its high accuracy and ability to handle large datasets [18].

- Gradient Boosting: This method also uses an ensemble of decision trees, but in a sequential manner, with each tree correcting the errors of the previous one. It is known for its ability to handle noisy data and high-dimensional feature spaces [19].
- Support Vector Machines (SVM): SVM is a linear or nonlinear classification algorithm that finds the best boundary (hyperplane) to separate the different classes [20].
- Neural Networks: These models are highly flexible and can handle complex relationships between features. They consist of layers of interconnected nodes that process the input data and produce an output[21].

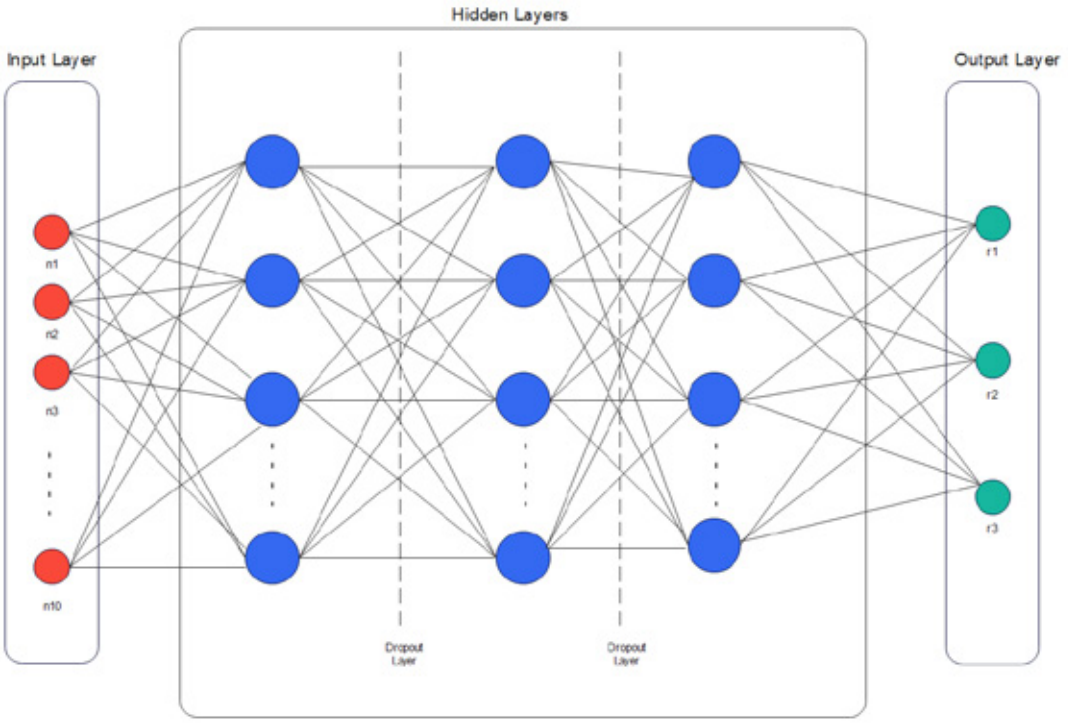


Figure 6: Proposed Neural Network Architecture

The proposed neural network is designed using the Sequential class in Python Keras to create a linear stack of layers. The architecture consists of three fully connected layers with 64, 32, and 16 neurons, respectively. The input to the neural network is a 10-dimensional vector. To address the issue of overfitting, Dropout layers are introduced. Dropout randomly deactivates a fraction of neurons during training, which helps prevent the model from relying too heavily on specific neurons and encourages better generalization to unseen data.

The final layer is a Dense layer with 3 neurons, and the sigmoid activation function is applied. Since this is a multi-class classification problem, the sigmoid function maps the output to a probability distribution across the three classes [21]. This means that the model predicts the probabilities for each class, and the class with the highest probability is considered the predicted class.

To train the model, the categorical_crossentropy loss function is utilized. This loss function is commonly used for multi-class classification tasks and aims to minimize the difference between the predicted probability distribution and the true labels of the data [22].

```
RandomForestClassifier [0.84101599 0.83623529]
SVC [0.81091251 0.81835294]
DecisionTreeClassifier [0.80997178 0.82635294]
xgb.XGBClassifier [0.83678269 0.83952941]
```

Figure 7: Accuracies taken after cross-validating the models

```
eval_model = model.evaluate(X_train, y_train)
eval_model
```

```
133/133 [==============================] - 0s 1ms/step - loss: 0.2710 - accuracy: 0.8579
[0.2710111141204834, 0.8579157590866089]
```

Figure 8: Accuracies taken after evaluating the Neural Network

The models were cross-evaluated, and their accuracies were obtained as shown in Figure 7. The RandomForrestClassifier achieved an accuracy of 84%, while the XGBoostClassifier achieved an accuracy of 83%. Subsequently, the normalized dataset was used to train a neural network, which achieved an accuracy of 85%, as shown in Figure 8.

4 Evaluation

When evaluating the performance of a learning algorithm in classification tasks, it is crucial to use appropriate metrics. Accuracy and error rate are the most commonly used metrics, but they tend to favor the majority class in imbalanced datasets. Therefore, these metrics do not adequately reflect the classifier's ability to predict examples from minority classes. To address this issue, alternative criteria have been proposed, which can be broadly categorized into single-class and overall performance measures [23]. The former assesses the classifier's performance in one class, while the latter provides an intuitive interpretation of its performance across all classes.

A confusion matrix can be developed to represent the classification problem and determine the classification performance, just like in binary classification problems. The performance metrics accuracy, precision, and recall may be generated from the confusion matrix and are the foundation for the performance measures of interest, F1 measure and G-mean. The Jaccard similarity index between the predicted labels and the actual labels is used to calculate accuracy. While recall is a measure of completeness that reflects the proportion of relevant items that are recommended, precision is a measure of exactness that describes the proportion of predicted items that are relevant [23].

		Predicted
Actual	True Positive	False Negative
	False Positive	True Negative

Figure 9: Confusion Matrix

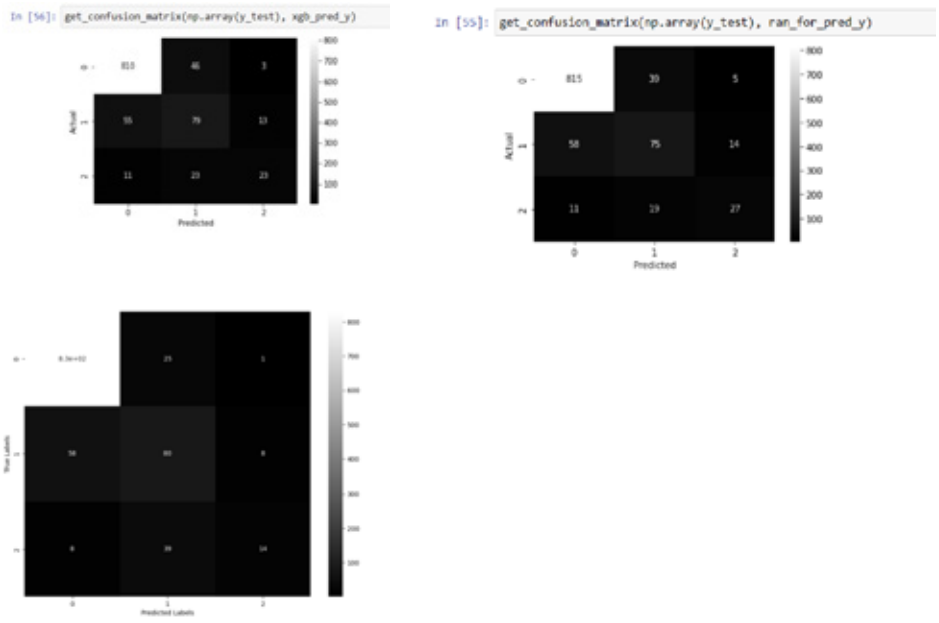


Figure 10: Confusion Matrix of the results predicted by the RandomForest Classifier, XGBoost model and the Custom Neural Network

The classifier performance in imbalanced data scenarios can be assessed using various evaluation metrics. However, Precision, Recall, F-measure, AUC, G-mean, Kappa, and MCC are the most commonly used metrics in both single-class and overall performance measures. These metrics are calculated from the confusion matrix, which is a two-dimensional matrix that provides an insightful representation of how the classifier performs in predicting new samples [22].

In Figure 8, the confusion matrices obtained for the most accurate models are presented.

	precision	recall	f1-score	support
0	0.92	0.95	0.94	859
1	0.56	0.51	0.54	147
2	0.59	0.47	0.52	57
accuracy			0.86	1063
macro avg	0.69	0.64	0.67	1063
weighted avg	0.85	0.86	0.86	1063

	precision	recall	f1-score	support
0	0.97	0.90	0.93	862
1	0.49	0.70	0.57	137
2	0.59	0.64	0.61	64
accuracy			0.86	1063
macro avg	0.68	0.75	0.71	1063
weighted avg	0.89	0.86	0.87	1063

	precision	recall	f1-score	support
0	0.92	0.94	0.93	859
1	0.53	0.54	0.54	147
2	0.59	0.40	0.48	57
accuracy			0.86	1063
macro avg	0.68	0.63	0.65	1063
weighted avg	0.85	0.86	0.85	1063

Figure 11: Classification report of the results predicted by the RandomForest Classifier, XGBoost model and the customNeural Network

The classification reports for the prediction results of RandomForrestClassifier, XGBoostClassifier, and the custom neural network are presented in Figure 11. The reports indicate that the neural network outperforms

the other models based on the macro average [24], which is a metric that can accurately measure class accuracy in an imbalanced dataset. Therefore this neural network architecture is well-suited for multi-class classification problems and can provide accurate predictions for the given classes.

5 Conclusion

The study has shown that it is possible to predict the next transaction day of a customer using a framework that combines unsupervised and supervised learning models. The framework achieves a satisfactory level of prediction accuracy. After evaluating various machine learning approaches, it was found that a fine-tuned multiclass neural network can provide accurate results for predicting customer transaction days. The study emphasizes the importance of considering customer lifetime value when predicting the next transaction day. This approach not only helps businesses forecast future transactions but also improves customer retention.

The proposed framework can be further improved by experimenting with more recent real-world transaction datasets, enabling it to learn the latest customer behaviors and provide better predictions. Additionally, it is important to validate the algorithm using real-time transaction systems to ensure its compatibility with marketing models and strategies. The next transaction day predictor has potential benefits for the environment, such as reducing the number of paper inserts in newspapers, improving inventory planning, and lowering logistics costs. Additionally, it can help businesses provide more personalized services to customers, ultimately improving customer satisfaction and retention.

ACKNOWLEDGMENTS

The author of this research paper acknowledges the guidance and evaluation insights received for the project by Mr. Udayanga Rathnasekara and the team. Additionally, the author extends sincere thanks to the supervisor for their continuous support and assistance during the research.

REFERENCES

[1] Epsilon Marketing, “The power of me: The impact of personalization on marketing performan…,” Jan. 04, 2018. https://www.slideshare.net/EpsilonMktg/the-power-of-me-the-impact-of-personalization-on-marketing-performance/1

[2] Y. Sun, D. Cheng, S. Bandyopadhyay, and W. Xue, “Profitable Retail Customer Identification Based on a Combined Prediction Strategy of Customer Lifetime Value,” Midwest Social Sciences Journal, vol. 24, no. 1, 2021, doi: https://doi.org/10.22543/0796.241.1053.

[3] E.-M. Nyckel, “Ahead of Time: The Infrastructure of Amazon’s Anticipatory Shipping Method,” JSTOR, 2021. https://www.jstor.org/stable/j.ctv1xcxr3n.18 (accessed Apr. 17, 2023).

[4] P. Kotler and G. M. Armstrong, Principles of Marketing. Prentice Hall, 2010. Accessed: Apr. 17, 2023. [Online]. Available: https://books.google.lk/books?hl=en&lr=&id=cRsxku7O06UC&oi=fnd&pg=PR13&dq=Kotler

[5] R. T. Rust, K. N. Lemon, and V. A. Zeithaml, “Return on Marketing: Using Customer Equity to Focus Marketing Strategy,” Journal of Marketing, vol. 68, no. 1, pp. 109–127, Jan. 2004, doi: https://doi.org/10.1509/jmkg.68.1.109.24030.

[6] S. Gupta, D. R. Lehmann, and J. A. Stuart, “Valuing Customers,” Journal of Marketing Research, vol. 41, no. 1, pp. 7–18, Feb. 2004, doi: https://doi.org/10.1509/jmkr.41.1.7.25084.

[7] D. P. Jaiswal, S. Kumar, and P. Mukherjee, “Customer Transaction Prediction System,” Procedia Computer Science, vol. 168, pp. 49–56, 2020, doi: https://doi.org/10.1016/j.procs.2020.02.256.

[8] R. Esmeli, M. Bader-El-Den, and H. Abdullahi, "Towards early purchase intention prediction in online session based retailing systems," Electronic Markets, Dec. 2020, doi: https://doi.org/10.1007/s12525-020-00448-x.
[9] "Predicting customer lifetime value -e-commerce use case." Available: https://arxiv.org/ftp/arxiv/papers/2102/2102.05771.pdf
[10] S. Monalisa, P. Nadya, and R. Novita, "Analysis for Customer Lifetime Value Categorization with RFM Model," Procedia Computer Science, vol. 161, pp. 834–840, 2019, doi: https://doi.org/10.1016/j.procs.2019.11.190.
[11] M. Droomer and J. Bekker, "USING MACHINE LEARNING TO PREDICT THE NEXT PURCHASE DATE FOR AN INDIVIDUAL RETAIL CUSTOMER," South African Journal of Industrial Engineering, vol. 31, no. 3, Nov. 2020, doi: https://doi.org/10.7166/31-3-2419.
[12] T.-C. Lu and K.-Y. Wu, "A transaction pattern analysis system based on neural network," Expert Systems with Applications, vol. 36, no. 3, pp. 6091–6099, Apr. 2009, doi: https://doi.org/10.1016/j.eswa.2008.07.073.
[13] A. Martínez, C. Schmuck, S. Pereverzyev, C. Pirker, and M. Haltmeier, "A machine learning framework for customer purchase prediction in the non-contractual setting," European Journal of Operational Research, vol. 281, no. 3, pp. 588–596, Mar. 2020, doi: https://doi.org/10.1016/j.ejor.2018.04.034.
[14] R. Sathya and A. Abraham, "Comparison of Supervised and Unsupervised Learning Algorithms for Pattern Classification," International Journal of Advanced Research in Artificial Intelligence, vol. 2, no. 2, 2013, doi: https://doi.org/10.14569/ijarai.2013.020206.
[15] D. Chen, S. L. Sain, and K. Guo, "Data mining for the online retail industry: A case study of RFM model-based customer segmentation using data mining," Journal of Database Marketing & Customer Strategy Management, vol. 19, no. 3, pp. 197–208, Aug. 2012, doi: https://doi.org/10.1057/dbm.2012.17.
[16] F. Safari, N. Safari, and G. A. Montazer, "Customer lifetime value determination based on RFM model," Marketing Intelligence & Planning, vol. 34, no. 4, pp. 446–461, Jun. 2016, doi: https://doi.org/10.1108/mip-03-2015-0060.
[17] M. Khajvand, K. Zolfaghar, S. Ashoori, and S. Alizadeh, "Estimating customer lifetime value based on RFM analysis of customer purchase behavior: Case study," Procedia Computer Science, vol. 3, pp. 57–63, 2011, doi: https://doi.org/10.1016/j.procs.2010.12.011.
[18] A. Chaudhary, S. Kolhe, and R. Kamal, "An improved random forest classifier for multi-class classification," Information Processing in Agriculture, vol. 3, no. 4, pp. 215–222, Dec. 2016, doi: https://doi.org/10.1016/j.inpa.2016.08.002.
[19] M. Rapp, E. L. Mencía, J. Fürnkranz, V.-L. Nguyen, and E. Hüllermeier, "Learning Gradient Boosted Multi-label Classification Rules," arXiv:2006.13346 [cs, stat], Jun. 2020, Available: https://arxiv.org/abs/2006.13346
[20] W. Aoudi and A. M. Barbar, "Support vector machines: A distance-based approach to multi-class classification," IEEE Xplore, Nov. 01, 2016. https://ieeexplore.ieee.org/document/7777430 (accessed Apr. 10, 2023).
[21] G. Ou and Y. L. Murphey, "Multi-class pattern classification using neural networks," Pattern Recognition, vol. 40, no. 1, pp. 4–18, Jan. 2007, doi: https://doi.org/10.1016/j.patcog.2006.04.041.
[22] Y. Ho and S. Wookey, "The Real-World-Weight Cross-Entropy Loss Function: Modeling the Costs of Mislabeling," IEEE Access, vol. 8, pp. 4806–4813, 2020, doi: https://doi.org/10.1109/access.2019.2962617.
[23] J. Tanha, Y. Abdi, N. Samadi, N. Razzaghi, and M. Asadpour, "Boosting methods for multi-class imbalanced data classification: an experimental review," Journal of Big Data, vol. 7, no. 1, Sep. 2020, doi: https://doi.org/10.1186/s40537-020-00349-y.
[24] X. Deng, Q. Liu, Y. Deng, and S. Mahadevan, "An improved method to construct basic probability assignment based on the confusion matrix for classification problem," Information Sciences, vol. 340–341, pp. 250–261, May 2016, doi: https://doi.org/10.1016/j.ins.2016.01.033.
[25] M. Grandini, E. Bagli, and G. Visani, "Metrics for Multi-Class Classification: an Overview," arXiv:2008.05756 [cs, stat], Aug. 2020, Available: https://arxiv.org/abs/2008.05756

Empowering Early Childhood Education Students: A User-Centered Approach to Enhance Spelling Skills

SHIDN De Silva. Author[1]

Department of Information Technology, Faculty of Computing, General Sir John Kotelawala Defence University, Ratmalana, Sri Lanka, 36-it-0010@kdu.ac.lk

MTA Wickramasinghe. Author[2]

Department of Information Technology, Faculty of Computing, General Sir John Kotelawala Defence University, Ratmalana, Sri Lanka, thisaranianandi@gmail.com

D Gunasekara. Author[3]

Department of Information Technology, Faculty of Computing, General Sir John Kotelawala Defence University, Ratmalana, dgunasekara@kdu.ac.lk

ABSTRACT

This research is to create an augmented reality (AR) spelling application for use in early childhood education. It can be difficult for parents and teachers to teach children spelling skills because the conventional approaches frequently do not capture their interest. This article provides AR technology as a method to create an interactive and immersive experience that can increase children's recognition and recall of individual letters within words. The main objectives of this research are to identify obstacles to language learning in early childhood education, develop a mobile application using augmented reality technology to improve spelling skills, and compare the efficiency of this application and traditional teaching methods in improving children's spelling. The study uses both qualitative and quantitative data and explores the application of a mixed-methods methodology. Parents and preschool teachers from Sri Lanka who take part in surveys and interviews make up the research population. According to the findings of the requirement analysis, youngsters are not sufficiently engaged by traditional teaching methods, and new and inventive methods are required. The study also describes the hardware and software prerequisites for creating AR applications, including ARCore, the Sceneform SDK, the Android SDK, and physically based rendering (PBR). In the suggested solution architecture, virtual 3D objects and the 3D letters of the provided word are displayed using augmented reality (AR) technology. Convolutional Neural Network (CNN) has been used to recognize the 3D letters presented in an AR environment. A trained recurrent neural network (RNN) model has been used to forecast the pronunciation of a letter in the context of the job of tapping on 3D letters. Ultimately, this study highlights how augmented reality technology has the potential to increase educational chances for children by offering a practical remedy for improving language and spelling learning.

KEYWORDS

Augmented Reality, Childhood Education, Mobile Application

1 Introduction

The development of technology has significantly altered the way we learn, live, and work. Education is one of the fields where technology has had a significant impact. Recently, there has been a rise in interest in using technology to improve education, especially in the early years. While teaching spelling to words using traditional methods, teachers and parents encounter several issues and difficulties. This can lead to a lack of interest in learning and can make it difficult for children to retain the information. Teachers and parents may have limited time to devote to teaching spelling skills. Children may find it challenging to remember the facts as a result, which may result in a lack of enthusiasm for learning. Parents and teachers may be limited in time when it comes to spelling instruction. Children may also have various learning styles and varied speeds of learning. There is an issue that traditional teaching strategies don't meet every child's needs specifically.

And teaching children how to pronounce words correctly can be difficult, especially if the parent or instructor is not a native speaker of the language. This innovative solution will address the problem by giving an effective solution for enhancing spelling skills and language learning in young children. One of the newest technologies augmented reality (AR), has the potential to significantly improve educational opportunities for children. AR is a technology that creates an interactive and immersive experience by superimposing digital content onto the actual world. Applications for augmented reality have been created for a variety of academic subjects, including physics, math, and spelling. This research study focuses on the development of an AR-based spelling application for 3-5-year-old children. The application's potential to improve children' ability in recognition and memorization of particular letters inside words is specifically explored. Additionally, aim to investigate how the application can improve children's language learning and pronunciation skills. The main objective of the research is to determine any difficulties or restrictions associated with language learning in early childhood education. Then develop a mobile application using Augmented Reality technology to enhance spelling skills for early childhood education. Finally aims to compare the effectiveness of this application and conventional teaching techniques in terms of increasing children's spelling skills.

2 Literature Review

A comprehensive review of the relevant literature on early childhood education, spelling skills development, and mobile learning technology will be conducted. This will inform the development of this application and the research questions for the study.

2.1 Related Research

Preschool children aged 3 to 5 develop their learning abilities by engaging with their environment. As they observe adults utilizing computers, they naturally desire to do so as well. However, at this tender age, they may not possess the necessary motor and cognitive skills to effectively operate computer systems.[1]

In recent years reliance on computer systems has increased in the fields of entertainment and e-learning. This revolution has made many schools take advantage of computers in early year's education programs.[1] There are many current children' language learning applications available in the market. Each of these systems has its pros and cons, making it difficult to determine which one is the best option. This review discussed some of the most popular language-learning applications for preschoolers, including their strengths and weaknesses. Some of these systems may use innovative technology such as augmented reality, gamification, and voice recognition to make language learning more interactive and engaging. However, these systems may also have limitations in terms of their accuracy, cost, and user experience. Ultimately, the choice of a language learning application will depend on the individual needs and preferences of each child and their parents.

The usefulness of games in enhancing children's learning has long been a subject of interest for researchers in this field. One study proposed the creation of an online multiplayer game as a means of delivering educational content to children. This research [2] postulated that the feedback, challenge, and enjoyable experience provided by the game would improve the cognitive and learning capabilities of children. The study further suggested that these games rely on recall and recognition principles and increase motivation for learning, cooperation, and adaptation. The integration of mobile technology into language education is happening at an exponential pace, providing students with the flexibility to utilize mobile-assisted language learning (MALL) without limitations of time or location. Devices such as cell phones, iPods, tablets, laptops, and iPads are utilized in mobile learning to support language acquisition. There are numerous applications specifically designed for individuals learning English as a second language.[5] As these applications aim to improve the language proficiency of learners, a major focus should be placed on developing their listening, speaking, reading, and writing skills through the use of mobile technology. [5]

The research on "EasyMath" application development aims to investigate the cognitive and motor development aspects of preschool children to design an edutainment system that is not only simple and engaging but will also provide them with useful learning content. "EasyMath" application, includes simple interfaces, includes only essential materials, and the ability to give appropriate feedback.

Moreover, from this application children can take little problem-solving experience, handle pointing devices, and use graphical icons. [1] This study [4], focuses on the development of an app aimed at teaching the basics of English, Spanish, and German to Arabic-speaking children. The app provides instruction on the alphabet, basic words, and phrases, and uses voice and image recognition to create an engaging and productive learning experience. The application also includes a simple quiz to assess the children's understanding of the material they have learned. The app may only focus on a limited number of words or phrases, leading to limited exposure to the target language. In line with this study [6], the "StoryKit" app is a highly effective mobile application designed for children between the ages of 4 and 8. It provides students with a platform to create and share multi-modal stories. The app combines text, illustrations, photographs, and sound effects to make storytelling an interactive and engaging experience. According to the source cited [6], "StoryKit" is highly rated among users, making it a popular choice for students looking to bring their stories to life. But the limitations were not improved children' pronunciation skills and interactive content. In accordance with this study, the "iWrite Words" app [8] is a great app for writing which allows students to practice drawing and writing words by using their fingers. Students can write letters, numbers, and sentences. But this application focuses on improving children' pronunciation skills and memorization abilities. In this research's conclusion [5] development of "Children Learn to Read" is an android application that is built for android to practice pronunciation skills. But this is built for only android devices. The limitations were no use of memorization techniques and lack of interaction. This research [7] indicates "Starfall App" and its website offer a variety of activities that aid students in reading for Kindergarten and elementary school. The app and website aim to motivate students to learn the sounds of letters by demonstrating pronunciation and encouraging students to learn through exploration and inquiry. But this application doesn't evaluate the student's progression of language development. Then currently explored the most effective preschool language learning apps available in 2023. Per this study's findings, "Duolingo" [9] is an application that provides game-based lessons in over 35 different languages, aimed at helping students practice speaking, listening, reading, and writing. It goes beyond just teaching vocabulary, by also offering written explanations of important grammatical concepts, helping mature learners gain a comprehensive understanding of the target language. The app features a news feed where adolescent learners can learn about additional ways to support language learning, including through TikToks. Parents will appreciate the daily reminders which encourage consistent language practice. Although Duolingo is marketed for children aged 4 and above, users need to have a solid foundation in reading and writing to maximize the benefits of the lessons. [9]

LingoDeer [10] is a user-friendly language-learning app that offers a customized experience. It allows learners to select their preferred narrators, script style, and dashboard color. The app features flashcards for memory retention and includes interactive stories to practice listening and speaking skills. However, it has some limitations such as limited language options, repetitive content, and a lack of in-depth courses. It is not available on a desktop and some courses may not have enough material to thoroughly learn a language. Additionally, there is a limited opportunity for speaking practice.

This research indicates "Little Chatterbox for Children" [11] Children can explore and learn a variety of languages with Little Chatterbox, from Spanish and French to German and Mandarin. Through the use of vibrant, appealing visuals and videos of native speakers pronouncing the words to teach vocabulary to young children. There are nine languages available on the app. It is one of the best language applications for children to aid with vocabulary acquisition in both native and foreign languages, even though we prefer apps that concentrate on interactive features and activities. But these applications are more expensive and also the interactivity of these applications is less.[11] Moving on to technology usage Augmented reality has been integrated into the development of children learning apps to provide an immersive and interactive experience for children, enhancing their engagement and motivation in learning.

A review of Augmented Reality (AR) technology as a teaching aid for foreign language instruction is provided in this work [13]. The analysis revealed that AR technology has several advantages for language learning, but it is not yet prepared for full incorporation into language programs. The report also offers recommended applications and concrete ideas for AR-enhanced exercises in four language skills. For teachers, teacher educators, researchers, and those who create course materials, this review has a number of consequences.

A general study [12] on augmented reality base language learning apps shows that this study aimed to enhance the educational experience and improve the children's understanding of vowel usage and numbers at "Juana Alarco de Dammert" nursery school in Trujillo, using a mobile app with augmented reality and markers in 2017. The specific goals were to boost the academic performance of vowel usage and numbers through the augmented reality app, assess the impact of the app on academic performance, and determine the average usage time of the augmented reality app based on markers. But the accuracy of those applications is at a low level and children' interaction and learning progress were low according to the evaluations.
In conclusion, the current preschool language learning applications still have room for improvement in terms of their level of engagement and effectiveness for young learners. The integration of technology such as Augmented Reality could enhance the learning experience by providing interactive and immersive environments for children to practice language skills. Additionally, incorporating virtual objects into the learning process could further promote children' interaction and engagement in language learning. It is crucial for these applications to continue to evolve in order to meet the needs of young learners and provide them with the best possible learning experiences.

3 Methodology

This research project employs a mixed-methods methodology and uses both qualitative and quantitative data. Parents and preschool teachers from Sri Lanka will make up the fifty-sample population, which will be encouraged the take an active part in surveys and interviews.

3.1 Requirement Analysis

Researchers have identified two key areas that need concentrated attention through in-depth interviews and surveys: the present teaching tactics used to assist youngsters to recognize and memorize specific letters within a word, and the child's desire for fun and interesting learning activities. The population's responses offer important information about the efficacy of the current teaching techniques and the need for original perspectives that can better cater to the interests and preferences of the kid.

[1] What strategies have you found helpful when teaching children to recognize and memorize individual letters within a word?

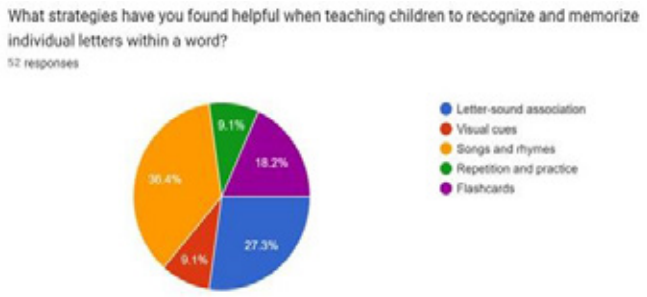


Figure 1: Responses for question 01

The question probably referred to efficient strategies for helping children identify and memorize the individual letters in a word. Their responses to various teaching methods, such as flashcards, visual cues, songs and rhythms, repetition and practice, and letter-sound association. Considering the results, the response then states that these tactics were applied equally. To ensure that the application is helpful in teaching children to recognize and remember specific letters within a word, all of these features should be included when developing it.

[2] Do you find that your child is engaged and happy when participating in the activities used to teach letters within a word?

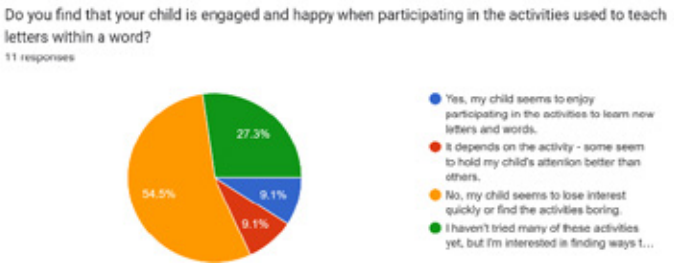


Figure 2: Responses for question 02

According to the responses gathered, it seems that most children are not engaged in the conventional activities used to teach letters inside a word. Concerns about children quickly losing interest and being bored have been raised by teachers and parents, making it difficult to keep them interested in their lessons. It seems that just a small percentage of children prefer conventional approaches. In order to engage young minds in the learning process and keep their attention, it is necessary to use a creative and unique approach.

3.2 Software and Hardware

Certain hardware and software requirements are necessary for the development of an application. The Android SDK, which comprises SDK build tools, Android emulator and Android device, android platform tools, SDK platform, and Google APIs, is required to create the application. The Sceneform SDK, which enables the development of dynamic AR apps without requiring knowledge of OpenGL, is another essential tool. A physically based renderer, a high-level scene graph API, and an Android Studio plugin make up the Sceneform SDK. The application's scene graph is a data structure that specifies the connections between virtual objects using transformable nodes that anchor the 3D objects into the scene. Physically based rendering (PBR) is a rendering technique that makes sure that all 2D and 3D models in the display have accurate lighting, allowing for realistic representations of various surface types. ARCore, which uses object tracking, light estimation, and environmental awareness to project 3D models into the environment, enables Android handsets to access AR features without the need for extra sensors.

3.3 System Solution Architecture

The proposed solution application has been designed with a well-defined system architecture, which is depicted in the diagram below. The interrelationship between components and how together function as a group to accomplish their intended functionality is illustrated in the diagram. The procedure of the application was discussed below in detail.

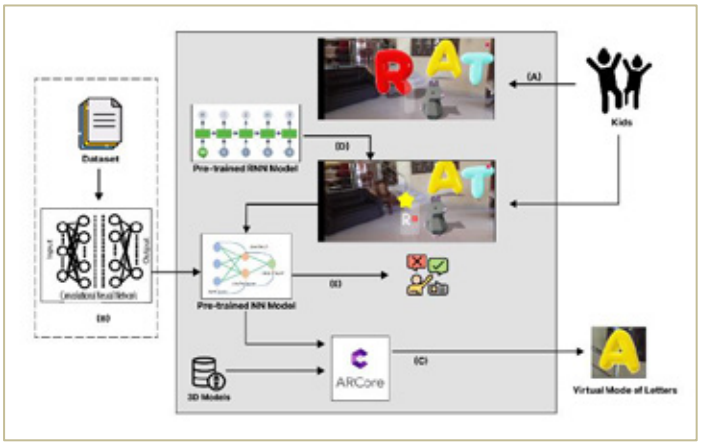


Figure 3: System Solution Architecture

This application displays the virtual 3D item that represents the first letter of a given word along with the 3D letters of the given word using augmented reality (AR) technology. The application utilizes augmented reality (AR) to overlay virtual objects with 3D letters of the supplied word and the 3D virtual item that represents that word on top of the real-world image that was acquired by the camera when the user points their device's camera at a surface.

For example, if the word is "CAT," the application shows "3D virtual CAT" displays on the floor and 3D letters "C" "A" and "T" floating in the air in front of the camera, visible through the device's screen. The child must next learn the proper letter placement for the word "CAT" before clicking on the letters in that order. When a child sees an object associated with a word, it will be easier for them to remember the word's letters. The app will then quickly indicate whether the chosen letters for the phrase are accurate or incorrect. This gives children a fun and engaging method to participate in the educational process while also assisting them with visualizing the letters in words.

Using Java, ARCore enables developers to create augmented reality applications for Android devices. For a comprehensive AR experience, ARCore tracks the physical environment and precisely positions virtual objects within it using computer vision techniques. For augmented reality apps, developers would create 3D models using Sceneform. They make significant letters and objects using 3D modeling software like Blender to make 3D models of them for this application. Developers used pre-built 3D animal items and alphabetical letters for this project. Once the 3D models have been located, they are exported inside a Sceneform-compatible format, such as glTF or OBJ. After that, the models are imported into the Sceneform Android Studio plugin, where they can be altered and placed in an augmented reality environment. The procedures to display the AR models through ARCore briefly can be discussed as follows.

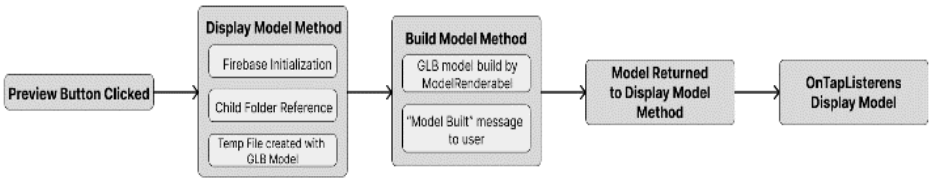


Figure 4 : Block diagram to display AR model

When Children turn on the ARMode button and click on the flashcard surface, users have been directed to the AR fragment page. Next, the display model technique is used, which involves initializing the Firebase database, attaching the 3D model to the scene graph's child node, and creating a temporary file to download the model file. The 3D model is created at runtime using the build model method, provided to the mechanism that displays models and correlated with such a node in the scene graph. When the on-tap listener is enabled, the model is displayed on the tapped node in the AR fragment, and a transformable node is employed to resize and shift the models horizontally over the scene graph.

The process of using Convolutional neural networks (CNN) to recognize the 3D letters displayed in an augmented reality environment involves several steps. A dataset comprising graphics of individual letters in different lighting and orientations is first gathered and categorized. To increase the generalizability of the model, the images are then preprocessed, which includes scaling, normalization, and data augmentation. The next model is the CNN, which has many layers of dense, pooling, and convolutional layers. The preprocessed image of the 3D letter presented in the augmented reality environment serves as the model's input. The most likely alphabetic letter that the input image represents is produced as an output prediction by the model after it has processed the image as input throughout its layers.

In order to reduce the error between the predicted output and the true label, the trained CNN model is employed, the preprocessed dataset is divided into sets for training and validation, and the trained model uses gradient descent and backpropagation. To make sure the model is not overfitting to the training data, the validation set is used to evaluate the model. The trained CNN model has been used to instantly identify the 3D letters demonstrated in the augmented reality environment after it has been validated.

The 3D letter input image is fed into the model, which then predicts the associated letter and displays it in the augmented reality environment for the kid to interact with.
CNN offers a number of advantages over neural networks (NNS) for the recognition of 3D letters shown in an augmented reality environment. The whole user experience is improved by CNN's capacity to extract spatial characteristics from 3D images to enable precise and effective letter recognition. CNN's architecture is specifically created to handle such information, ensuring accurate recognition of the letters in the augmented reality scenario, in contrast to NNS, which may struggle with the intricacies of 3D data. CNN is the go-to option for real-time and trustworthy letter identification in augmented reality applications due to its versatility and speed.
Then the process of tapping on a letter to hear its pronunciation to spell out words should be considered. Recurrent neural networks (RNN) have a number of benefits over classic neural networks (NNS), including the ability to process tapping on an AR letter to hear its sound when typing out words. Context awareness, made possible by RNN's sequential memory, offers more precise and contextually appropriate pronunciation feedback. The overall spelling experience in the augmented reality environment is improved by RNN's capacity to capture temporal dependencies in language acquisition, in contrast to NNS.

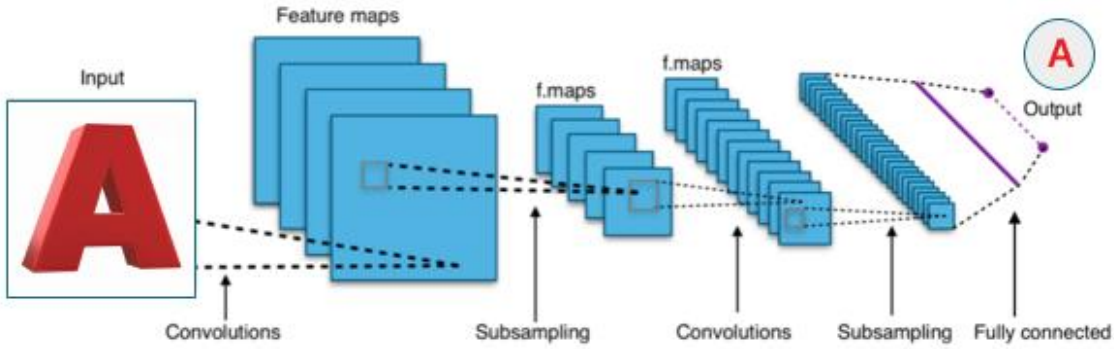


Figure 5 : The procedure of the CNN model

Recurrent Neural Networks (RNNs) have been effectively applied in the task of tapping on 3D letters and pronunciation to model the sequential nature of language and speech. RNNs are a class of neural networks that are designed to operate on sequential data by processing one input at a time and maintaining a "memory" of the previous inputs to inform the processing of subsequent inputs.
A trained RNN model has been utilized to predict the pronouncing of a letter or word given its sequence of taps in the context of the job of tapping on 3D letters and pronunciation. A set of taps that individually match a 3D letter in the augmented reality environment serve as the model's input. Each feature vector in the taps' encoded form indicates a different aspect of the tapped letter, such as its placement, size, and orientation. The trained RNN model processes each feature vector in the input sequence one at a time while changing its internal state. The model then generates an output, which is a forecast of how the tapped letter or word should be pronounced. A softmax layer is used to compute the output, which results in a probability distribution over a number of potential pronunciations. A labeled dataset of taps and their related pronunciations served as the model's training data, and its parameters were adjusted using methods like backpropagation via time. To provide feedback on the accuracy of the letters for a given word, the app likely uses pre-recorded audio to pronounce the word. This feedback can be used to let the user know if they have correctly spelled the word or if there are errors that need to be corrected.
When the app uses pre-recorded audio, it will contain a database of correct pronunciations for each word. When a user taps on the letters to form a word, the app compares the letter order to the word's proper spelling and plays the corresponding voice for the right pronunciation. The app may offer remedial feedback if the user spells the word incorrectly by playing a slow-talking audio clip or by requesting that they try again.

4 Implementation

Here it shows the application is deployed in sections as below.



Figure 6 : Selection of Animal (RAT)



Figure 7 : AR Preview

In Figure 6. It shows the list of animals to be selected for a kid to play the game. When the kid clicks on "RAT" and kid has to move toward to AR Mode of the selected animal. Then Figure 7. Shows the next phase that is archived by the kid. When the selected animal is a RAT, the Word is "RAT," the application shows "3D virtual RAT" displays on the floor and 3D letters "R" "A" and "T" floating in the air in front of the camera, visible through the device's screen.



Figure 8: Click on Letters



Figure 9: Make Letters in Order

After displaying all the letters and the object related to the word, here In Figure 8 shows the next phase. Children have the ability to memorize the order of letters to the word "RAT" and click on the letters in order. When the kid clicks on the letter, the kid could be able to hear the sound of pronouncing the letter loudly. If the letter clicked by the kid is in the correct order it displays in the environment as shown in Figure 3. The application will give a notification if the selected letter is not in the correct order to make the word "RAT". The kid has to continue finding the correct letters for words as shown in Figure 9. When the task is over children can take feedback on the Letters for whether the word is correct or wrong. Figure 10 and Figure 11 show the feedback notifications.



Figure 10 : Feedback for correction

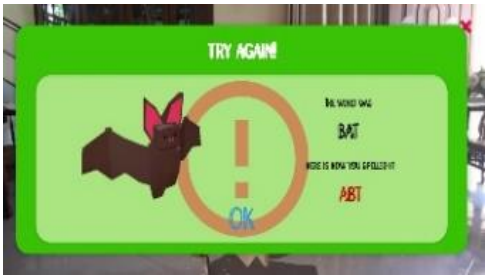


Figure 11 : Feedback for wrong

In this application, children have the ability to make the words without using AR mode also.



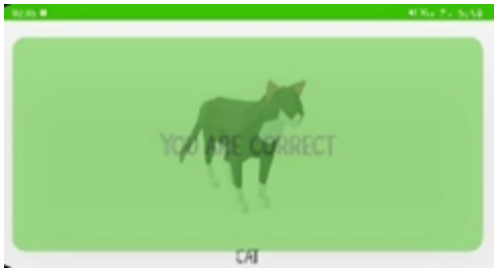


Figure 12: Making the words without using AR mode

Figure 13: Feedback.

Here in Figure 12, shows that the children could see the letters of the word "CAT" and the image of the "CAT". The kid can make the letters in order to make the correct word. Figure 13 shows that the feedback notifications are generated as the above function.

5 Results and Evaluation

In order to study teachers' and parents' cognitive perception and intention to use this application by children 15 statements on the 7-point Likert scale were listed and utilized for in-depth analysis. The top seven statements are mentioned here. Since "The child is engaged and happy when participating in the activities used to teach letters within a word." was the highest-ranked statement, it was deemed essential to ensure that further elaboration was heard from the teachers and parents.

Below Table 1 shows the results reveal that the app is useful for teaching and memorization of letters within words as well as the association of words with objects in early childhood. By using the application, which enables them to build words inside a real-time interactive environment and obtain feedback, children are interested and delighted. The capabilities of the app complement traditional learning methods. The application has the potential to serve as a useful tool for children who are developing literacy skills overall.

Table 1: Teachers' and parents' perceptions and intentions to use the educational game.

Rank Order	Student	Mean (M)	SD
1	The child is engaged and happy when participating in the activities used to teach letters within a word.	6.02	1.14
2	The use of this app enables the kid to memorize the word related to an object (animal).	5.64	1.54
3	The use of this app enables the kid to memorize the letters within a word.	5.01	1.67
4	The use of this app enables the kid to make the letters in order to make a word.	4.96	1.19
5	The use of this app enables giving feedback to the kid.	4.35	1.68
6	The use of this game makes a real-time interactive environment to make a word.	5.85	1.24
7	This app consists of features that tally with the traditional learning strategies used.	4.98	1.36

Through rigorous functional testing, we achieved over 90% accuracy in letter recognition and word association tasks in the augmented reality environment. The feedback delivery and real-time interactivity also scored above 90%, ensuring a seamless and engaging learning experience. These positive outcomes confirm the effectiveness of the proposed system in teaching children to recognize letters and memorize words, making it an excellent tool for language learning. These results validate the system's efficiency in delivering precise and reliable language learning experiences, highlighting its potential as a valuable tool for children's education.

6 Conclusion

In conclusion, the development of an augmented reality learning application provides a unique and interactive way for users to engage with educational content. Users can have interesting and immersive learning experiences thanks to technology like augmented reality, machine learning algorithms, and mobile device capabilities. The application's capacity to offer individualized feedback and advance linguistic abilities can be improved through the use of speech recognition and natural language processing algorithms. Additionally, the application's development process can be streamlined and made more efficient overall by using design patterns and development frameworks. Testing and evaluation are essential to guaranteeing any application's effectiveness and user happiness. To ensure that the application's performance and user experience satisfy the demands and expectations of its users, developers must constantly monitor and update the application.
Overall, the development of AR learning applications has the potential to revolutionize the way of approaching education and provide new and exciting opportunities for learners of all ages.

REFERENCES

[1] T. Yasir, "Child Computer Interaction: A Case of Preschool Edutainment Systems," 2018 International Conference on Smart Computing and Electronic Enterprise (ICSCEE), Jul. 2018, doi: 10.1109/icscee.2018.8538391.

[2] S. Y. Cheung and K. Y. Ng, "Application of the Educational Game to Enhance Student Learning," Frontiers in Education, vol. 6, Mar. 2021, doi: 10.3389/feduc.2021.623793.

[3] M. Kljun, A. Brodnik, and A. I. Starcic, "LMS in the pre-school education program," 28th International Conference on Information Technology Interfaces, 2006., Cavtat, Croatia, 2006, pp. 277-282, doi: 10.1109/ITI.2006.1708492.

[4] S. Harous, F. Alkhoori, A. Alzaabi, and S. A. Saedi, "Alturjoman Alsagheer: Basic language app for children," 2017 International Conference on Infocom Technologies and Unmanned Systems (Trends and Future Directions) (ICTUS), Dubai, United Arab Emirates, 2017, pp. 441-443, doi: 10.1109/ICTUS.2017.8286048.

[5] R. Gangaiamaran and M. Pasupathi, "Review on Use of Mobile Apps for Language Learning," International Journal of Applied Engineering Research, vol. 12, pp. 11242-11251, 2017, [Online]. Available: https://www.ripublication.com/ijaer17/ijaerv12n21_102.pdf

[6] A. Quinn, B. Bederson, E. Bonsignore, and A. Druin, "StoryKit: Designing a Mobile Application for Story Creation By Children And Older Adults," accessed Feb. 03, 2023, [Online]. Available: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=5e5b4fb0e6d524b45e0149f79cf895178f1021a1

[7] H. P. Parette and C. Blum, "Using Flexible Participation in Technology-Supported, Universally Designed Preschool Activities," TEACHING Exceptional Children, vol. 46, no. 3, pp. 60-67, Jan. 2014, doi: 10.1177/004005991404600307.

[8] H. Alkhamis, "EWU Digital Commons Implementing iPad apps for elementary language arts students," [Online]. Available: https://dc.ewu.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1187&context=theses

[9] M. Shortt, S. Tilak, I. Kuznetcova, B. Martens, and B. Akinkuolie, "Gamification in mobile-assisted language learning: a systematic review of Duolingo literature from public release of 2012 to early 2020," Computer Assisted Language Learning, pp. 1-38, Jul. 2021, doi: 10.1080/09588221.2021.1933540.

[10] Techjury, “LingoDeer Review - Can It Help You Master a New Language?,” [Online]. Available: https://techjury.net/reviews/lingodeer-review/#gref.

[11] E. Ltd, “10 Best Language Learning Apps for Children - Download Now!,” Educational App Store, [Online]. Available: https://www.educationalappstore.com/best-apps/10-best-language-learning-apps-for-children.
[12] E. Cieza and D. Lujan, “Educational Mobile Application of Augmented Reality Based on Markers to Improve the Learning of Vowel Usage and Numbers for Children of a Kindergarten in Trujillo,” Procedia Computer Science, vol. 130, pp. 352–358, 2018, doi: 10.1016/j.procs.2018.04.051.
[13] C. G. Karacan and K. Akoğlu, “Educational Augmented Reality Technology for Language Learning and Teaching: A Comprehensive Review,” Shanlax International Journal of Education, vol. 9, no. 2, pp. 68–79, Mar. 2021, doi: 10.34293/education.v9i2.3715.

Developing a System to Transliterate Singlish Twitter Posts to Sinhala

M. D. C. Amarasekara †
Department of Information Technology, University of Moratuwa, Sri Lanka, dilusha.18@itfac.mrt.ac.lk

R. A. D. P. Rajapaksha
Department of Information Technology, University of Moratuwa, Sri Lanka, dammika.18@itfac.mrt.ac.lk

H. M. D. T. Jayarathna
Department of Information Technology, University of Moratuwa, Sri Lanka, theekshana.18@itfac.mrt.ac.lk

H. M. G. K. Karunarathna
Department of Information Technology, University of Moratuwa, Sri Lanka, gajitha.18@itfac.mrt.ac.lk

I. T. S. Piyatilake
Department of Computational Mathematics, University of Moratuwa, Sri Lanka, thilinisp@uom.lk

C. P. Wijesiriwardana
Department of Information Technology, University of Moratuwa, Sri Lanka, chaman@uom.lk

† Corresponding Author

ABSTRACT

Due to the continued growth of Sinhala microblogs and the increasing popularity of Singlish language use on Twitter, it is important to have a mechanism to translate Singlish posts into Sinhala.This paper proposes a novel approach for transliterating Singlish text on the Twitter platform into the Sinhala language. The proposed approach uses a rule-based methodology that incorporates grammatical structures and patterns in the Sinhala language. N-gram analysis and corpus dataset technology are used to improve the accuracy of the transliteration process. Using these strategies, the system significantly improves its accuracy in capturing the intended meaning of Singlish posts in Sinhala. Experimental results demonstrate the effectiveness and potential of the proposed approach, highlighting its ability to accurately transliterate Singlish content into Sinhala on the Twitter platform, thus bridging the gap between the two languages.

Keywords: *Singlish, N-gram analysis, Rule-based approach,BLEU,GLEU*

I. INTRODUCTION

People from all cultures and origins are now more connected than ever due to the growth of social media platforms [1]. Due to the brevity and real-time nature of these platforms, users can share their opinions and interact with a large audience quickly. Twitter is one such platform which was introduced in 2006. The number of Twitter users has increased tremendously in recent years. Twitter's influence goes beyond simple communication; it has developed into a potent medium for people to express their identities, opinions, and distinctive linguistic idioms. Singlish is a sub-language of Sri Lanka. It is a distinct linguistic style, has arisen on Twitter and gained popularity among different communities around the country. It includes local slang, colloquialisms, and grammatical structures peculiar to its speakers. Sinhala is a diverse language with its own distinct features, alphabet, and cultural significance. The transcription of Singlish Twitter messages into Sinhala aims to overcome the linguistic gap and enhance cross-cultural understanding [2]. This transliteration projection is important for maintaining language diversity and encouraging inclusivity in the digital age. Twitter's global reach gives viewpoints that would otherwise be neglected or ignored to a platform. We ensure that individuals who communicate in Singlish can be understood and appreciated by a wider audience, including those who predominantly speak Sinhala, by facilitating the transliteration of Singlish Twitter postings to Sinhala.

The objective of this study is to discuss the obstacles and opportunities that arise throughout the transliteration process, taking into account the linguistic and cultural intricacies of both Singlish and Sinhala. Meantime we are developing transliteration system that capture the essence of the original message while keeping the integrity of both languages by researching the

linguistic traits, lexicon, and grammar of Singlish. We can use such kind of implementation to examines the transliteration of Singlish Twitter messages to Sinhala, emphasizing Twitter's value as a forum for expression and communication. This will facilitate to build an inclusive digital environment that embraces the variety of language and cultural identities by encouraging cross-cultural understanding, conserving linguistic diversity, and developing cultural relationships.

II. RELATED WORK

Macketanz et al. [3] presented a machine-learning approach to transliteration, using a combination of neural networks and rule-based techniques. They first described the challenges of transliteration, including the need to handle large volumes of data and the difficulty of preserving the meaning of the text when converting between scripts. They then describe their hybrid approach, which uses a neural network to learn how to map characters from one hand to another, and a set of rules to resolve conflicts or ambiguities in the transliteration process. Kay and Rineer [4] presented a method for transliterating personal names between different scripts, focusing on the challenges and limitations of such transliteration in the context of cross-language information retrieval. They described a rule-based approach to transliterating personal names, which combines phonetic and semantic information to map characters from one script to another. They evaluate the performance of their approach on a dataset of transliterated personal names between English and Arabic and show that it outperforms a number of other transliteration methods. Dhindsa et al. [5] proposed a transliteration system for English to Hindi transliteration that uses a conjunction-based approach. It consists of two phoneme-based modules. Module 1 was implemented using the phonetic dictionary method and module 2 was implemented using the rule-based (5-gram format) method. The proposed system's word accuracy is 70.22% compared to 58.73% of the Google input tool. Karunaratne et al. [6] suggested techniques for Singlish sentence prediction in SMSs. The proposed techniques are based on a Genetic algorithm that takes into account time series and user classifications. When selecting the best approach for prediction, several methods have been compared using the Genetic algorithms method. However, because the meaning of sentences should be generated by a sequence of separate words, this research is primarily focused on sentence prediction rather than word prediction.

Dias et al. [7] conducted research to answer the problem of matching or searching for proper names (people and place names). The goal of this study was to create a rule-based engine capable of matching Sinhala proper names even if they were spelled differently than the search query. For Sinhala proper names, an 85% success rate has been reached.Sonnadara et al. [8] presented a machine-learning approach to correct spelling errors in Sinhala. They used a neural network to learn how to detect and correct errors in the text. The authors evaluate the performance of their approach on a dataset of Sinhala text and show that it outperforms a rule-based approach. Wathsala et al. [9] created a Sinhala spell checker utilizing *n*-gram analysis. Three types of errors-pronunciation and spelling, reflexive and dental dissonance the "na-na-la-la" dissonance, and the pronunciation and spelling of reflexive and palatal sibilants-were discovered through this research. This method creates a list of distinct words by encoding the input text into distinctive Sinhala words. The corpus is then used to compare each word. The word is sent for additional processing if it is not found in the corpus but is removed from the list of unique terms if it is.

In our approach, we introduced both English and Sinhala corpus datasets to streamline the transliteration process and reduce the number of words requiring further processing. By incorporating these datasets, we aimed to enhance the efficiency of the system by quickly identifying and validating common words. Additionally, we introduced a spell checker module, leveraging machine learning techniques and Natural Language Processing (NLP) algorithms. This spell checker module further improved the accuracy of the system by detecting and correcting spelling errors in the transliterated text. Through the integration of corpus datasets and the implementation of advanced spell-checking techniques, we are optimistic about achieving higher accuracy levels in the transliteration process.

III. METHODOLOGY

The Sinhala language is often considered a challenging language due to its unique characteristics, one of which is the existence of words that have similar sounds but different meanings. When translating a specific language into Sinhala, one must be cautious and think carefully to ensure accurate and meaningful translations. To address this issue, we have developed an approach that incorporates a transliteration module and an error correction module. The transliteration module aids in converting words from one script to another, facilitating a smoother translation process. Additionally, the error correction module helps identify and rectify any potential mistakes or inaccuracies in the translation, ensuring a more reliable output. By combining these modules, we can overcome the difficulties posed by the complexity of the Sinhala language and provide a more effective solution for translation needs. Figure 1 shows the high-level diagram of the proposed approach.

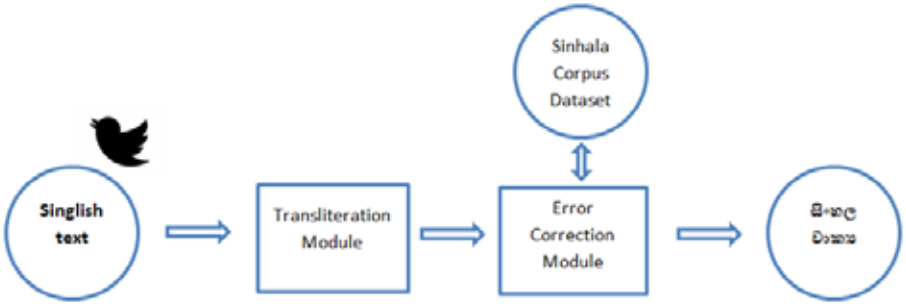


Figure 1: High-level diagram of the proposed approach.

To further enhance the accuracy level of our error correction module, we have incorporated a Sinhala corpus dataset. By adding this valuable resource to our system, we can leverage a vast collection of Sinhala language texts, which enables the module to better understand the language's grammar, vocabulary, and usage patterns. This integration allows the error correction module to make more informed and precise corrections, resulting in a significant improvement in accuracy. By continuously updating and refining the corpus dataset, we ensure that our error correction module stays up-to-date. Therefore, the system is capable of handling a wide range of Sinhala language nuances and intricacies.

A. Transliteration Module

For the development of our transliteration module, we have employed a rule-based approach. This approach involves creating a set of predefined rules that govern the conversion of words from one script to another, specifically targeting the transliteration from a particular language to Sinhala. Rule-based approaches are significant in transliteration tasks as they provide a systematic and structured framework for mapping sounds and characters between languages. These rules take into account phonetic and orthographic similarities, allowing for accurate and consistent transliterations [10]. The importance of a rule-based approach lies in its ability to handle various linguistic patterns and complexities present in languages, ensuring reliable and high-quality transliterations. By meticulously crafting and refining these rules, we strive to achieve precise and faithful representations of words and phrases in Sinhala, thereby facilitating effective communication across language barriers. Sinhala is a syllable-based language. Our first method was to break a word down into syllables and execute a replacement per syllable. Table 1 shows the suggested Sinhala transliteration mapping for Sinhala constants. The proposed Sinhala transliteration mapping for Sinhala vowels are shown in Table 2.

Table 1. Proposed Sinhala transliteration table for Sinhala constants.

Sinhala	English	Sinhala	English
ක	k,c	ම	m
ග	g	ය	y
ච	ch	ර	r
ජ	j	ල	l
ට	t	ව	w,v
ත	th	ස	s
ද	d	ශ	sh
න	n	හ	h
ප	p	ඤ	gn
බ	b		

We have introduced the following set of rules specifically designed to transliterate Singlish letters into Sinhala unicode characters.

a) Vowel modifier such as {◌aa } - /ruu/, {◌a} - /ru/.
For example, ‘k’ + ‘ruu’ = ‘ක’ + ‘◌aa ’ = ‘කaa’
‘Kruurathwaya’ = ‘කaaරත්වය ’

b) Constant with ‘Rakaransaya’ and the vowel modifier is replaced.
For example, ‘th’ + ‘r’ + ‘a’ = ‘ත’ + ‘්’+‘◌ා’= ‘ත්‍රා’
‘thrasaya’=’ත්‍රාසය’

c) Constant with other vowel modifiers is replaced.
For example, ‘k’ + ‘o’ = ‘ක’ + ’◌ො’ = ‘කො’
‘Kotuwa’ = ‘කොටුව’

d) Constant without vowel modifiers are replaced with 'Halkirima'(්ර).
For example, 'k' = 'ක' + '්' = 'ක්'
'Wakkada' = 'වක්කඩ'

e) Single vowels can be replaced.
For example, 'e' = 'එ'
'Eka' = 'එක'

Table 2. Proposed Sinhala transliteration table for Sinhala vowels.

Sinhala Vowel	Vowel Modifier	English
අ	-	a
ආ	ා	-
ඇ	ැ	-
ඈ	ෑ	-
ඉ	ි	i
ඊ	ී	-
උ	ු	u
ඌ	ූ	oo
එ	ෙ	e
ඒ	ේ	-
ඓ	ෛ	ai
ඔ	ො	o
ඕ	ෝ	-
ඖ	ෞ	au

B. Error Correction Module

The error correction module holds significant importance in ensuring accurate and precise translations. It comprises four sub-modules: the Sinhala corpus dataset, similar sound generation module, word distance calculation module, and best word detection module. The Sinhala corpus dataset module serves as the initial step in error correction, directly identifying words that exist within the dataset. These words are not further processed by the subsequent sub-modules. However, if a word is not found in the corpus dataset, it proceeds to the other sub-modules for further analysis. The similar sound generation module generates a list of words with similar sounds, offering potential alternatives for error correction. The word distance calculation module then measures the linguistic proximity between words, assisting in determining the most suitable correction based on context. Finally, the best word detection module considers various factors like frequency and contextual relevance to identify the most appropriate correction. By integrating these sub-modules, the error correction module ensures a robust and comprehensive approach to improving the accuracy and reliability of translations in the Sinhala language. Figure 2 shows the steps involved in error correction module.

The following sub modules are introduced to the proposed approach.

a) Sinhala corpus dataset: The Sinhala corpus dataset used in our system is compiled from a wide range of Sinhala news websites, capturing the most widely used words in the language
b) Similar sound generation module: The Similar sound generation module in our system employs a sophisticated algorithm to generate words with similar sounds by manipulating specific types of letters known as 'Mahaprana', 'Thaluja', and 'Murdhaja' letters. Proposed similar sounding letters with their categories are shown in Table 3.
c) Word distance calculation module: The word distance calculation module within our system plays a crucial role in assessing the linguistic proximity between a specific word obtained from the transliteration module and the output words generated by the similar sound generation module.
d) Best word detection module: The best word detection module in our system plays a crucial role in selecting the most appropriate word from the outputs of both the permutation module and the edit distance calculation module.

Table 3. Similar sounding letter groups.

Letter group	Sinhala letter list
Constants	{ක, ඛ}, {ග, ඝ, ඟ}, {ච, ඡ}, {ජ, ඣ}, {ට, ඨ}, {ඩ, ඪ, ඬ}, {ත, ථ}, {ද, ධ, ඳ, ඩ, ඪ, ඬ}, {ප, ඵ}, {බ, භ, ඹ}, {ම, ඹ}, {ල, ළ}, {ස, ශ, ෂ}, {ඥ, ඤ}.
Vowels	{අ, ආ, ඇ, ඈ}, {ඉ, ඊ}, {උ, ඌ}, {එ, ඒ}
Vowel modifier	{ැ, ෑ}, {ි, ී}, {ු, ූ}, {ෙ, ේ}, {ොෝ, ෞ}
Special characters	{ඞ්, ං}, {ය, ්‍ය(yanshaya)}, { ඉ, ඊ}

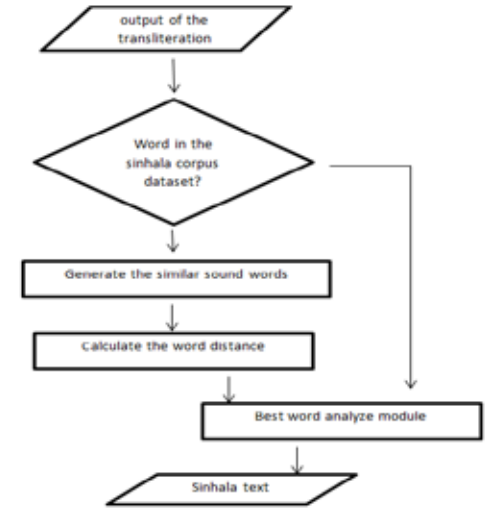


Figure 2: Flow chart of proposed solution for error correction module.

IV. RESULTS AND DISCUSSION

This study focused to improve the accuracy of transliterating Singlish posts from Twitter into Sinhala letters. The input dataset consisted of Singlish posts from Twitter, which underwent transliteration using a dedicated module. The transliteration process converted Singlish text into Sinhala letters. The resulting words were then checked against a Sinhala corpus dataset, consisting of commonly used Sinhala words. If a word matched the dataset, it was considered valid and no further processing was required. For non-matching words, subsequent modules were employed for error correction, including similar sound generation, word distance calculation, and best word detection. The multi-stage approach significantly enhanced accuracy by considering phonetic similarities, linguistic proximity, and context. Figure 3 shows the input singlish sentences and Figure 4 shows the generated output from our system.

Discuss about the strengths and weaknesses of the system

Similar-sounding words could be identified correctly in our system in the error correction part.

Ex:- 'wanaya' - වනය/වණය Our system can identify the best word.

Signlish language patterns are not able to identify our system.

Ex:- 'mk' - මක් (this is our system output). But this is the correct "මොකද කරන්නේ". Our system was not developed to identify Singlish patterns

we employ both BLEU(Bilingual Evaluation Understudy) and GLEU(Google-BLEU) metrics to assess the performance of our system. Using various test cases, we calculate the values of these metrics, ranging from 0 to 1. A higher value signifies a higher level of accuracy in the translation. This evaluation process allows us to evaluate the effectiveness and precision of our translation mechanism in transforming Singlish posts into Sinhala. Table 4 shows the different tests that were used in our system. Table 5 shows the output of the BLEU and GLEU values. These test cases output sentences compared with the man-written sentences using these metrics.

Table 4. Test cases used to evaluate proposed system

Tests	Scenarios
Test 1	Output of the transliteration module before adding corpus dataset
Test 2	Output of the transliteration module after adding corpus dataset.
Test 3	Output of the preprocessing module

Table 5. Output values of proposed system

Tests	BLEU	GLEU
Test 1	0.2989	0.0830
Test 2	0.3178	0.1130
Test 3	0.3367	0.1329

The results demonstrated improved transliteration accuracy, making the system applicable for language translation, sentiment analysis, and social media analysis tasks. Further research can focus on expanding the corpus dataset and addressing Singlish variations on social media platforms to enhance the system's performance.

V. CONCLUSION

The proposed approach successfully improved the accuracy of transliterating Singlish posts from Twitter into Sinhala letters. Incorporating a multi-stage approach, including a transliteration module, Sinhala corpus dataset, and error correction modules, significantly enhanced the system's performance. The system demonstrated high accuracy in capturing the intended

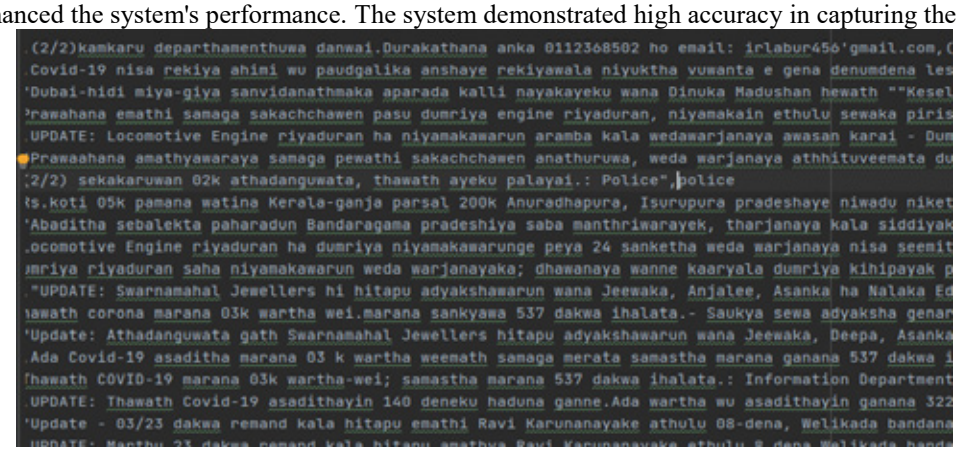


Figure 3: Singlish post (Input).



Figure 4: Output of our system.

meaning and accurately representing Singlish text in Sinhala. Furthermore, the researchers observed that the proposed system performed efficiently, requiring less time to process the transliteration task. This efficiency is attributed to the inclusion of the Sinhala corpus dataset, which allowed for the quick identification and retention of valid Sinhala words. The subsequent error correction modules further improved the accuracy of the system, providing alternative word options based on phonetic similarities, linguistic proximity, and contextual relevance. To further enhance the system, we can focus on expanding the corpus dataset. By incorporating a more comprehensive collection of Sinhala words, including domain-specific terminology and colloquial language, the system can better handle a wide range of transliteration scenarios.Moreover, the research highlighted the need for continuous refinement and optimization of the system to improve its overall performance. This can involve fine-tuning the transliteration rules and algorithms, incorporating additional linguistic rules and patterns, and exploring advanced techniques such as machine learning and neural networks.

Future works can be done as a method of increasing knowledge according to the changes in the Singlish language that take place from day to day. Can develop this system to identify the Singlish language patterns above mentioned. In addition, a Sinhala corpus dataset can be created that is updated in relation to a specific time frame.

REFERENCES

[1] X Zhang and A Ghorbani, "An overview of online fake news: Characterization,detection, and discussion," *Inf. Process. Manag.*, vol. 57, March 2020.

[2] K Shu, A Sliva, S Wang, J Tang, and H Liu, "Fake News Detection on Social Media:A Data Mining Perspective," vol. 19.

[3] V Macketanz, E Avramidis, A Burchardt, J Helcl, and A Srivastava, "Machine translation: Phrase-based, rule-based and neural approaches with linguistic evaluation," *Cybernetics and Information Technologies*, 2017.

[4] Brant N Kay and Brian C Rineer, "Approaches to Arabic Name Transliteration and Matching in the DataFlux Quality Knowledge Base," *SAS Institute Inc*, 2011.

[5] B K Dhindsa and D V Sharma, "ENGLISH TO HINDI TRANSLITERATION SYSTEM USING COMBINATION-BASED APPROACH," *International Journal of Advanced Research in Computer Science*, vol. 5, p. 8, October 2017.

[6] M S Karunarathne, "Sentence Prediction on SMS in Sinhala Language," *International Journal of Scientific and Research Publications*, vol. 3, p. 5, 2013.

[7] S C Fernando and G Dias. (2011) Inexact Matching of Proper Names in Sinhala. [Online]. http://dl.lib.mrt.ac.lk/handle/123/666

[8] Charana Sonnadara, Surangika Ranathunga, and Sanath Jayasena, "Sinhala Spell Correction ¬ A Novel Benchmark with Neural," University of Moratuwa, 2021.

[9] Eranga Jayalatharachchi , Asanka Wasala , and Ruvan Weerasinghe, "Data-Driven Spell Checking: The Synergy of Two Algorithms for Spelling Error Detection and Correction," University of Limerick, Limerick, 2015.

[10] V Macketanz, E Avramidis, A Burchardt, J Helcl, and A Srivastava, "Machine Translation: Phrase-Based, Rule-Based and Neural Approaches with Linguistic Evaluation," *Cybern. Inf. Technol.*, vol. 17, June 2017.

[11] Baljeet Kaur Dhindsa and Dharam Veer Sharma, *International Journal of Advanced Research in Computer Science*, vol. 5, p. 8, October 2017.

Helaa: A Sinhala Language-Based Programming Language

Rajitha Yasasri
University of Colombo School of Computing, University of Colombo, Colombo, Sri Lanka,
rajithayasssri@gmail.com

Damitha Karunarathna
University of Colombo School of Computing, University of Colombo, Colombo, Sri Lanka,
ddk@ucsc.cmb.ac.lk

ABSTRACT

Programming is an essential part of IT education. Most programming languages are mainly developed in English, which poses a language barrier to most people and limits the accessibility to learning how to code. This research proposes a novel programming language, called 'Helaa' which targets the Sinhala language. The proposed language is simple, imperative, and close to the Sinhala natural language. The language uses the Sinhala character set and employs commonly used Sinhala words as keywords. The compiler was developed in the Java programming language, but it is possible to develop it in other programming languages such as JavaScript and Python. Furthermore, the compiler construction is also discussed. The compiler is also able to give the exceptions in Sinhala language. Finally, a user survey and two user studies have been conducted to evaluate this work.

KEYWORDS

Non-English Programming Language, Sinhala, Compiler Design, Context-Free Grammar

1 Introduction

Starting from early low-level machine languages, programming languages have evolved around five generations and even to Quantum programming.

In this large number of programming languages, around a third of them are inspired by the English language. This does not count the programming languages that were created in a country where English is not the primary language. Such programming languages are Python, Ruby, and Lua. Almost all the mainstream programming languages use English keywords and are inspired by the English language. The reason for the English programming languages to become dominant is very simple. It's because computing technology was born and developed in English-speaking countries.

But with time, people identified the need to internationalize the programming languages and came out of the mainstream programming language design choice of English language use.

English is the lingua franca of academic works. But it is not the language with most native speakers. There are a lot of countries where a majority of their population is Non-Native English Speakers. To support their community, especially focusing on the educational aspect, and to address the need for more language research, the researchers came up with many non-English programming languages and related works. In this research work, the researchers are trying to achieve the same goal for the Sinhala language.

Sinhala is a language that belongs to the Indo-Aryan branch of the Indo-Iranian language family of the larger language family of the Indo-European language family. The language is primarily used in Sri Lanka which is a South Asian country. It is also one of the national and official languages used in Sri Lanka with more than 15 million native speakers. There are also a few million secondary speakers as well.

1.1 Research Objectives

The main objective of the research is to propose a design that is syntactically closer to the natural language, for a programming language inspired by the Sinhala language and develop a compiler for testing the design. So that this language could be used as an instrument to overcome the language barrier and broaden access for the students to technology.

1.2 Methodology

The main constructs of a programming language were first identified. And then the Sinhala codes were designed for these identified constructs. To develop the grammar, Liyanage et al.[7] grammar was expanded to cover these designs. The designs were tested and refined by generating the parse trees against the expanded grammar using a similar NLTK parser as in the [7]. Next, this initial grammar was developed into the programming language grammar by introducing rules for variable names, support for nesting, and more. At the same time, the compiler development was started, and added improvements to the programming language grammar. Lastly, the programming language was evaluated.

2 Non-English Programming Languages

There are many non-English programming languages developed and used throughout the world. Some of these languages are discussed in this section.

Phoenix [1] is an object-oriented language inspired by the Arabian language. The language uses a compiled approach so it can offer advantages like early error detection. The language is designed to be a general-purpose, high-level language so it can be used for a wide range of applications, ease of use, and abstractions. The language has limited portability since it is only developed for Windows. The language is similar to C# and it can be seen as an advantage for developers who are already familiar with C# like languages but this raises questions about the originality and innovation of the language.

Dolittle [2] is a very extensive and feature-rich non-English-based programming language for the Japanese language. It targets K12 Education (Kindergarten to grade 12) which is commendable. Because of this GUI, components, graphics, and animation capabilities are incorporated into the language making it feature-rich. The language follows a Prototype-based object-oriented model like Self and JavaScript languages. Complex components like timers, GUIs, image related codes can be developed with minimal complexity. Because the language is specialized in design, the range of its potential uses is unclear.

Wang [3] developed PseuToPy which acts as a natural language extension to Python programming language. The concept is intriguing. But this was created by editing the rules of Python grammar so it is still closely tied to Python language. The keywords of the

targeted languages can also be added. There should be specific rules implemented when editing the Python rules and adding keywords otherwise consistency would be a problem and may cause unintended consequences in the resulting language design. The developers have created the French version. A parser was developed using the lark-parser Python package. The main motivation for developing the language was to reduce the syntax errors caused by English keywords by non-native English speakers. However, there is no concrete evidence of whether they have achieved this goal or not.

Kamal et al. [4] have developed ChaScript which uses the Bengali character set and Bengali semantics which is a unique approach. They make an interesting distinction by referring to this work as a 'programming system' but not a programming language. The main motivation to develop this programming system is education. To help non-native English speakers, the language provides detailed information about parsing errors. The system was developed using ECMAScript. The authors claimed that it can be used to teach students by presenting the results of their experiments on 110 non-programmers using this system.

Srivastava et al. [5] developed a desktop application for the Hindi language with the objectives of helping beginners to start programming in Hindi and then for the smooth migration of Hindi to C, C++, or Java. Even though some control structures are different, they have used a template-based approach so the code is more similar to their counterpart equivalent codes in languages like C or C++ but that could create the learners an easier transition from this language to C or C++ or Java-like language. The application is portable because it is written in Java. An interesting feature of this application is that it translates the Java errors to Hindi, though to what extent this mapping is completed is not clear. They have published the tech stack, reference architecture, and test results.

Ceylonicus [6] is the first programming language that uses the Sinhala charset. The language has multilingual support meaning that the English character set is also supported. The implementation is done by using a Python interpreter. The language is more resembles Python programming language. Because of this dynamic typing and white space indentation are also there. The errors are given in the English language making challenges for users with limited English proficiency. Ceylonicus programming language was an inspiration for this research work and one of the aims of this research was to achieve more closeness to the natural language than Ceylonicus programming language.

3 Helaa programming language

The proposed language is a keyword-based programming language. The language only allows Sinhala characters. The Helaa code is compiled to byte code of the run time of the programming language by which the compiler is used. Currently, the compiler is fully implemented using the Java programming language. It is possible to develop compilers using other programming languages such as Python and JavaScript. The language is designed to be closer to the Sinhala natural language.

4 Features of the Helaa Programming Language

- Data Types (Integer and String)

- Conditional Structures (if and if else)
- For and while loops (can be nested)
- Arithmetic Operations (+, -, *, /, %)
- Logical and comparative operators
- Comments
- Exceptions can be given in the Sinhala language
- Usage of both infix and postfix notations
- Head-final design
- No need to indent
- Based on an expanded version of computational grammar for the Sinhala language[7]

5 Sample Codes

Figure 1 shows a sample code. The code only consists of variable assignments and print statements. Figure 2 shows the output of the code.

```
අංකය 5 යි
වචනය "ආයුබෝවන්" වේ
10 ලියන්න
අංකය ලියන්න
වචනය ලියන්න
"ආයුබෝවන්" ලියන්න
```

Figure 1: Sample code with variable assignments and print statements

```
10
5
ආයුබෝවන්
ආයුබෝවන්
```

Figure 2: Output of the above code (figure 1)

Figure 3 shows how basic arithmetic operations can be declared in the Helaa programming language using words. Note that not only this syntax but the mathematical expressions can also be defined using the standard mathematical notations in the Helaa programming language. Figure 4 shows the output.

```
පුටු 2 යි
මේස 4 යි
පුටු සහ මේස වල එකතුව ලියන්න
පුටු සහ මේස වල ගුණිතය ලියන්න
මේස වලින් පුටු අඩු කරන්න
පුටු වලින් මේස බෙදන්න
```

Figure 3: Basic arithmetic operations defined using words

```
6
8
2
2
```

Figure 4: Output of the above code (Figure 3)

Figure 5 shows an if-else construct and an if construct and Figure 6 shows its output. Nesting is supported and the end of a block is declared using the dot (.) character. Note the dot character at the end of the two if statements.

```
අංකය 10 යි

අංකය 200 ට වඩා විශාලයි නම්
"ඔව්" ලියන්න
නැතිනම්
"නැත" ලියන්න.

අංකය 200 ට කුඩා හෝ සමානයි නම්
"කුඩා හෝ සමානයි" ලියන්න.
```

Figure 5: An if-else construct and an if construct

```
නැත
කුඩා හෝ සමානයි
```

Figure 6: Output of the above code (Figure 5)

Figure 7 shows how a for loop is defined in the language. The code prints 10 to 19. 'අංක' is the loop iteration variable. As aforementioned, nesting is supported, and the end of a block is defined by using a dot character. Note the dot at the end of the code.

```
අංක 10 සිට 20 තෙක්
අංක ලියන්න.
```

Figure 7: For loop

Figure 8 shows how a while loop is defined using the language. The loop is an infinite loop. Note the dot at the end of the code.

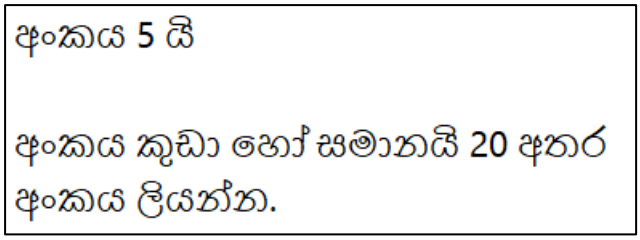


Figure 8: While loop

Errors can also be given in the Sinhala language. Figure 9 shows how the error is thrown when there is a division by zero.

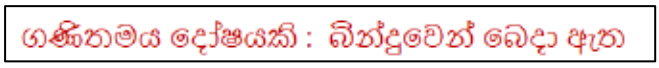


Figure 9: Division by zero error

6 Implementation

The compiler was implemented using the ANTLR 4. It is a powerful open-source parser generator. It supports generating parsers for many languages such as Java, Python, and JavaScript. Figure 10 shows the architecture diagram of the Java compiler.

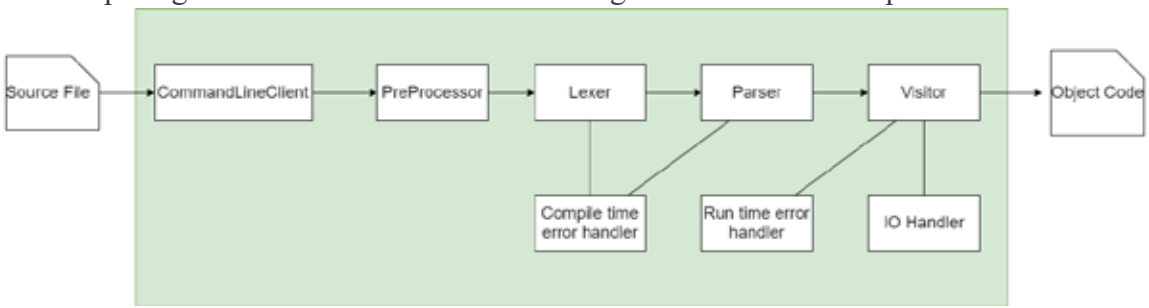


Figure 10: Architecture Diagram

The following are the main components of the Helaa compiler.

CommandLineClient – The command line client is the user-facing interface of the compiler. Currently, there are few compiler options that are supported. Not only a command line, but it is possible to develop a GUI client as well.
Preprocessor – The preprocessor takes the file name from the command line client and retrieves the source text. Currently, this only removes the comments.
Lexer and Parser – The lexer and parser are generated by the ANTLR 4 parser generator. Lexer identifies the basic lexical units and categorizes them into different lexical classes. The parser recognizes the meaningful structures using the lexer output.
Visitor component – This is the place where the parse tree manipulation happens. The run-time errors are caught and passed to the run-time error listeners.
Error Listeners – There are two error listeners: The compile time error listener and the run time error listener. The two listeners capture the compile time errors and the run time errors

respectively, and then map those errors to pre-built Helaa language exceptions which have the equivalent Sinhala language error messages.

7 Results and Evaluation

7.1 Test dataset

The compiler was tested against the designed Helaa codes, as single statements, as multiple statements combined, and also as nested statements up to 2 levels. Variable assignment, conditional statements (if-else), for and while loops, print statement, 4 basic arithmetic operations with modulus operation, and comments were executed as single statements. Codes with multiple statements combined are then tested. Section 5 discusses these multiple statement codes that were tested. Also, two erroneous codes (compile time and run time) were executed in order to test the error output. Figure 9 shows the output of the code tested for run time Arithmetic errors.

7.2 Accuracy

For the above-described dataset of codes, the compiler gave the expected outputs for every code segment including the expected output for two erroneous codes. Section 5 discusses these outputs along with the test source codes. But given the infinite number of combinations of statements including the nested statements, that can be fitted into a single code, more test codes should be executed. Even though the error mapping is not complete, errors can be present in different forms in a code, so it should also be tested more.

A user survey and user studies were conducted in order to evaluate the research work and to gather feedback.

7.3 User survey

A user survey was conducted to collect feedback from the software developers. The sample population was either university students or industry professionals. The students have at least completed a six-month industrial placement. 61 responses were gathered. 16 respondents were postgraduates and 45 of them were undergraduates.

The Ceylonicus language was used as a benchmark in this survey. The users were asked to give their opinions comparatively with the Ceylonicus language in multiple places. The users were asked to respond on different properties of the language such as expressiveness, readability, whether it is closer to the natural language or not, easiness to code, ease of learning, and similar properties. The researchers also asked questions like the opinion of the respondent on whether a Sinhala language-based programming language is needed and other similar questions.

Overall, the researchers found out that the language is better performing than the Ceylonicus programming language in some aspects like beginner-friendliness, closeness to natural language, readability, and but not in some aspects like easiness to code. When considering the programming language design, the researchers found that loops were

confusing. Most of the respondents found it difficult to identify how the loops work. The majority of the comments were positive and encouraging.

The question, 'In what language would you like to have the error messages?' was asked and, interestingly, the majority of the participants (figure 14) responded that both English and native language are needed at the same time, which is an unexpected answer. The researchers got to know that the English error messages which are thrown in every widely-used general-purpose languages, are not that descriptive. So, the developers find it difficult to locate and understand the problem with English-only error messages.

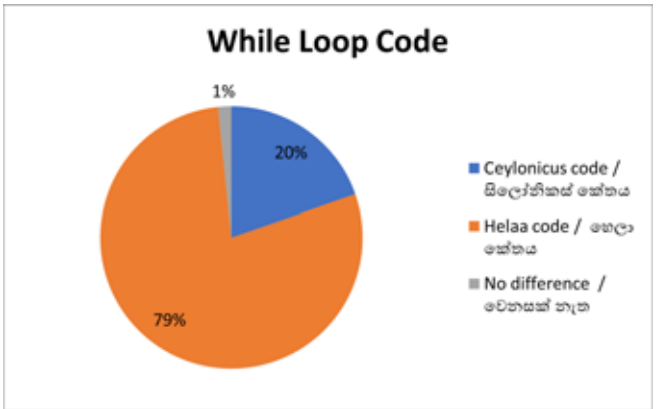


Figure 11: Equivalent while loop codes were given in both Helaa and Ceylonicus languages and the participants were asked to respond on what is closer to the Sinhala language. This chart illustrates the responses.

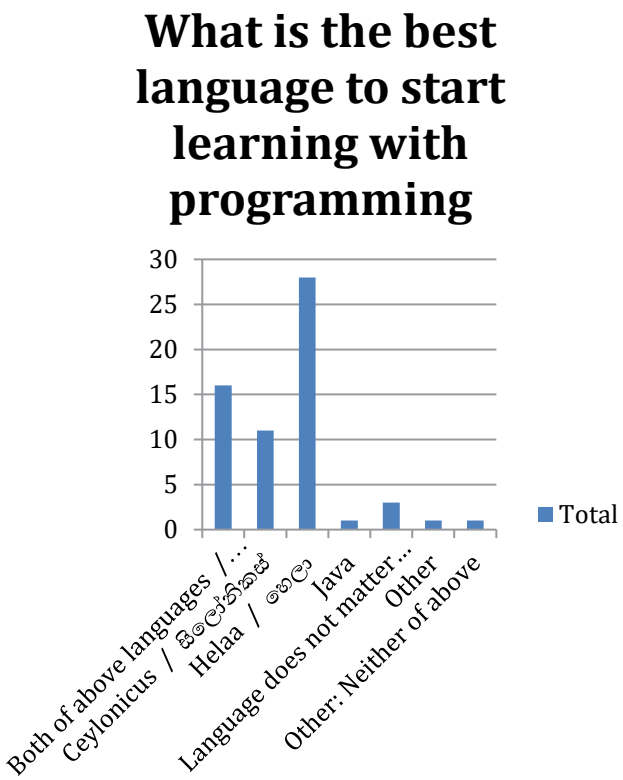


Figure 12: Responses to the question 'What is the best language to start learning programming?'

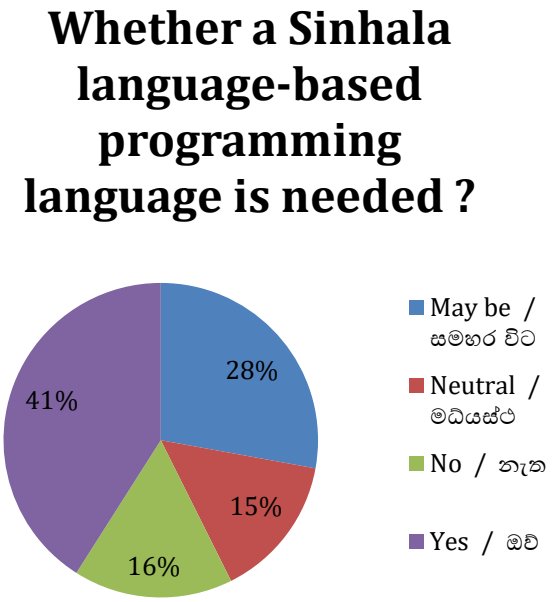


Figure 13: Responses to the question, 'whether a Sinhala language-based programming language is needed?'

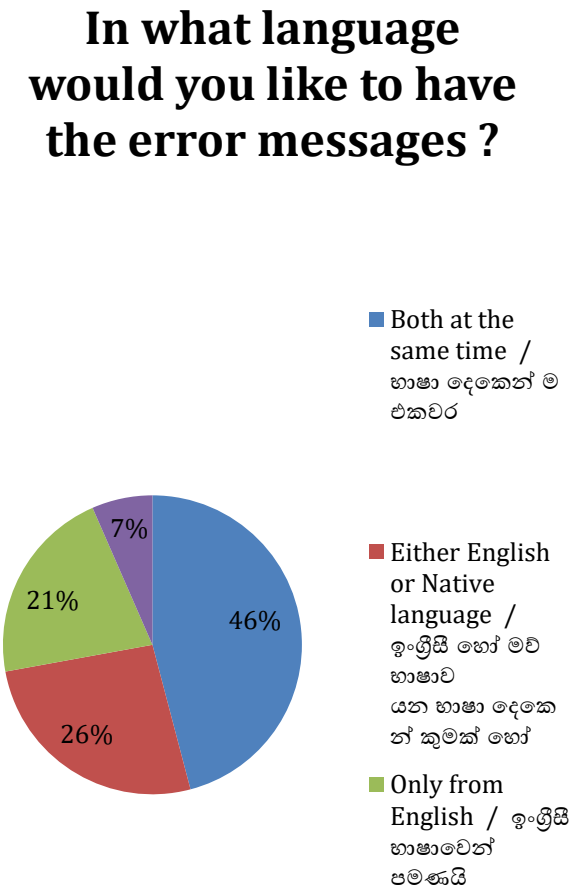


Figure 14: Responses to the question, 'In what language would you like to have the error message?'

7.4 User studies

There were two user studies conducted involving two students who had little to no knowledge of programming. The participants were first taught the main constructs in the Helaa programming language and then asked to develop a simple program for a pre-defined problem.

By engaging in this exercise, it was found that the beginners find it easier to understand the if-else statements in the language, but it is difficult to grasp the concept of a for loop. Finally, they were able to understand the concepts and develop solutions near the expected solution. Both participants learned a couple of new concepts on their own while involved in this exercise. One participant commented that the language is not complex, but it would

be better if it is simpler. Also, the participant commented that the language syntax is easy to remember to some extent.

8 Future Work

There are many possible future developments of this language exist.

Error mapping – The error mapping is yet to be fully completed. Because the exception hierarchy is huge and complex, and it is based on the underlying programming language of the compiler. Another main issue to be addressed is that there are no direct Sinhala translations for many English technical words that are used in error messages.

Remaining constructs – There are constructs that are not supported by this programming language. Some of the important constructs are the data structures (ex: arrays) and functions. Also, the language is not an object-oriented language, so how the object-oriented approach can be integrated into this kind of language is yet to be explored.

Morphology - Sinhala language is a morphologically rich language. These rich morphological properties of the language could be exploited when designing a Sinhala programming language so such a language could be simpler, more powerful, and closer to the Sinhala natural language.

ACKNOWLEDGMENTS

The researchers are grateful to all the participants in the user survey and the user studies which was carried out.

REFERENCES

[1] Y. Bassil, "Phoenix - The Arabic Object-Oriented Programming Language," *Int. J. Comput. Trends Technol.*, vol. 67, no. 2, pp. 7–11, 2019, doi: 10.14445/22312803/ijctt-v67i2p102.

[2] S. Kanemune and Y. Kuno, "Dolittle : an object-oriented language for K12 education," *EuroLogo2005*, pp. 144–153, 2005.

[3] P. Wang, "PseuToPy: Towards a Non-English Natural Programming Language," 2021. doi: 10.1145/3446871.3469787.

[4] A. Kamal, M. N. Monsur, S. T. Jishan, and N. Ahmed, "ChaScript: Breaking language barrier using a bengali programming system," *8th Int. Conf. Electr. Comput. Eng. Adv. Technol. a Better Tomorrow, ICECE 2014*, pp. 441–444, 2015, doi: 10.1109/ICECE.2014.7026875.

[5] S. Srivastava, U. C. Jaiswal, S. Lamba, and T. V Prabhakar, "Understanding Design Elements of Hindi Coding Application," *Bull. Tech. Comm. Learn. Technol. (ISSN 2306-0212)*, vol. 20, no. 2, pp. 19–26, 2020.

[6] "GitHub - RezSat/Ceylonicus: Ceylonicus is the first Sinhala Programming Language (first to support sinhala-based syntax). It can run sinhala and english syntaxes in a single code-base." https://github.com/RezSat/Ceylonicus (accessed Jul. 03, 2022).

[7] C. Liyanage, R. Pushpananda, D. L. Herath, and R. Weerasinghe, "A computational grammar of Sinhala," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 7181 LNCS, no. PART 1, pp. 188–200, 2012, doi: 10.1007/978-3-642-28604-9_16.

Alternative Transport Layer Protocol

Harshi P. S. Ratnayake

Informatics Institute of Technology affiliated to University of Westminster, Colombo 06, Sri Lanka, harshiratnayake@gmail.com

Geethapriya Liyanage

Department of Computer Science and Engineering, University of Moratuwa, Moratuwa, Sri Lanka, g_liyanage@hotmail.com

ABSTRACT

The internet mainly runs on TCP, one of the main protocols of transport layer. To start a TCP session, a three-way handshake takes place between the two parties involved in the communication, and this happens in plain text. The security protocols also run-on top of TCP and at the beginning of the communication those protocols also communicate in plain text until encrypted communication starts. Due to this, communications happening through TCP are subject to spoofing, sniffing, SYN flooding attacks, TCP reset attacks, and session hijacking attacks.

This research intended to find a comprehensive solution that would overcome these problems of TCP. For that purpose, TCP's security issues were analyzed using literature, observed inside a test environment, and collected information about TCP security problems and possible solutions by way of interviewing technical experts. Using those collected information a framework was developed identifying the main requirements for a secure transport layer protocol that would replace TCP. Then using that framework, a new protocol was designed. The proposed solution is called Alternative Transport Layer Protocol, which could be utilized as a guide to develop and implement a fully functional protocol in future. This paper presents a solution that would overcome the vulnerability of sequence numbers in TCP, which is the main area of the TCP header that would be manipulated by the attackers. The solution is called the question-and-answer method. In addition, this new protocol provides end-to-end communication, authentication, data confidentiality and integrity, handles the TCP re-transmission problem, and is suitable for large scale deployment.

The protocol was evaluated against the framework, other transport layer protocols, and by experts to receive feedback about the concept. Further enhancements to the protocol, followed by developing and testing this protocol in a live environment are the next steps to continue this research.

KEYWORDS

TCP, security, three-way handshake, attacks, DoS, man-in-the-middle, sequence number, countermeasures, framework, alternative, confidentiality, integrity, authentication, question-and-answer

1 Introduction

TCP is a connection-oriented protocol where the two end points; a client and a server, have to set up a communication channel using the well-known 3-way handshake [1]. The TCP protocol states that it guarantees the end-to-end connection over a communication path, and it is not restricted by the number of hops involved, but the ability to view the packets using a packet sniffer or analyze a packet using firewall questions the end-to-end communication. The acknowledgments (ACKs) are used to find out any lost data segments and those will be re-transmitted. TCP has guaranteed delivery, congestion control and flow control, but it was not built with the security features [1], [2]. The communications happening through TCP are protected by a set of security protocols such as secure sockets layer (SSL), and more recent transport layer security (TLS). These security protocols are built with middle box compatibility and backward compatibility

which make them vulnerable to version downgrade attacks [3]. Due to this, communications happening through TCP protocol are subject to man-in-the-middle (MITM) attacks including synchronization (SYN) flooding, spoofing, session hijacking, TCP reset (RST) attack, sniffing, and data tampering [4].

Once there is an established, successful system in place people will not try to replace it, but they will only do the patches to cover any loopholes and improve the efficiency [5]. Any new system built to replace that old system needs to bring a phenomenal improvement to the output received from the previous system, otherwise it would not be implemented worldwide. This is the case with the TCP not been replaced by another protocol for 40 years. When a protocol could not evolve because of its existing deployments limiting the areas that it could be extended or evolved, it is called ossified. TCP is an example of ossification [5].

The TCP was designed four decades back as an end-to-end communicating protocol. But the implementation of various middle boxes like proxy server, Network Address Translator (NAT) has violated the true end-to-end communication, thereby making TCP prone to man-in-the-middle attacks [5]. It was built without any security features, so it has security issues, that makes TCP communications vulnerable to attacks. As solutions SSL and TLS protocols are used on top of TCP. TCP three-way handshake happens in plain-text, and it is not secure. The applications that run on top of TCP are required to integrate security features to overcome security loopholes of TCP. Introducing a completely new protocol is a tedious, time and resource consuming task. Once a protocol is functioning well in large scale people in the industry are reluctant to replace it. Introducing a new protocol to replace TCP would be ground-breaking research, but still a doable task. Improving TCP is also another option that is available, but such improvements by previous researchers resulted in creating new vulnerabilities and will be discussed in the literature review.

In this research work, the main objective was to develop a framework for a secure transport layer protocol, and using that as a guide, to design a secure transport layer protocol. During the research necessary aspects for the new framework was identified, and those mainly cover integrity, confidentiality, and authentication, while preserving end-to-end communication. The re-transmission issue of TCP also needs to be addressed. Any solution that comes forward to replace TCP should be possible to be implemented worldwide, as TCP is the main protocol running the internet. Taking the developed framework as the guideline a new protocol was developed, named Alternative Transport Layer Protocol (ATLP). The designed protocol can be used to develop a fully operational protocol that would be suitable to replace TCP.

2 Literature Review

TCP has been in operation for more than 4 decades and it has been contributed evolution of the internet to this massive scale [6]. Reasons for TCP ossification were discussed in detail by a group of researchers, and they state that the transport layer has become rigid, further progression has become tough and a difficult task [5]. Even though TCP is still in operation it has many weaknesses that have led to attacks on communication systems using it. Those weaknesses found in the literature will be discussed in the coming paragraphs.

The research by Nguyen, Mimura and Tanaka states that in virtual environments denial of service (DoS) attacks were possible by use of TCP re-transmission before the three-way-handshake completes, because the servers start to re-transmit when the ACKs are not received [7].

Reference [8] states that Retransmission Time Out (RTO) is important in a TCP communication to achieve the reliability aspect of the protocol. That means in a TCP communication when the sender sends a packet and if acknowledgement is not received the RTO timer will be expired and sender can assume the packet is lost and do a retransmission. In an ideal world the retransmission time out is equal to the Round-Trip-Time

(RTT), but in the real world it is important to set the RTO based on calculations. If there's a large RTO time it will allow time for an attacker to initiate a DoS attack [8].

In a TCP communication if the three-way handshake is not finished, those connections are called half-open connections. The server must store those half-open connections in a queue [9]. An attack can be initiated by filling up that queue quickly, so that server doesn't have space to store the new half-open connections. In other words, the server will not be able to accept new SYN packets, and nobody will be able to connect to the server after that. In such an attack the attacker just does two things:

1. Sends lots of SYN packets to the server, from different IP addresses [9].
2. Does not complete the 3rd step of the TCP three-way handshake [9].

TCP saves details about half-open connection as Transmission Control Block (TCBs) in a queue. This makes TCP vulnerable to DoS or DDoS attacks, as an attacker can send forged packets to take up the space in that queue and make the space in that queue unavailable for new legitimate connections [9].

The researcher Steven Michael Bellovin has conducted research on the Transport Layer protocols and security. In his first paper he states that TCP/IP is widely deployed in those days despite serious security flaws inherent. The types of attacks he has identified were sequence number spoofing, routing attacks, source address spoofing, and authentication attacks [10]. The sequence number spoofing attack is an exploitation of the three-way handshake, where a man-in-the-middle predict the sequence number and send a sequence number to impersonate a trusted host. The trusted host will receive it and attempt to reset the connection to the server as he didn't send it. To avoid such a reset the intruder will choose a time where the real host is down or flood it with connection requests that will result in queue overflows and the ACK from the server will be lost. But the intruder continues by sending the ACK he predicted for the server, impersonating the host and keep the communication on [11].

During this literature review it was discovered new solutions designed for the transport layer are not widely adopted or accepted by the network community over the years. Most of the protocols proposed in the literature have not taken security aspects into consideration during the design phase. Even though several researchers have proposed new protocols to replace TCP like Stream Control Transmission Protocol (SCTP), and Datagram Congestion Control Protocol (DCCP), none of the protocols were widely accepted in the industry. The large-scale adoption of those protocols on the internet has been hindered by the ossification of the networking equipment [12]. On the other hand, QUIC protocol developed by Google which is implemented in the user space rather than in the kernel, does not require any upgrades to operating system and is currently in operation [12]. But QUIC is based on User Datagram Protocol (UDP), which does not guarantee packet delivery.

One of the latest research projects done in the transport layer domain is the TCP Wave. Altogether two research papers were published: TCP Wave, and TCP Wave over Linux [13], [14]. The latter research was on implementation of TCP Wave on Linux platform and related work. TCP Wave is an alternative transport protocol proposed in this research, which supports burst transmissions that are taking place in the modern networks, and removes the ACK based transmission paradigm. But security was not discussed in TCP Wave.

Another research has proposed a new protocol called "Graphene" which is designed to address security concerns in cloud environments [3]. The protocol is designed for faster reconnection and high performance, with a focus on architecture, protocol design, and implementation. According to the researchers, Graphene provides good protection against many attacks possible inside clouds and has been verified through evaluation. They note that while cloud platforms rely on TLS version 1.3 and 1.2, and DTLS to secure data-

in-transit, these platforms can still become vulnerable to related attacks [15]. The involvement of central key server (CKS) makes their proposed protocol vulnerable.

A group of researchers conducted a study in 2017 and found that transport layer is subject to ossification, making further advancement difficult [5]. They proposed a solution to overcome the challenges caused by middleboxes and limitations of APIs on TCP's evolution. That research team developed an adaptable transport layer framework that allows transport layer solutions to integrate and cooperate in a flexible, application-independent manner. The researchers identified several key requirements for such a framework, which include:

1. API flexibility
2. Deployability
3. Extensibility
4. Guided parameter value selection
5. Scalability

The research team categorized the five key requirements and further subdivided them into more specific requirements [5]. The study highlights the necessity for a transport layer framework that can provide a solution to the issue of ossification and allow the implementation of novel network and transport functions in a more adaptable and non-specific manner. As a drawback of this research, it was evident that the proposed framework has no allocation for security requirements. It could have been a better framework if security requirements were also analyzed in depth and integrated.

Much research has been conducted since the inception of TCP to improve its performance and security [5]. The research work discussed in the above paragraphs proposes new protocols that would complement the TCP to improve either performance, efficiency, or the security. But none of the researchers have come up with a single solution that would replace TCP integrating the security aspects and avoid the usage of sequence numbers.

3 Observations in the Test Environment

To observe whether TCP three-way handshake have a security issue or issues, two machines were configured in a virtual environment as a client and server and opened a TCP communication and tried to interfere into the communication and view traffic.

For this simulation, C socket programming was used to configure the sockets for the client Linux machine and the server Linux machine. Simple text message was retrieved from the server to the client. The communication happening between the two machines was captured from another virtual machine using Wireshark software. The captured packets were analyzed to check whether there are any security loopholes. The packet content was visible to the packet sniffer: source and destination IPs, source and destination ports, SYN, SYN-ACK, ACK handshakes, sequence numbers, and the plaintext message communicated in the payload. Figure 1, and Figure 2 depict the observations in this test environment. These observations proved the existence of security issues in TCP. TCP reset, TCP session hijacking, and SYN flooding attacks were observed to be possible within this environment.

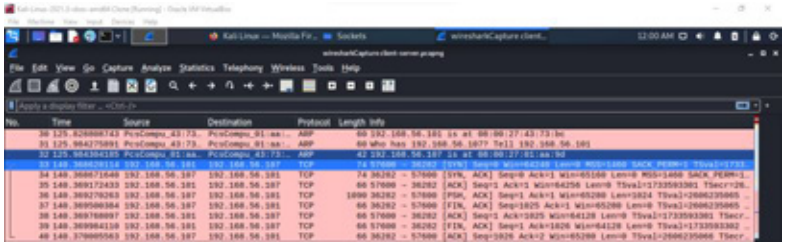


Figure 1: TCP three-way handshake packet capture revealing the source, destination IP and port addresses, and sequence numbers

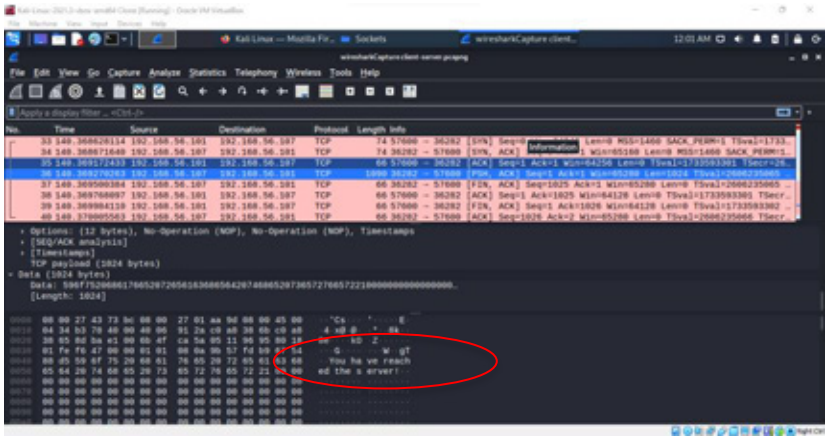


Figure 2: Payload data revealed in the TCP communication

Any third party sitting in the middle can alter the packets sent between the server and the client, and thereby interfere into the communication. For this the attacker needs to have data about the source and destination and needs to find the sequence number [16]. Even though there are schemes proposed to make the initial sequence number (ISN) as random as possible, with the very fast networks of nowadays, this 32-bit sequence number space can be easily used-up within few seconds and it will get wrapped around [17]. The sequence number guessed or derived by the attacker could become a valid sequence number in an already started communication.

To test a SYN flooding attack, such an attack was directed towards the server using a tool called Synflood, it is a tool available in many Linux versions [18]. When the SYN flood attack was initiated, after a few seconds, the server became unresponsive. When the packets were analyzed, it could be observed, the server responded only to 5 of the SYN packets came from the forged IP addresses. After that the server became unresponsive. The output of the packet capture is shown in the Figure 3.

In a TCP reset attack an attacker spoofs a TCP packet from client to server or server to client, with RST flag raised. To make a successful TCP reset attack, source IP, source port, destination IP, and destination port, needs to be correctly filled in the spoofed packet. In addition to that, the sequence number of the spoofed packet needs to be correct, otherwise the receiving party will discard it. If an attacker becomes successful in finding the correct sequence number, they can break the connection between the server and the client. Figure 4 and Figure 5 show the implementation and the packet sniffer output of a TCP reset attack using a third virtual machine to send a forged packet and close the TLS connection between client and server machines in the same virtual environment.

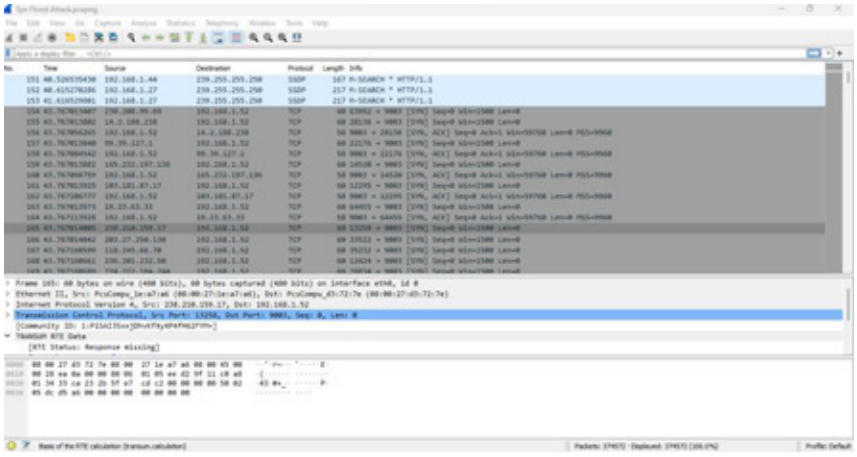


Figure 3: Packet capture of the SYN flooding attack

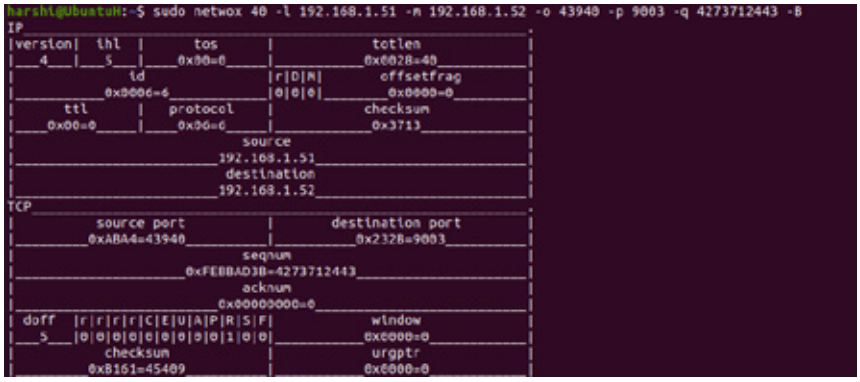


Figure 4: Implementing a TCP reset attack

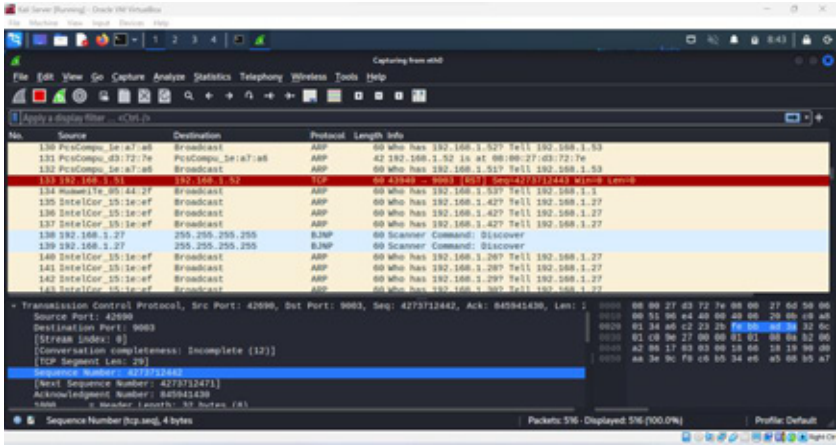


Figure 5: The captured forged TCP Reset (RST) packet

4 Framework for a Secure Transport Layer Protocol

Rather than modifying the existing TCP, to integrate security, which was the usual path the previous researchers have followed, it would be beneficial for the body of knowledge to come up with a conceptual protocol to replace TCP. Introducing a completely new protocol is a tedious, time and resource consuming task. Once a protocol is functioning well in large scale people in the industry are reluctant to replace it. Introducing a new protocol to replace TCP would be ground-breaking research, but still doable. Using the results from the data collection phase of the research, the researchers were able to design the framework for a secure transport layer protocol. The 6 main areas were considered when designing this framework for a secure transport layer protocol, and those six main areas are listed below.

1. The TCP was designed four decades back as an end-to-end communicating protocol. But the implementation of various middle boxes like proxy server, Network Address Translator (NAT) has violated the end-to-end communication, thereby making TCP prone to man-in-the-middle attacks [5]. It was built without any security features, so it has security issues, that makes TCP communications vulnerable to attacks [19]. Therefore, ensuring end-to-end communication was identified as an essential aspect to be covered by the framework.
2. To ensure the end-to-end communication channel, authenticating the two parties involved in the communication is necessary. By analyzing the data collected through interviews it was revealed that the majority of the interviewees agreed to this as a necessary requirement. Therefore, an authentication mechanism is identified as the second aspect for the framework.
3. TCP handshake happening in plain-text and any man-in-the-middle can view the content of the packets that are being transmitted. Even if SSL/TLS is used to provide encryption, the acknowledgements for the TLS packets also take place using TCP as it was revealed in the observations. Confidentiality of communication is a vital aspect as per the information security guidelines given by National Institute of Standards and Technology (NIST) [20]. Therefore, confidentiality of the data was identified as the 3rd most essential aspect for the framework development.
4. Any data packets captured by an adversary could be subject to alterations, re-transmissions or replays [9]. The integrity of the data is compromised in such scenarios. Integrity is identified as the 4th essential aspect.
5. Re-transmission problem of TCP arises when the client does not acknowledge the received packets and the server will again send the packet thinking the packets are lost. If the server receives several of these kinds of connections, the server resources will be taken up and the server will become unresponsive. If the same thing is happened at the client side in parallel, both server and client will get into a deadlock, because client is waiting for another set of packets and do not acknowledge the retransmitted packets, but simply drop those and vice versa [7]. The developed framework is required to address this re-transmission problem of TCP, and that is considered as the 5th aspect for the framework.
6. TCP is a worldwide accepted and deployed protocol, and any protocol that would replace TCP needs to have that level of acceptance and deployability in large scale [5]. This criterion was identified as the 6th aspect for the framework.

Papastergiou et al. have identified changing the socket API as a limiting factor, and their framework was discussed in the literature review [5]. But the expert opinion at the interviews revealed kernel level programming can be done for a new protocol and API cannot be considered as a limitation. Since this is a framework for a new protocol, backward compatibility and middle box compatibility were not required to be integrated, as those will result in vulnerabilities. The developed framework is shown in Figure 6.

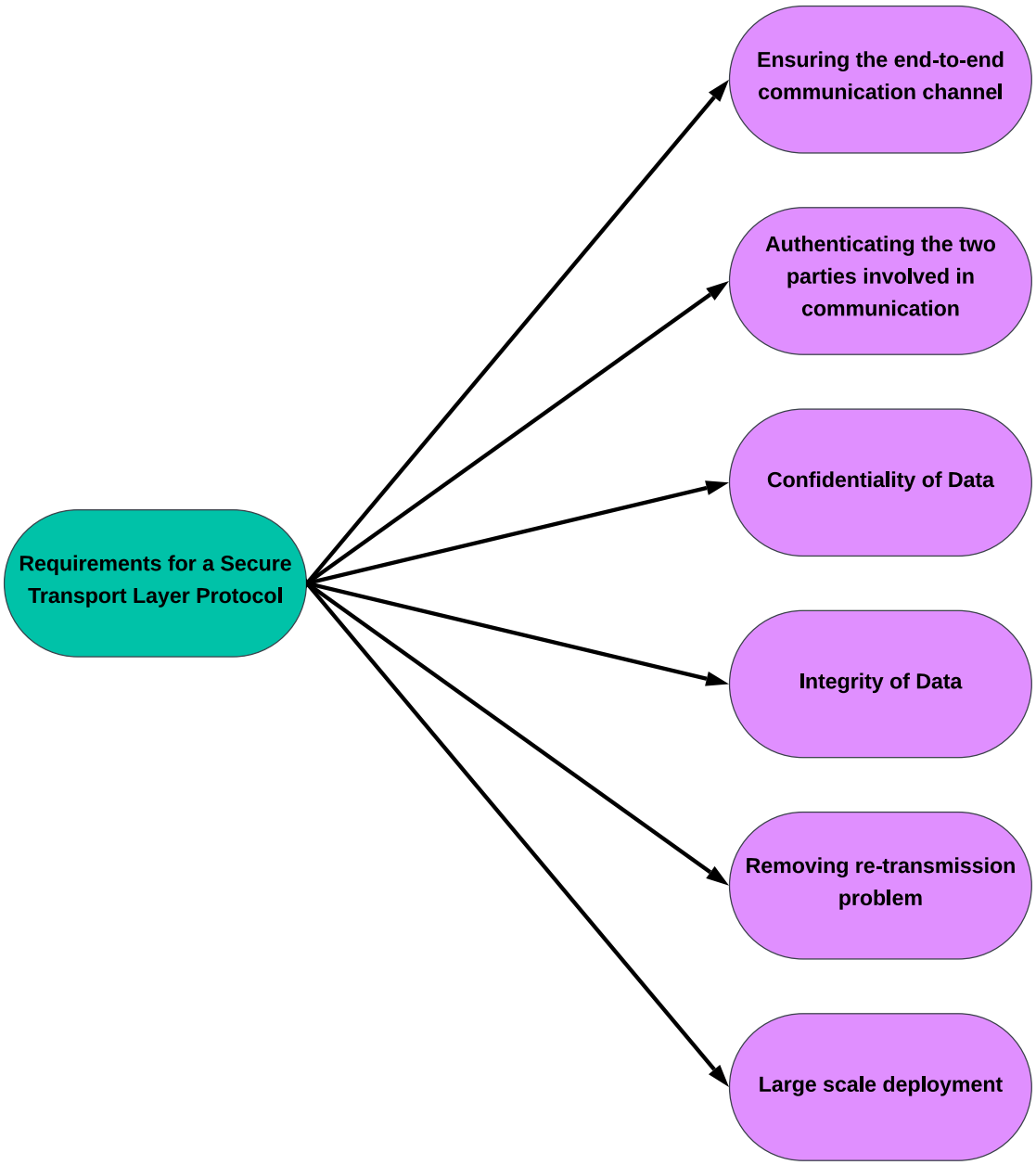


Figure 6: Framework for the secure transport layer protocol

5 Proposed Conceptual Protocol for Transport Layer

Taking the above conceptual framework as the guide a new protocol for the transport layer was designed as the outcome of this research. This conceptual protocol was named Alternative Transport Layer Protocol (ATLP). The functional requirements for the protocol were well defined before finding solutions for each requirement. The functional requirements of ATLP were identified by the researchers and listed below.

1. Protocol packet format

2. Handshakes – establishing a connection.
3. State transitions
4. A replacement for the sequence numbers – question-and-answer method
5. Authentication – using digital signatures
6. Special features

5.1 Protocol Packet Format

ATLP's packet structure is defined in a way to include the source port, destination port, signature, public key, certificate, flags, and packet length. Like TCP, in ATLP data segments are sent as internet datagrams. Therefore, the source IP address and destination IP address come from the Internet Protocol (IP) and the packet structure is shown in Figure 7.

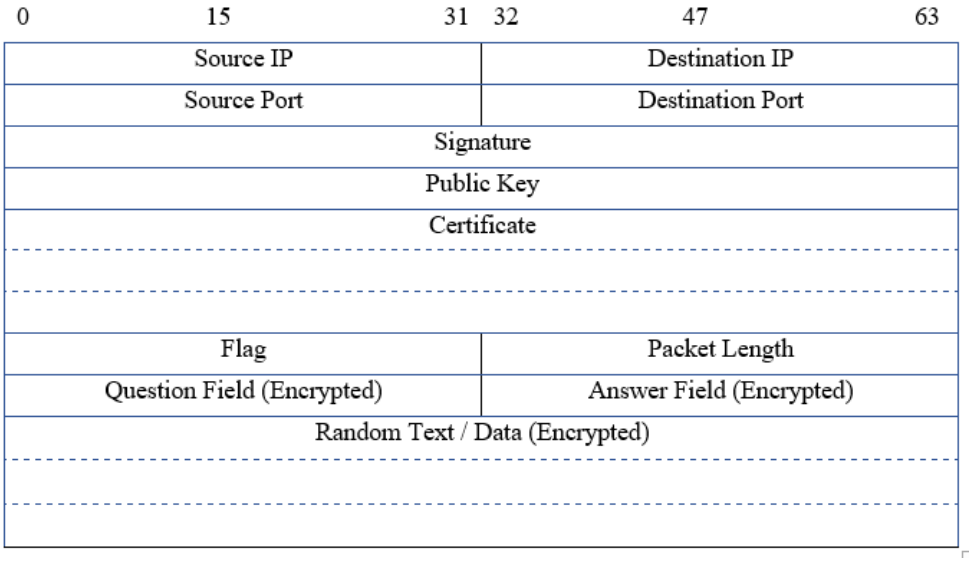


Figure 7: The structure of the protocol packet

5.2 Handshakes of Alternative Transport Layer Protocol

The handshakes of the client-server communication in ATLP are depicted in Figure 8. The client and the server are termed as "A" and "B" respectively. Both A and B have self-generated public and private key pairs and certificates. There are six message flags: Hello 1, Hello 2, Data, Close 1, Close 2, and Wait, and for special connection scenarios Abort flag can be used. When the protocol evolves there is enough space in the Flag field to introduce new flags. The steps involved in establishing a connection are discussed in detail below.

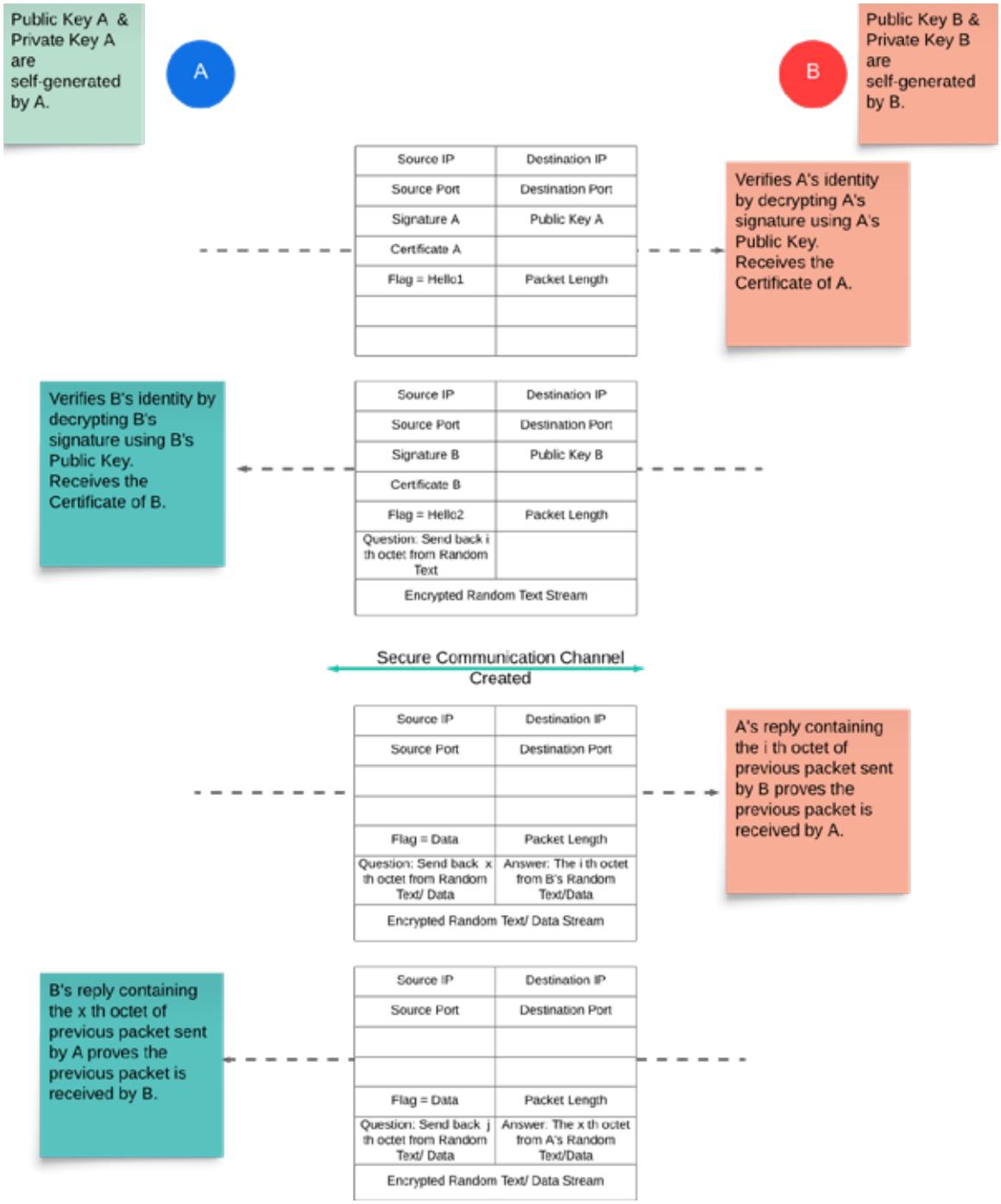


Figure 8: Alternative Transport Layer Protocol – Handshakes

1. To start the communication A sends the first packet with flag set to “Hello 1”. This packet contains the source and destination IP addresses, source, and destination port numbers. In addition, it contains the signature of A, public key of A, Certificate of A, and the packet length. This packet is named “First Handshake”.

2. When this message is received by B, B can verify A's identity by use of the signature and the certificate. A's signature is the hash of A's certificate generated using the private key of A. A's public key is now informed to B. B will compare the recalculated hash of A's Certificate and the decrypted A's signature. If two values are not equal, that implies A's certificate is tampered with, and the packet will be dropped by B. If A's signature is verifiable, B sends the response packet with the flag set to "Hello 2". This packet too contains the source and destination IP addresses, source, and destination port numbers, signature of B, public key of B, and certificate of B. But in addition, it has two other fields, an encrypted question field, and an encrypted random text field. Those are encrypted using A's public key. The question field asks for an octet number from the random text field, where A must send back as the response. This packet is named the "Second Handshake".
3. When A receives the second handshake message, A can verify B's identity, as B's signature is there. Also, B's public key is now delivered to A. To continue the communication, A must respond with an answer to the question asked by B. Receiver A decrypts the random text field using A's private key, finds and locates the octet from that random text, and places that in the next packet as the answer. Before sending the response packet, A adds the data that needs to be transmitted, encrypts it with A's private key, and sets the flag to "Data". Then adds an octet number as the question asking B to send back, when responding or acknowledging the receipt of the packet, both question and answer fields are encrypted. Communication will be continued between A and B using this question-and-answer method to acknowledge the receipt of the packets.
4. After the second handshake "a secure communication channel" could be established between A and B; where both A and B can encrypt and decrypt the communication between them using private and public key pairs.
5. To end the communication A or B can send the packet with flag "Close 1", again with a random text or last part of the data encrypted with the relevant party's key. The receiving party will respond back with the answer, and the flag "Close 2" if no more packets are to be sent and release the resources in use. If there are more packets to be sent the flag would be "Wait". When the "Close 2" is received by the other party, it will close the connection, releasing the resources, if it is a server, it will return to "Listen" state. If the "Wait" message is received, the other party will not close the connection, but wait for more packets.

5.3 State Transition Diagram

The state transition diagram for this conceptual protocol is shown in Figure 9. There will be 12 states defined for a connection in this protocol, those states are listed below.

1. Listen
2. Hello 1 Sent
3. Hello 1 Received
4. Hello 2 Sent
5. Hello 2 Received
6. Connection Established
7. Close 1 Sent
8. Close 1 Received
9. Close 2 Sent
10. Close 2 Received
11. Wait
12. Closed Connection

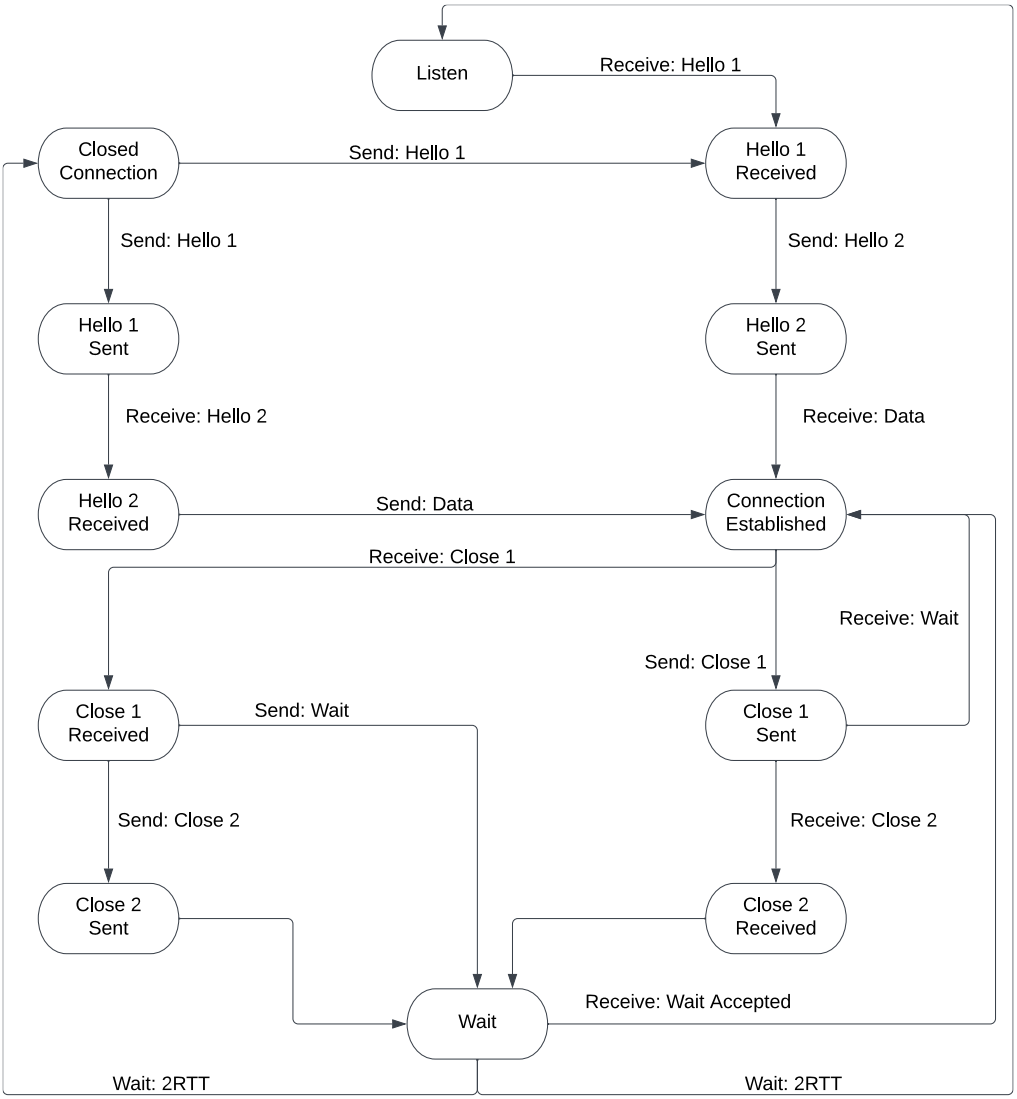


Figure 9: Alternative Transport Layer Protocol - State Transition Diagram

The transitions from one state to another take place, based on the current state and the flag of the packet received in that state. For example, server B will be in the "Listen" state at the beginning of the communication and client A will be in "Connection Closed" state. With the receipt of "Hello 1" the state transitions will be started and will progress according to the diagram and reach the "Connection Established" state. To close the connection either server or the client send the "Close 1" packet, and if the other party sends back the "Close 2" packet, the connection will be closed, and the client will return to "Connection Closed" state and server will return to "Listen" state after a time interval equal to 2 times the RTT. If the other party sends a "Wait" packet, the connection will go to the "Wait" state at that end and the other end will move from "Close 1 Sent" state to "Connection Established" state. After receipt of "Wait Accepted" the wait requested side will again move to the "Connection Established" state.

5.4 Replacing Sequence Numbers - Question & Answer Method

Usage of sequence numbers was completely avoided, as most of the attacks have taken place by threat actors guessing or predicting the next sequence number to be used, when the signature of a TCP connection is known. In ATLP sequence numbers were replaced by question-and-answer method. To achieve the guaranteed delivery feature, the data sender asks the receiving party to reply with a certain octet of the message, in the next packet. Since the packet data is encrypted, man-in-the-middle (MITM) cannot find the message content and send a forged packet with the answer. Only the authenticated receiver can respond with the correct answer. This method provides guaranteed delivery as well as avoid any MITM attacks.

5.5 Authentication - Use of Digital Signatures

For the authentication of the client and the server this conceptual protocol will use digital signatures. The signature generation and verification processes are depicted in Figure 10. Both client and server will have self-generated public/private key pairs and a self-certificate. This conceptual protocol will not involve any Certification Authority (CA) or Public Key Infrastructure (PKI) servers.

At the client end, the self-certificate is fed into the cryptographic hash function generator, it generates a hash of fixed length. This hash is then sent through the encryption algorithm and encrypted using A’s private key and the output is the signature SA. Then the certificate and the signature are transmitted to the server side. The server B receives the message, and the certificate is sent through the hash function to calculate the hash of the received certificate. In addition, the server, i.e., the recipient, decrypts the signature SA using A's public key which was received with the message itself. If the calculated hash and the decrypted signature SA are equal the receiver can accept the signature to be valid and certificate content to be intact, that means integrity is preserved. Here the integrity is preserved because only A has its private key, and only A could have generated the encrypted signature which could be correctly decrypted with A’s public key.

The same processes are repeated for the server side’s signature verification. After the first two Hello packets, i.e., Hello1 and Hello2 packets, signature fields will be empty, because now both parties have received the certificate and the public key of the other party and encrypted communication has started. After Hello2, and authentication and integrity of the message is secured by the encrypted data.

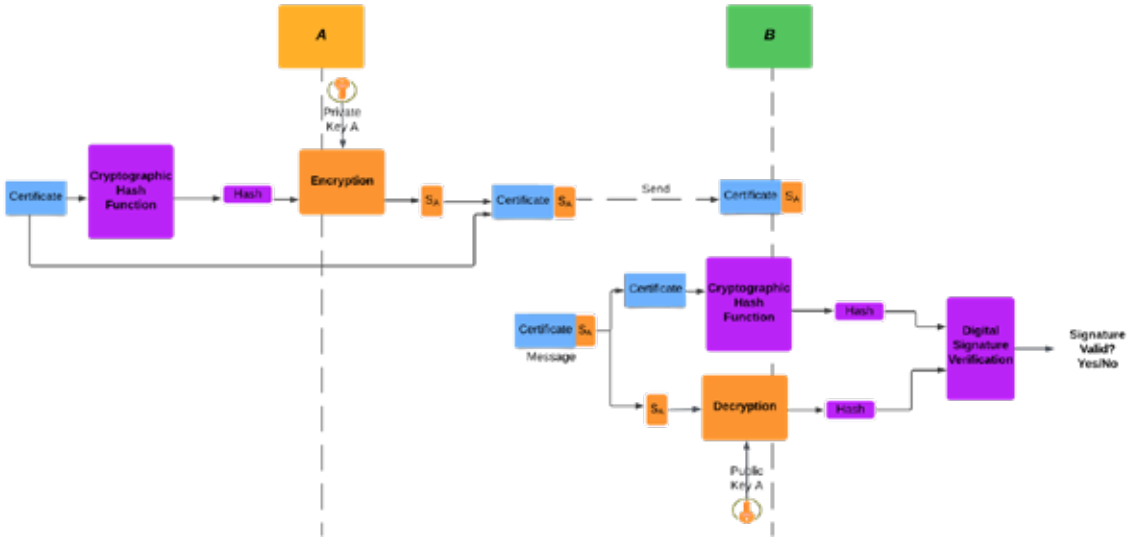


Figure 10: Alternative Transport Layer Protocol - Signature generation and verification processes

5.6 Features of the Protocol

RFC 4101 states that any protocol needs to answer three main questions: 1. The problem the protocol is trying to solve, 2. Type of data that is being transferred, and 3. The important, but barely visible features of the protocol [15]. For ATLP these questions were answered, as below.

1. The problem the protocol is trying to solve – Security vulnerabilities of TCP imposed by 3-way handshake.
2. Type of data that is being transferred – traffic between a client and a server through the internet.
3. Important features that are not obvious – listed below.

Therefore, the features of the protocol were defined referring to the handshakes, state transition diagram, and other specifications of the protocol.

1. Signatures - Ensure the authentication of the server and the client
2. Self-generated certificate - reduce the hassle of contacting certification authority (CA) every time the certificate expires, and any vulnerabilities associated with CA
3. Self-generated public/private key pairs - Overcomes the involvement of a CA, and thereby any vulnerabilities from the communication with the CA.
4. Client and Server drops packets that arrive without a signature - protection from spoofing attacks
5. Question-and-answer method - Used as the acknowledgement from the receiving party and the readiness to receive next packet.
6. Flags – The flags used in the communication are Hello 1, Hello 2, Data, Close 1, Close 2, Wait, and Anomaly.

6 Evaluation of the Conceptual Protocol

Evaluation of the protocol is carried out in two separate aspects. Firstly, the protocol was analyzed against the developed framework for the secure transport layer protocol and TCP. Then it was compared against the existing protocols and frameworks discussed in the literature.

6.1 Evaluating against the framework and TCP

The conceptual protocol was evaluated for its effectiveness in achieving the 6 aspects defined in the framework developed for a secure transport layer protocol and TCP.

1. Ensuring the end-to-end communication channel – In this proposed protocol the communication from client to the server is end-to-end encrypted after the first packet. The question-and-answer method of acknowledging makes it difficult for a third party to interfere into the communication channel. In TCP this end-to-end communication channel can be violated by the middleboxes but in ATLP it is ensured.
2. Authenticating – This conceptual protocol achieves this requirement using signatures transferred between two parties in the two handshakes. The self-certificate is sent to the other party encrypted using the private key as a signature, so that other party can view the certificate by decrypting it with the public key. This confirms the identity of the sender, as only the sender can encrypt with the private key of the corresponding public key, whereas, TCP does not have a mechanism to authenticate.
3. Confidentiality – Confidentiality of the data is secured by this new proposal. Only the first packet is sent as plain text to start a communication. From the second Hello onwards the random text fields or the data fields are encrypted using the public key of the receiver and receiver can decrypt using the corresponding private key. Unlike TCP from the second Hello onwards the random text fields or the data fields are encrypted.
4. Integrity – Data integrity is also preserved from the first handshake onwards, because in the first handshake the client's certificate is sent as the payload of the message, and the certificate's encrypted hash value is also sent to the serve side. If the certificate is altered by some MITM, it will be detected at the receiving end. Thereby the integrity is preserved by this conceptual protocol. From the second handshake onwards, communication is encrypted. Therefore, integrity is ensured throughout communication.
5. Removing re-transmission problem – In the proposed conceptual protocol this problem of TCP is solved by use of a timer. The timer is set to two times the RTT, for the receiver to respond to the received packet. After the elapse of the timer, the server will re-transmit the packet. Server will make 3 re-transmission attempts, and after that the connection will be closed, and server will move to the listening state. Due to this method, server resources will not be taken up by an unnecessarily long re-transmission queue, as server will close the connections after 3 attempts.
6. Large scale deployment – The proposed protocol uses self-certification and self-generated public/ private key pairs for clients and servers. Therefore, the requirement for a CA and public key infrastructure (PKI) has been avoided in this proposal. This makes large scale deployment feasible like TCP, but more secure, as security vulnerabilities posed by PKI are avoided.

6.2 Evaluating against existing protocols and frameworks

The proposed protocol was compared against the protocols and frameworks discussed in the literature review.

1. SCTP has similar features as TCP, and specifically it uses similar methods for windowing, error detection and re-transmissions [12]. Therefore, it is vulnerable to TCP's re-transmission problem. The proposed protocol does not continue to re-transmit, it rather makes 3 re-transmissions and then closes the connection completely if a response is not received. This ensures the availability of the server, because if the client is purposely avoiding answering re-transmit attempts, by this method server avoids a possible DoS attack.
2. DCCP is similar to UDP, and it does not provide reliable data transfer or ordered delivery [12]. Therefore, it is suitable for applications that do not seek reliability. DCCP does cover security aspects. The ATLP is equipped with acknowledgements, and thereby provides guaranteed delivery, together with security.
3. TCP Wave is another protocol that was proposed in recent years [14]. TCP Wave has omitted to address security side at the design stage or the implementation stage of the protocol. In ATLP security is considered from the framework development stage to the design stage.

4. "Graphene" is a protocol developed to address security concerns in cloud environments [21]. Graphene ensures authenticity of the users of the cloud using a Central Key Server (CKS) mechanism. The design was done without any middlebox or backward compatibility in mind, as those can contribute to version degradation attacks. Central key server (CKS) security is not discussed in their research. If the security of the CKS is compromised and private keys of all the clients and cloud servers are leaked, there can be negative consequences. Further, it is observed that communications required by client and cloud server for registration, cancellation, and distribution of keys with the CKS and the security aspects of such communications are also overlooked. Since ATLP uses self-generated keys, and does not rely on a central server, these security drawbacks were avoided in ATLP's design.
5. Papastergiou et al., propose a framework in their research [5]. Their research identifies the requirement for comprehensive transport layer framework to overcome the ossification and enable the use of new network and transport functions in a more flexible and non-application-specific way. The drawback of this research is that it was evident that the proposed framework has no allocation for security requirements. It could have been a better framework if security requirements were also analyzed in depth and integrated. The framework proposed in the ATLP research has considered the security side and has given more weight to that aspect during the research.

7 Future Research Direction

TCP was introduced to the world through RFC 793 in 1981 and has been in use for 4 decades and same is true with the Internet Protocol (IP) [1], [22]. Reference [23], has first proposed IP version 6 in 1995, but until recent years it was not widely implemented. It requires considerable amount of time for a protocol to evolve and receive wide acceptance.

No single solution has emerged yet to totally replace TCP and overcome its security issues. The future research efforts will lead to further develop this conceptual protocol towards fully operational transport layer protocol. The expert feedback indicated the question-and-answer method to be an innovative practical concept and it has added value to this protocol. The designed protocol ATLP could be presented as a Request for Comment (RFC) document to IETF. This research could be further advanced to the stage of developing and testing this protocol. The developed conceptual protocol needs to be evaluated for its effectiveness as an alternative to TCP, and how to implement it using the kernel level programming. Further, the protocol needs to be evaluated for the vulnerabilities and effectiveness of security features integrated into it.

8 Conclusion

To overcome security issues and other issues of TCP, many solutions were proposed in the literature, and some were implemented. But all those solutions run on top of TCP. None of the solutions offer the opportunity to replace TCP. Therefore, those solutions fail to provide a comprehensive protocol to replace TCP. This research's main technical contribution was finding an effective security solution for vulnerabilities of TCP. This framework identifies the major aspects to consider when developing a protocol for the transport layer. The proposed framework and the protocol could be used to develop a fully operational secure transport layer protocol in the future, expert feedback and the results from the evaluations which compare ATLP against the framework, TCP, and existing protocols confirm this outcome.

REFERENCES

[1] J. Postel, "RFC793: Transmission Control Protocol - DARPA Internet Program Protocol Specification," Internet Engineering Task Force (IETF) Request for Comments 793, 1981. https://www.rfc-editor.org/rfc/rfc793.html (accessed Aug. 07, 2022).

[2] W. Eddy, "RFC 9293: Transmission Control Protocol (TCP)," Internet Engineering Task Force (IETF) Request for Comments 9293, 2022. https://www.rfc-editor.org/rfc/rfc9293.pdf

[3] A. Faisal and M. Zulkernine, “A secure architecture for TCP/UDP-based cloud communications,” in International Journal of Information Security, Apr. 2021, vol. 20, no. 2, pp. 161–179. doi: 10.1007/s10207-020-00511-w.
[4] B. Harris and R. Hunt, “TCP/IP security threats and attack methods,” Computer Communications, vol. 22, no. 10, pp. 885–897, 1999, doi: 10.1016/S0140-3664(99)00064-X.
[5] G. Papastergiou et al., “De-Ossifying the Internet Transport Layer: A Survey and Future Perspectives,” IEEE Communications Surveys and Tutorials, vol. 19, no. 1, pp. 619–639, 2017, doi: 10.1109/COMST.2016.2626780.
[6] Benj Edwards, “The Foundation of the Internet: TCP/IP Turns 40,” https://www.howtogeek.com/, 2021. https://www.howtogeek.com/751880/the-foundation-of-the-internet-tcpip-turns-40/ (accessed Feb. 06, 2023).
[7] S. D. Nguyen, M. Mimura, and H. Tanaka, “Abusing TCP Retransmission for DoS Attack Inside Virtual Network,” in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 10763 LNCS, Springer Verlag, 2018, pp. 199–211. doi: 10.1007/978-3-319-93563-8_17.
[8] A. M. Ali, “Using the TCP Retransmission Timeout,” 2018, no. 1.
[9] W. Du, “Attacks on the TCP Protocol,” Computer Security: A Hands-on Approach, pp. 42–52, 2017.
[10] S. M. Bellovin, “Security Problems in the TCP/IP Protocol Suite,” ACM SIGCOMM Computer Communication Review, vol. 19, no. 2, pp. 32–48, 1989.
[11] S. M. Bellovin, “A look back at ‘security problems in the TCP/IP protocol suite,’” Proceedings - Annual Computer Security Applications Conference, ACSAC, pp. 229–249, 2004, doi: 10.1109/csac.2004.3.
[12] M. Polese, F. Chiariotti, E. Bonetto, F. Rigotto, A. Zanella, and M. Zorzi, “A Survey on Recent Advances in Transport Layer Protocols,” IEEE Communications Surveys and Tutorials, vol. 21, no. 4, pp. 3584–3608, Oct. 2019, doi: 10.1109/COMST.2019.2932905.
[13] A. Abdelsalam, M. Luglio, C. Roseti, and F. Zampognaro, “TCP Wave: A new reliable transport approach for future internet,” Computer Networks, vol. 112, pp. 122–143, 2017, doi: 10.1016/j.comnet.2016.11.002.
[14] A. Abdelsalam, M. Luglio, N. Patriciello, C. Roseti, and F. Zampognaro, “TCP Wave over Linux: A disruptive alternative to the traditional TCP window approach,” Computer Networks, vol. 184, no. October 2020, 2020, doi: 10.1016/j.comnet.2020.107633.
[15] Rescorla E, “Writing Protocol Models,” Internet Engineering Task Force (IETF) Request for Comments 4101, 2005. https://www.rfc-editor.org/rfc/pdfrfc/rfc4101.txt.pdf
[16] F. Gont and S. Bellovin, “RFC 6528: Defending against Sequence Number Attacks,” Internet Engineering Task Force (IETF) Request for Comments 6528, 2012. https://www.rfc-editor.org/rfc/pdfrfc/rfc6528.txt.pdf
[17] L. L. Peterson and B. S. Davie, Computer Networks: A Systems Approach Fifth Edition. 2012. doi: 10.1016/b978-0-12-385059-1.00014-4.
[18] Canonical Ltd., “Ubuntu Manpage: netwox - examples/tools of the network library netwib,” Ubuntu Manpage Repository, 2019. https://manpages.ubuntu.com/manpages/focal/man1/netwox.1.html (accessed Mar. 14, 2023).
[19] Z. Kanmai, “TCP/IP protocol security problems and defenses,” Proceedings - 2020 International Conference on Intelligent Computing and Human-Computer Interaction, ICHCI 2020, pp. 117–120, 2020, doi: 10.1109/ICHCI51889.2020.00033.
[20] M. Nieles, K. Dempsey, and V. Y. Pillitteri, “An introduction to information security evaluation,” NIST Special Publication 800-12 Revision 1, 2017, doi: 10.6028/NIST.SP.800-12r1.
[21] A. Faisal and M. Zulkernine, “A secure architecture for TCP/UDP-based cloud communications,” International Journal of Information Security, vol. 20, no. 2, pp. 161–179, 2021, doi: 10.1007/s10207-020-00511-w.
[22] J. Postel, “RFC 791: Internet Protocol,” Internet Engineering Task Force (IETF) Request for Comments 791, 1981. https://www.rfc-editor.org/rfc/rfc791 (accessed Oct. 30, 2022).
[23] S. Deering and R. Hinden, “RFC 1883: Internet Protocol, Version 6 (IPv6),” Internet Engineering Task Force (IETF) Request For Comments 1883, 1995. https://www.rfc-editor.org/rfc/rfc1883

iCIIT 2023

THE 4TH INTERNATIONAL CONFERENCE ON INNOVATIONS IN INFO-BUSINESS & TECHNOLOGY

Conference Track

TECHNOLOGY OF BUSINESS



AlzhCare: A cognitive therapy application for Alzheimer's Patients with smart results, and a safety tool to support their Caregivers

Jesica Rayen
Student, Department of Business School Informatics Institute of Technology, Colombo, Sri Lanka, jesica.2018812@iit.ac.lk

Namal Malalasena
Lecturer, Department of Business School Informatics Institute of Technology, Colombo, Sri Lanka, namal.m@iit.ac.lk

ABSTRACT

Alzheimer's disease is a progressive neurological disorder characterized by memory loss and cognitive decline. Memory loss, confusion, issues with language communication, behavioral changes and mood swings are common symptoms of the disease. As the leading cause of dementia in adults, it poses significant challenges for caregivers, who face difficulties in providing care due to the mental impairments of the patients and their own busy lifestyles. Additionally, medical professionals strive to slow down the disease's progression through cognitive enhancement therapies. This research project aims to analyze the obstacles faced by caregivers and provide a viable solution to support stakeholders in delivering optimal care, enhancing cognitive abilities, and ensuring patient safety.

The research methodology included a comprehensive literature review and interviews with domain experts such as psychologists, psychiatrists, and caregivers themselves. The findings revealed that caregivers experience high levels of stress, encounter various difficulties, and face safety concerns while managing and caring for Alzheimer's patients. Tracking the patient's current mental status is crucial due to the gradual loss of cognitive abilities over time. Moreover, maintaining the patient's cognition at a standard level necessitates continuous support and cognitive enhancement therapies under the guidance of medical professionals.

To address these challenges, AlzhCare introduces a mobile application designed to assist medical professionals and caregivers. The application enables them to track assigned cognitive stimulation therapies and monitor the patient's memory level as needed. Additionally, it provides a safety tool for caregivers, allowing them to monitor the patient's location, heartbeat, blood pressure, and oxygen levels, ensuring vigilant care and enhancing patient safety.

Keywords
AlzhCare, Alzheimer's disease, caregivers, safety tool, cognitive enhancement therapies, medical professionals, memory level, challenges.

1 Introduction:

Alzheimer's disease, a neurological condition, poses significant obstacles for patients, resulting in cognitive impairment, memory loss, and difficulties in performing everyday tasks. [20] The burden of care falls heavily on the shoulders of dedicated caregivers, who tirelessly support and assist their loved ones throughout the disease's progression.

In addition, medical professionals encounter unique challenges in providing optimal care to Alzheimer's patients. Tailoring cognitive therapies to suit individual needs, monitoring vital signs and health parameters in real-time, and ensuring the safety of patients during potential wandering episodes are just some of the intricate demands faced by medical practitioners.

To deeply explore and address these critical issues, the research adopts a rigorous approach. It involves a meticulous review of existing literature on Alzheimer's care, caregiving challenges, and advancements in cognitive therapy. Furthermore, the study conducts insightful interviews with experienced caregivers and doctors, obtaining first-hand perspectives and insights into the complexities of caring for Alzheimer's patients.

By utilizing both primary research that is interviews with caregivers and medical professionals, and secondary research methods, such as analyzing reputable scholarly articles, this research article seeks to provide a comprehensive understanding of the multifaceted challenges affecting all stakeholders involved in Alzheimer's care. By identifying and analyzing these challenges, the study endeavors to develop strategic and innovative technological interventions, ultimately striving to improve the overall quality of care and support for individuals grappling with Alzheimer's disease. Through this collective effort, we aim to make a significant impact in the lives of Alzheimer's patients, caregivers, and medical professionals, enhancing their well-being and enhancing the effectiveness of Alzheimer's care and support services.

2 Problem Statement

Alzheimer's disease is a progressive neurological disorder that affects millions of individuals worldwide. [7][29] As the most common cause of dementia in adults, it poses significant challenges for both patients and their caregivers. The disease's hallmark symptoms include memory loss, confusion, communication difficulties, behavioral changes, and mood swings. [19] Currently, there is no known cure for Alzheimer's disease, necessitating a focus on managing symptoms, maintaining cognitive function, and ensuring patient safety. [29]

The early signs of the disease include forgetting recent events, occasions, conversations. As it progresses, it will develop severe memory impairment and fail the ability to carry out day to day tasks. [16] Alzheimer's disease is not a normal part of aging; it is the result of complex changes in the brain that begin way before the symptoms show up and lead to the loss of brain cells/neurons and their connections. [20]

Dementia affects more than 50 million people around the world, with a new case of 3 every second in the world.[6] The number of people affected by dementia is expected to increase significantly in the coming years. By 2030, it is estimated that the global prevalence will reach 82 million, and by 2050, it is projected to almost triple to 152 million. [4] Here is the graphical illustration of global projections of dementia prevalence by Geographic region by Alzheimer's research UK.[12]

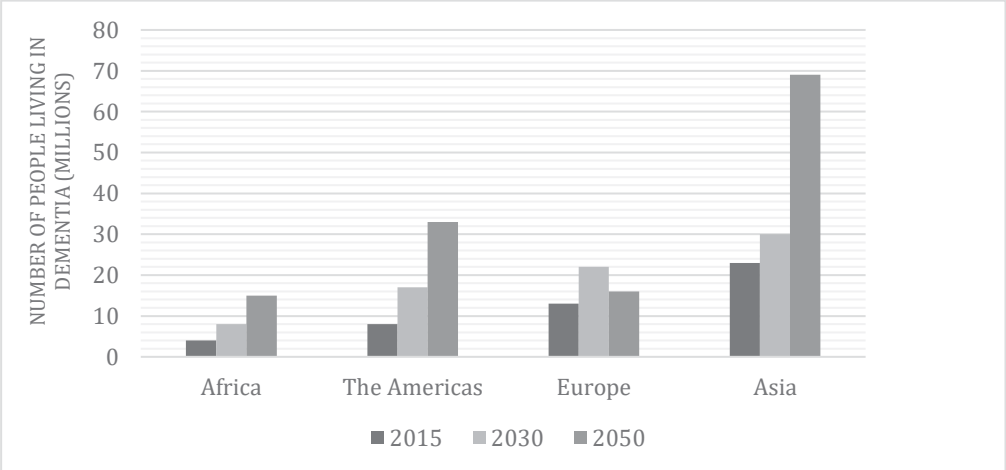


Figure 1: Global Projections Of Dementia Prevalence By Geographic Region, Alzheimer's Research UK

According to [27] the disease progresses in stages, with the severity of symptoms, 1. Asymptomatic 2. Mild Cognitive Impairment (MCI) 3. Mild Dementia 4. Moderate Dementia 5. Severe Dementia. The early stage of Alzheimer's is when experiencing memory loss, cognitive shortcomings, in which the stage where the disease is often diagnosed. MCI is when more memory problems faced than their age, which has high risk of developing Alzheimer's Disease where a regular consultation of doctor or specialist is required. Mild dementia affects the areas of brain where the language control, reasoning, sensory procession and conscious(ness). There is more confusion in recognizing family and friends. Severe dementia is where the person cannot communicate and completely dependent on others for care. [20] While medical treatments play a vital role in managing Alzheimer's disease at each stage, it is crucial to acknowledge the significance of cognitive treatments in improving the quality of life for patients . [22] However, research shows that approximately 55% of caregivers encounter challenges in diagnosing, evaluating, and monitoring the memory levels of patients [6]. One contributing factor is the patients' reluctance to frequent hospital visits for treatment, as revealed through interviews with caregivers of Dementia patients in Western Province. Additionally, interviews with doctors who are handling the Dementia patients in Sri Lanka indicated that 65% of them struggle with assigning, tracking, and conducting cognitive brain exercises aimed at delaying memory loss. They express a need for an application that can facilitate these activities.

Cognitive stimulation therapy helps the patients to use the parts of the brain that are working to help the ones which are not working at the early stages of dementia. Through the interviews conducted and research done with the Doctors, patients can be supported with an application to do the cognitive therapies at the first three stages. However, interviews with caregivers indicate that many do not prioritize cognitive stimulation therapies due to a lack of attention and time, resulting in the patient's cognitive decline. During a visit to the Lanka Alzheimer's Foundation and an interview conducted with the office manager, it was observed that patients engage in brain exercises during their visits to the center, which has shown positive effects on their mood. However, it was noted that due to limited resources, scheduling constraints and to give opportunities to others, patients can only visit the center once a week. This visit was found to provide significant relief to their caregivers too.

One notable consequence of the brain changes caused by Alzheimer's disease is the impact on patients' moods and behaviors. These changes may manifest as depression, apathy, mood swings, irritability, aggressiveness, wandering, distrust in others, and delusions [19]. Consequently, caregivers must exercise constant vigilance and supervision to ensure the safety and well-being of the patients. Interviews conducted with Dementia doctors and caregivers have emphasized the crucial need for attentive care and supervision due to the patients' unpredictable mood swings and behavioral fluctuations. The strain and challenges associated with managing these behaviors can lead to caregiver distress and mental health problems.

Additionally, it was evident that the behavioral changes exhibited by individuals with Alzheimer's disease can also increase the risk of accidents and instances of wandering. Caregivers must be proactive in preventing accidents and ensuring the safety of the patients. Wandering behavior poses a significant concern as it can result in individuals becoming disoriented, lost, and exposed to potential dangers.

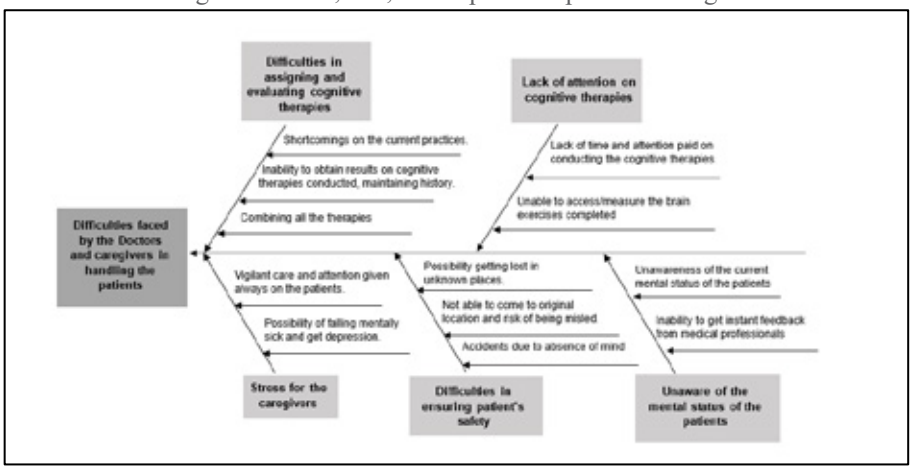


Figure 2: Fishbone Diagram categorising the potential causes of the problem domain of this research

Review of existing literature on Alzheimer's care, caregiver challenges, and cognitive therapy interventions. Additionally, interviews with caregivers and medical professionals will provide first-hand insights into the complexities of caring for Alzheimer's patients. Through this comprehensive research, the study aims to develop and validate the effectiveness of AlzhCare as a transformative solution, providing valuable data-driven insights and practical strategies to improve Alzheimer's care and support services.

Ultimately, this research endeavors to contribute to the advancement of Alzheimer's care practices, alleviate caregiver burdens, and enhance patient well-being, making a meaningful impact in the lives of those affected by this debilitating disease.

2.1 Issues faced by caregivers and medical professionals when handling the patients

Alzheimer's disease is a chronic and progressive neurodegenerative disorder characterized by the deterioration of cognitive and functional abilities. Caregivers and medical professionals involved in the care of individuals with Alzheimer's disease encounter numerous challenges throughout the different stages of the disease. Each stage of the disease presents specific symptoms and manifestations, including but not limited to memory loss, impaired judgement, confusion, repetitive questioning, wandering, getting lost, impulsive behavior, and a diminished capacity for empathy [20]

2.1.1 Caregiver Issues

Caring for individuals with Alzheimer's disease poses significant challenges for caregivers, including the patients' inability to perform day-to-day activities, memory impairments affecting family bonds, and the need for specialized attention and care as the disease progresses. In the earlier stages, patients may require assistance with basic activities of daily living, while those in advanced stages of dementia necessitate constant supervision and care. The demanding nature of caregiving places considerable strain on caregivers, leading to stress, social isolation, and stigma “Almost half (49%) of paid health and long-term care professionals said they felt stressed or under pressure some of the time, while 37% said they felt stressed often or all of the time.” [16] Furthermore, caregivers must be attentive to the patients' heightened risks of falls, wandering, absentmindedness, mood swings, and behavioral changes, which can contribute to their own mental health problems, including depression, anxiety, impatience, and emotional exhaustion. This research examines the challenges faced by caregivers and the impact on their mental well-being when providing care for individuals with Alzheimer's disease.

2.1.2 Medical professional Issues

Diagnosing and caring for individuals with Alzheimer's Disease presents several challenges for medical professionals. The early-stage symptoms of the disease can be subtle and easily overlooked, making diagnosis a complex task. Effective communication with Alzheimer's patients is another significant challenge, requiring caregivers to demonstrate patience, empathy, and specialized skills [28]. Additionally, limited access to patients' information and a lack of remote monitoring capabilities hinders the ability to track patients' conditions and provide coordinated care. Assigning brain exercises to patients on a frequent basis, which supports their cognitive state, is also one of the challenges faced. Furthermore, ensuring the safety and security of Alzheimer's patients is a primary concern highlighted during interviews with medical professionals.

2.1.3 Risks in handling the patients

As Alzheimer's Disease progresses through its stages, the care and attention required for patients becomes increasingly crucial. With worsening symptoms, patients may experience behavioral changes, mood swings, and a loss of independence. Below are a few risks identified when handling Alzheimer's Disease patients and emphasize the need for high attention and vigilance as the disease progresses.

- As the disease advances, patients may face challenges with mobility and balance, increasing the risk of falls and injuries. The absence of mind, a common symptom of Alzheimer's, can contribute to accidents such as gas leaks posing real-life dangers shared during the interview conducted.
- Exhibit difficult behavior such as agitation, aggression and wandering which may cause the risk of getting lost, physical harm or injury, frustration, and potential behavioral outbursts.
- Handling patients with Alzheimer's Disease can also lead to a risk for mental healthcare of caregivers as they can be emotionally demanding, leading to stress, anxiety, and depression.
- Alzheimer's patients may be at an increased risk of developing and transmitting infectious diseases. This susceptibility can be attributed to factors such as compromised immune function, impaired judgment leading to inadequate hygiene practices, and difficulties in communicating symptoms.

3 Literature review

Alzheimer's disease is the leading cause of dementia, affecting memory, thinking, and behavior, leading to a progressive decline in cognitive and social skills. [20] It is a neurodegenerative disorder characterized by the accumulation of beta-amyloid plaques and neurofibrillary tangles, disrupting brain cell communication, and causing memory loss and cognitive decline. [7][29] Symptoms start with mild memory loss and progress to severe impairment, resulting in loss of independence. Risk factors include age, family history, genetics, lifestyle, and environmental factors. Globally, there are around 50 million AD patients, and this number is expected to double every five years, reaching 152 million by 2050. [5] Currently, there is no cure for Alzheimer's, and available treatments only aim to manage symptoms without complete recovery. [29]

Cognitive Therapy and Alzheimer's Disease: Numerous studies have emphasized the role of cognitive therapy in managing cognitive decline and memory impairment in Alzheimer's patients. Cognitive stimulation therapies (CST) have been shown to enhance cognitive processing, improve memory, and slow down cognitive decline [15]. Other research has demonstrated that regular cognitive exercises and brain training can positively influence cognitive function and promote better cognitive outcomes [18].

Importance of Early Diagnosis: Research has emphasized the significance of early diagnosis in Alzheimer's disease, as it allows for timely interventions and treatment planning. Diagnostic tools like the Mini-Mental State Examination (MMSE) and Montreal Cognitive Assessment (MoCA) have proven effective in identifying cognitive impairments and determining the disease's progression [23] Early detection can improve the effectiveness of cognitive therapies and enhance patients' quality of life.

Challenges Faced by Caregivers and Medical Professionals: Studies have highlighted the challenges faced by caregivers and medical professionals in providing care for Alzheimer's patients. Caregivers often experience high levels of stress, emotional exhaustion, and social isolation due to the demanding nature of caregiving [8]. Medical professionals may struggle with accurate diagnosis, communication with patients, and limited access to patient information [13]. AlzhCare, with its remote access and cognitive therapy features, could potentially alleviate some of these challenges.

Role of Technology in Alzheimer's Care: Research has shown that technology can play a vital role in improving Alzheimer's care. Mobile applications like AlzhCare have the potential to provide personalized cognitive therapies, enhance patient engagement, and improve cognitive function [10]. Online platforms can also offer flexibility and convenience for caregivers and patients, allowing them to access cognitive therapies remotely [28].

Overall, the literature review highlights the significance of cognitive therapy in Alzheimer's care and the potential benefits of AlzhCare in addressing the challenges faced by patients, caregivers, and medical professionals. The studies and research support the notion that AlzhCare can be a valuable tool in improving the quality of life for individuals with Alzheimer's disease and enhancing their overall well-being.

4 Research Methodology

4.1 Research Design

The research has adopted a mixed-methods approach, combining both qualitative and quantitative data collection methods. The study aims to explore the efficacy of AlzhCare, a cognitive therapy application, in improving care and support for Alzheimer's patients. The research involves in-depth interviews with medical professionals and caregivers, as well as a literature survey to gather existing research related to Alzheimer's care, cognitive therapies, and the use of technology in healthcare.

4.2 Sampling

A purposive sampling method will be employed to select participants for the interviews. Medical professionals, including neurologists, psychiatrists, geriatric specialists, and general practitioners, who have experience in treating Alzheimer's patients and are familiar with cognitive therapies, were included. Caregivers, both family members and professional caregivers, who are actively involved in providing care for individuals with Alzheimer's disease were also included in the study.

4.3 Data Collection

a. Interviews: Semi-structured interviews will be conducted with the participants to gather in-depth insights into their experiences with Alzheimer's care and the use of AlzhCare. The interview questions will cover topics such as the challenges faced in caregiving, the importance of cognitive therapies, the role of technology in supporting patient care, and the perceived benefits and limitations of using AlzhCare.

b. Literature Survey: A comprehensive literature survey will be conducted to identify and review existing research related to Alzheimer's care, cognitive stimulation therapies, and the use of technology-based interventions for patients with dementia. Academic databases, research journals, and reputable sources will be utilized to collect relevant literature for analysis.

Through this comprehensive research methodology, the study aims to provide a comprehensive understanding of the effectiveness of AlzhCare as a cognitive therapy application and its potential to improve the quality of care and support for Alzheimer's patients, as perceived by medical professionals and caregivers. The insights obtained from the interviews and literature survey will contribute valuable knowledge to the field of Alzheimer's care and inform the development of targeted interventions for this challenging disease.

5 About AlzhCare

Alzheimer's Disease poses significant challenges for medical professionals and caregivers in providing effective care and ensuring the safety of patients. In this context, the development of innovative technologies can play a crucial role in enhancing care delivery and improving patient outcomes. AlzhCare is a mobile application with the support of a wearable smart device to address these challenges by providing a comprehensive platform for cognitive stimulation therapies and safety monitoring for Alzheimer's Disease patients.

Cognitive Stimulation and Therapy Tracking: AlzhCare offers a user-friendly interface for medical professionals to assign cognitive stimulation therapies to patients based on their specific memory levels. By utilizing the application, medical professionals can easily monitor the progress of patients and evaluate their cognitive performance through generated reports and analytics. This feature enables medical professionals to provide valuable feedback and guidance to caregiver's real time, facilitating optimized care plans tailored to the individual needs of each patient.

Safety Monitoring and Caregiver Support: Recognizing the safety concerns associated with Alzheimer's Disease, AlzhCare incorporates advanced monitoring capabilities to enhance patient safety and provide support to caregivers. The application allows caregivers to track the location of patients with the help of a wearable smart device in case they become lost or disoriented, minimizing the risk of accidents or wandering-related incidents. Additionally, caregivers can conduct regular check-ups on vital signs such as heartbeat, blood pressure, and oxygen levels. Moreover, AlzhCare provides real-time monitoring of gas and smoke levels around the patients, ensuring a safe environment for their well-being using the application and through the smart device.

5.1 Originality of AlzhCare

The AlzhCare mobile application exhibits originality through its unique and comprehensive set of features that cater to the specific needs of medical professionals, caregivers, and Alzheimer's patients.

- Role-Based Registration: The landing page's role-based registration allows users (doctors, patients, and caregivers) to create profiles specific to their roles. This personalized approach ensures that each user has access to the relevant functionalities and information, streamlining their experience with the application.
- Caregiver Assignment: The feature that enables doctors to assign caregivers to specific patients is innovative as it facilitates efficient care coordination. This ensures that patients receive dedicated and personalized support, leading to better patient outcomes and caregiver satisfaction.
- Customized Brain Exercises: The ability of doctors to assign tailored brain exercises to patients addresses the individual cognitive needs of each patient. This feature allows for a personalized approach to cognitive stimulation, optimizing the effectiveness of the exercises.
- Real-Time Feedback and Scoring: The application's capability to provide real-time feedback and scoring for completed brain exercises enhances patient engagement and motivation. Patients receive immediate insights into their performance, enabling them to track progress and work towards cognitive improvement.
- Scoreboard and Analytics: The inclusion of a scoreboard and analytics section for doctors and caregivers to view patient scores on brain exercises allows for data-driven decision-making. This data analysis contributes to a more comprehensive understanding of patient cognitive abilities and progress.
- Location Tracking and Safety Alerts: The incorporation of location tracking and safety alerts sets AlzhCare apart from conventional caregiving solutions. Caregivers can monitor patients' locations in real-time, enhancing patient safety and reducing the risk of wandering-related incidents.
- Health Monitoring: AlzhCare's feature that enables the monitoring of patients' vital signs, such as heartbeat, blood pressure, and oxygen levels, is a novel addition. This health monitoring capability empowers caregivers to promptly address any health concerns and provide timely medical assistance.
- Smart Device Integration: The application is capable to integrate with smart devices to monitor gas/smoke levels and other environmental factors further ensures patient safety and wellbeing. This innovative use of smart technology adds a layer of safety and security to the caregiving process.

5.2 Technologies used

The development of the AlzhCare application incorporates a robust technological stack to ensure efficient and seamless functionality. The front-end development is powered by Flutter, an open-source UI software development kit renowned for its ability to create cross-platform mobile applications using a single codebase. Flutter's framework enables the construction of responsive user interfaces, delivering high-speed performance and a smooth user experience. For the backend development, Node.js and Express.js languages are utilized. Node.js allows the execution of JavaScript on the server-side, facilitating efficient server operations. Express.js, on the other hand, provides a flexible and reliable framework for constructing robust backend applications, ensuring the stability and responsiveness of the application's backend functionalities.

The application's data storage and management are facilitated by MongoDB, a popular NoSQL database. MongoDB offers a flexible and scalable solution for document-oriented data storage. Leveraging MongoDB's capabilities, the application benefits from features such as high availability, automatic backups, and seamless deployment. To enhance the security of user accounts, the AlzhCare application incorporates Two-Factor Authentication. This security measure adds an extra layer of protection by requiring users to enter a code received via email before gaining access to their accounts, ensuring the confidentiality and integrity of user data.

Firebase Cloud Messaging is integrated into the application as a cross-platform messaging solution. It enables the delivery of alerts to caregivers regarding location updates, care indicators, and brain exercise reminders. This functionality enhances communication and facilitates timely updates for caregivers, ensuring they stay informed about the well-being and cognitive stimulation of the patients. Sophisticated algorithms are employed within the application to create customized brain exercises, accurately assess cognitive abilities, and calculate overall scores. These algorithms provide valuable analytics and insights for medical professionals and caregivers, enabling data-driven decisions regarding patients' cognitive performance in areas such as language, memory, and attention.

5.3 Uniqueness of the application

To promote cognitive function improvement, the application offers a range of brain exercises accompanied by colorful images and engaging user interfaces. These elements captivate patients' attention and foster their interest in participating in exercises, thereby enhancing their cognitive abilities.

All three user roles - doctors, patients, and caregivers - are seamlessly integrated within the same application platform.

Instant Feedback from the medical professionals on the brain exercises completed by the patients, which helps caregivers to have proper track on the memory levels of the patients and have instant connection with medical professionals in case of emergency. Care tips and guidance for the caregivers to handle the patients can be shared through the same by the medical professionals.

Tracking the location of the patients by the caregivers remotely, as patients have wandering behaviours which increases the chances of getting lost. This feature will help the caregivers to ensure the safety of the patients if they get lost. Further alerts to notify caregivers if patients venture beyond predetermined boundaries.

Patients can pin and save their native location in their profile, providing them with a means to safely return home if they become disoriented or lost.

The application employs care indicator alerts to inform caregivers of any abnormal fluctuations in the patients' vital signs, such as heartbeat, oxygen level, and blood pressure. These alerts enable caregivers to exercise extra caution and promptly address any concerning health changes.

To ensure the safety of patients from potential gas leakage accidents, the application monitors gas/smoke levels in their surroundings. If the levels exceed predefined limits, the application sends out alerts to caregivers, allowing them to take immediate action and mitigate potential risks.

5.4 Evaluation and Expert Validation

AlzhCare has undergone an evaluation process involving domain experts and industry professionals to ensure its effectiveness and usability in real-life caregiving scenarios. The expertise of these professionals has contributed to refining the application's features, ensuring that it addresses the genuine challenges faced by medical professionals, caregivers, and patients in Alzheimer's Disease care.

6 Project Methodology

The initial requirements for the project were identified through research. However, to ensure alignment with stakeholder expectations and to accommodate changes in the business domain, interviews were conducted to gather feedback and insights. This iterative process facilitated the refinement and alteration of requirements throughout the development lifecycle. The time constraints associated with the project necessitated an efficient and effective development approach. Agile methodology, particularly the Scrum approach, was well-suited to address this challenge. It emphasized the delivery of functional software, which was regularly reviewed and tested by stakeholders to ensure their expectations were met.[27][26]

The Scrum project management approach promoted transparency, with regular meetings providing updates on project goals and progress to the supervisor. This facilitated continuous feedback, allowing for adjustments and improvements to the project plan, thereby enhancing the efficiency and effectiveness of the project management process. Additionally, the Scrum methodology enabled the author to deliver the expected outcomes of stakeholders promptly and respond to changing demands. Its iterative and incremental nature allowed for the delivery of working application components in short intervals, enabling quick feedback and continuous improvement.

To ensure time efficiency and cost-effectiveness, the Lean Software Development methodology [24] was incorporated into the development approach. This methodology aimed to eliminate waste, optimize flow, and amplify learning. By focusing on delivering a high-quality mobile application quickly and efficiently, the author prioritized key features valued by customers within the limited time frame. The iterative nature of Lean Software Development aligned well with the application's development process. Delivering software in small increments enabled the author to receive feedback and guidance from the supervisor, ensuring the inclusion of the right features and functionalities while facilitating continuous learning and meeting stakeholder needs.

The development methodology for the AlzhCare application followed an Agile approach, specifically utilizing the Scrum project management approach and incorporating elements of Lean Software Development. This iterative and incremental methodology allowed for flexibility, adaptability, and the efficient delivery of a functional and high-quality mobile application while ensuring stakeholder expectations were met.

7 Findings and Discussions

7.1 Findings

Cognitive Stimulation Tracking: The AlzhCare mobile application effectively tracks and manages cognitive stimulation therapies assigned to Alzheimer's patients. Medical professionals and caregivers can easily schedule and monitor the therapy sessions, ensuring that patients receive the necessary mental stimulation regularly. This feature has shown promising results in enhancing cognitive function and memory retention in patients, as consistent engagement in cognitive activities has been linked to improved cognitive abilities.

Memory Level Monitoring: The application's ability to monitor the patient's memory level in real-time is an asset for medical professionals and caregivers. Through cognitive assessment tools integrated into the app, caregivers can gain insights into the patient's cognitive status and detect any cognitive decline promptly. This feature has enabled early interventions and adjustments to therapies based on the patient's cognitive progression.

Safety Monitoring: The safety tool integrated into the AlzhCare application has proven to be a game-changer for caregivers. By tracking the patient's location and vital signs, such as heartbeat, blood pressure, and oxygen levels, caregivers can ensure the patient's safety and well-being even from a distance. This feature has significantly reduced the risk of accidents, or wandering behavior in Alzheimer's patients, providing peace of mind to both caregivers and families.

7.2 Discussion

The AlzhCare mobile application has the potential to revolutionize Alzheimer's care by empowering medical professionals and caregivers with efficient and effective tools. The findings indicate that the application's cognitive stimulation tracking feature has a positive impact on the cognitive abilities of patients. Regular engagement in cognitive exercises and activities can slow down cognitive decline, improve memory, and enhance overall cognitive function.

Moreover, the real-time monitoring of the patient's memory level has proven invaluable in providing personalized care and interventions. Identifying early signs of cognitive decline allows medical professionals to adjust treatment plans and tailor therapies to address specific cognitive challenges, leading to better outcomes for patients.

The safety monitoring aspect of the application has greatly benefited caregivers by providing them with vital information about the patient's well-being. This real-time tracking allows caregivers to respond promptly to any medical emergencies or wandering incidents, preventing potential harm to the patient, and easing the burden of caregiving.

8 Conclusion

The study has shown the major difficulties faced by Alzheimer's patients, medical professionals, and caregivers due to the cognitive impairment of the patients while treating and taking care of them. This research paper has demonstrated that AlzhCare is built to assist the Alzheimer's patients and caregivers in its best possible way. Through AlzhCare the above stake holders can benefit from cognitive stimulation, remote monitoring, safety and security, social engagement, and a safety tool to support the caregivers. The technologies built in AlzhCare have the potential to empower the patients, promote their independence, enable them to live fulling lives and it also reduce the burden and stress faced by the caregivers while taking care of their loved ones and ensuring their safety.

As future work, the research and development can be made in mood predictions by camera. AI using machine learning algorithms and computer vision techniques to analyse facial expressions and emotional status of the patients to be extra precautionary on handling the patients. Advance video and audio-based brain stimulation exercises using virtual reality augmented reality, and gamification. Further Machine Learning models for building advanced predictive analytics and insights on cognitive abilities of patients which can provide personalized exercise recommendations. Natural Language processing AI to translate the application will immensely help the old aged to understand and use the application easily.

Technology has the potential to transform Alzheimer's care, providing a ray of hope for individuals affected by the diseases and their families. By harnessing the power of technology, substantial strides can be made towards improving the lives of those with Alzheimer's disease by providing comprehensive, personalized, and dignified care.

9 Acknowledgements

The author of this research paper acknowledges the help, support and direction given by many people including domain specialists and experts, also special thanks to Mr. Namal Malasena and all the lecturers who gave their guidance.

10 References

[1] Age Space. "Treatments Used for Dementia," Age Space, Available: https://www.agespace.org/dementia/treatments-used-dementia.

[2] Alzheimer's Association. "What Is Alzheimer's?," Alzheimer's Association, Available: https://www.alz.org/alzheimers-dementia/what-is-alzheimers.

[3] Alzheimer's Association. "Alzheimer's & Dementia Treatments," Alzheimer's Association, Available: https://www.alz.org/alzheimers-dementia/treatments.

[4] Alzheimer's Disease International. "World Alzheimer Report 2015," Available: https://www.alzint.org/u/WorldAlzheimerReport2015.pdf.

[5] Alzheimer's Disease International. (n.d.). Dementia Statistics. Retrieved from https://www.alzint.org/about/dementia-facts-figures/dementia-statistics/\

[6] Alzheimer's Disease International. "Dementia Statistics," Alzheimer's Disease International, Available: https://www.alzint.org/about/dementia-facts-figures/dementia-statistics/.

[7] "Alzheimer's Disease," NHS inform, Available: https://www.nhsinform.scot/illnesses-and-conditions/brain-nerves-and-spinal-cord/alzheimers-disease#:~:text=Alzheimer's%20disease%20is%20a%20life,can%20cause%20problems%20with%20swallowing.

[8]Alzheimer's Disease International. "World Alzheimer Report 2022," Available: https://www.alzint.org/resource/world-alzheimer-report-2022.

[9] Apostolova LG, Thompson PM, Rogers SA, et al. "Subregional hippocampal atrophy predicts Alzheimer's dementia in the cognitively normal." Neurobiol Aging. 2006 Dec;27(12): 1678-88. DOI: 10.1016/j.neurobiolaging.2005.09.017. Available: https://pubmed.ncbi.nlm.nih.gov/12399581/.

[10] [Author(s)]. (2022, December). Comprehensive Review on Alzheimer's Disease: Causes and Treatment. Molecules, [Volume Number](Issue Number), [Page Numbers]. DOI or URL if available.

[11] Cowen, P., Harrison, P., & Burns, T. (2012). Shorter Oxford Textbook of Psychiatry. DOI: 10.1093/med/9780199605613.001.0001.

[12] Dementia Statistics. "Prevalence and Incidence of Dementia," Dementia Statistics, Available: https://dementiastatistics.org/about-dementia/prevalence-and-incidence/.

[13] Hofmann, [First Author Initials]., [Second Author Initials]., [Third Author Initials]., & [Last Author Initials]. (2021). [Title of the Paper]. [Journal Name], [Volume Number](Issue Number), [Page Numbers]. DOI or URL if available.

[14] J. Li, X. Yuan, X. Tang, and M. Wang, "A Comprehensive Review of Web Personalization Technologies in Advanced Recommender Systems," IEEE Access, vol. 6, pp. 40128-40143, 2018. DOI: 10.1109/ACCESS.2018.2854122. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6140014/.

[15] Loraine, [First Author Initials]., [Second Author Initials]., [Third Author Initials]., & [Last Author Initials]. (2014). [Title of the Paper]. [Journal Name], [Volume Number](Issue Number), [Page Numbers]. DOI or URL if available.

[16] Mayo Clinic, "Alzheimer's Disease: Symptoms & Causes," Mayo Clinic, Available: https://www.mayoclinic.org/diseases-conditions/alzheimersdisease/symptoms-causes/syc-20350447.

[17] MDPI. "Biochemistry Section - Special Issue: Bioactive Phytochemicals and Functional Food Components in Health and Disease," Molecules, vol. 25, no. 24, article 5789, 2020. Available: https://www.mdpi.com/1420-3049/25/24/5789.

[18] Missouri State Medical Association (MSMA). (n.d.). Home. Retrieved from https://www.msma.org/

[19]National Institute on Aging. "Alzheimer's Diagnostic Guidelines Updated for the First Time in Decades," NIA News, Available: https://www.nia.nih.gov/news/alzheimers-diagnostic-guidelines-updated-first-time-decades.

[20] National Institute on Aging (NIA). "What Is Dementia?," NIA, Available: https://www.nia.nih.gov/health/what-is-dementia.

[21] National Institute on Aging (NIA). "What Causes Alzheimer's Disease?," NIA, Available: https://www.nia.nih.gov/health/what-causes-alzheimers-disease.

[22] NHS. "Dementia - Treatment," NHS, Available: https://www.nhs.uk/conditions/dementia/treatment/#:~:text=Acetylcholinesterase%20inhibitors&text=

Donepezil%20(also%20known%20as%20Aricept,treat%20more%20severe%20Alzheimer's%20disease.
[23] Physiopedia. (n.d.). Dementia. Retrieved from https://www.physiopedia.com/Dementia?utm_source=physiopedia&utm_medium=search&utm_campaign=ongoing_internal
[24] Roadmunk. "Types of Software Development Methodologies," Roadmunk, Available: https://roadmunk.com/guides/types-of-software-development-methodologies/.
[25] Soto, C.J., John, O.P., Gosling, S.D., & Potter, J. (2011). Age Differences in Personality Traits From 10 to 65: Big Five Domains and Facets in a Large Cross-Sectional Sample. Journal of Personality, 79(2), 323-345. DOI: 10.1111/jopy.12050. Available: https://onlinelibrary.wiley.com/doi/10.1111/jopy.12050.
[26] Synopsys. "Top 4 Software Development Methodologies," Synopsys, Available: https://www.synopsys.com/blogs/software-security/top-4-software-development-methodologies/.
[27] Teamwork. "Project Management Methodologies," Teamwork, Available: https://www.teamwork.com/project-management-guide/project-management-methodologies/.
[28] WebMD. "Dementia Treatments Overview," WebMD, Available: https://www.webmd.com/alzheimers/dementia-treatments-overview.
[29] World Health Organization. "Dementia," World Health Organization, Available: https://www.who.int/news-room/fact-sheets/detail/dementia.

The Impact of Electronic Word-of-Mouth on Customer Purchase Intention

Study on the Sri Lankan Fashion Industry

Vishaleni Sureshan†

School of Business, Informatics Institute of Technology, Wellawatte, Colombo, Sri Lanka, sureshan.20191041@iit.ac.lk

Shonali Aponso

School of Business, Informatics Institute of Technology, Wellawatte, Colombo, Sri Lanka, shonali.a@iit.ac.lk

ABSTRACT

The research was conducted with the intention of exploring the impact of electronic word-of-mouth on customer's purchasing intention in the Sri Lankan Fashion industry. A pilot study was conducted initially to confirm that a research problem existed, and was identified that the presence of e-WOM is weak due to lack of captivating marketing strategies by fashion brands of Sri Lanka. Then the study examined and identified the different dimensions of e-WOM which were argument quality, information credibility, review quantity and source credibility. The deductive approach and quantitative methodology were chosen in testing the relationship between each of these dimensions and its impact on customer purchase intention by building a conceptual framework and hypotheses. Then a survey was carried out, where 250 valid responses were collected and was run for a series of tests using the SPSS software to analyze the data.

The data analyzation led to the findings that all 4 variables (argument quality, information credibility, review quantity and source credibility) positively and significantly impacted customers' purchasing intention, and out of which "source credibility" was considered the most influential. The study then concluded with a discussion on the findings, limitation of the research study and by providing recommendation to fashion brands to improve their marketing strategies and encourage e-WOM.

KEYWORDS

Electronic word-of-mouth, customer purchase intention, fashion industry

1 Introduction

The evolution of internet and social media platforms has changed the lifestyle of individuals globally and have become a powerful tool for marketers to communicate and promote their products and services. Numerous corporations including fashion companies have now embraced Web 2.0 to inform, understand and connect with their consumers and build brand awareness (Bhat & Bhat, 2020). This has made electronic word-of-mouth and online shopping to evolve into a socio-economic activity, providing a platform for consumers to voice out their experiences (Wang et al., 2015). e-WOM is crucial for fashion brands and if utilized properly, will lead to increase in brand credibility, sales, reach and engagement.

The Fashion industry of Sri Lanka is a growing sector and has been affected negatively in the last few years by Covid and economic crisis, tasked with a responsibility to recover quickly (Wattegama, 2019). Due to this, digital marketing spending in the recent years by the fashion brands had become inconsistent and hindered the development of effective marketing strategies (Statista, 2022), resulting in Sri Lankan Fashion industry to be perceived to have weak digital word-of-mouth presence.

The research endeavors with the research question "What is the impact of e-WOM on the customer purchasing intentions in the Sri Lankan fashion industry? with the formulation of the following objectives to guide the study, 1) to identify the dimensions of e-WOM that influences the customer purchase intention of

fashion products, 2) to analyze the relationship of the identified e-WOM dimensions on the customer purchase intention, 3) to provide recommendation for fashion brands to formulate digital marketing strategies that are effective in creating engagement among customers.

A wealth of research has been carried out locally, observing the impact of e-WOM on purchase intentions in different industries of Sri Lanka such as the retail, cosmetic, hospitality, fast-food and mobile-banking (Jayarathne, 2018; Rathnaya & Jayasuriya, 2021; Nushra & Mubarak, 2022; Shashikala & Thilina, 2020; Senayake & Mudiyanse, 2022), yet there is inadequate research that have been carried out to explore the impact of e-WOM on purchase intention in the fashion industry of Sri Lanka. Hence, this study seeks to contribute to fill the knowledge void in this specific area mainly focusing on e-WOM generated through social media platforms in Sri Lanka. It aims to improve e-WOM by providing insights for fashion brands and recommending targeted strategies that assist the recovery of the fashion industry.

2 Literature Review

2.1 Key Terms

2.1.1 Electronic Word-of-Mouth

Electronic word-of-mouth (e-WOM) is "the dynamic and the ongoing information exchange process between potential, actual or former consumers regarding a product, service, company or brand which is available to a multitude of people and institutions via the Internet" (Henning Thurau et al., 2004). The first ever research study on e-WOM analyzed the transfer of basic information between senders and receivers (Bansal and Voyer, 2000). The evolution of internet and Web 2.0, have allowed individuals to disseminate information and to display their preferences among friends and social networks (Alboqami et al., 2015).

The favorable impacts of e-WOM on consumer intentions have been studied in previous researches focusing on the aspects such as social media platforms (Samoggia et al. 2020), online review characteristics (Zinko et al., 2020), and brand perception (Tajuddin et al, 2020) and its impact on the customer's attitude and behavior, to build EWOM. But there are still gaps in our knowledge of its long-term consequences, cultural variances, information adaptability and complex source credibility variables. More research in such knowledge gaps will assist to look into customer emotions' effect on e-WOM effectiveness, social media algorithms and its impact on e-WOM visibility, and whether there are any moral ramifications to employing e-WOM for marketing. Hence through this research it is aimed to provide insightful information that can be used to improve e-WOM strategy and customer engagement for fashion brands.

2.1.2 Purchase Intention

Purchase intention is a combination of consumers' interest in purchasing and possibility of buying a product (Khan, 2019). Customer purchase intention is the construct that gives the confidence, interest and strength within the customer in purchasing a product (Thamizhvanan and Xavier, 2013). The purchase intention has been used broadly as a focal construct in order to indicate the consumers' buying behavior in the market research (Liew & Falahat, 2019).

2.2 Frameworks and Models

2.2.1 Elaboration Likelihood Model

The elaboration likelihood model is a dual-process theory of persuasion that was proposed (Petty et al., 1983), in-order to understand how individuals process information shared in different medias, to form attitudes and behavioral intentions in decision-making (Wang, 2015). The central route involves high level cognitive processing, where individuals evaluate the information that is presented to them, while the peripheral route involves low-level cognitive processing where individuals who lack motivation, make decisions depending on source credibility or the quantity of information (Filieri & McLeay, 2013; Yang et al., 2022). Businesses can use this model to tailor e-WOM strategies based on customer motivation and engagement for maximum persuasive impact.

2.2.2 Information Adoption Model

In e-WOM or online information sharing, the way an information is understood and processed varies depending on each person, as the same message can generate various outcomes, due to the difference in beliefs, attitudes, and experiences of individuals (Chaiken & Eagly, 1976). Hence, in order to understand how each individual adopts the information differently and how it influences their behavioral intention, Sussman and Siegal (2003) introduced the Information Adoption Model (Wang, 2016).

2.2.3 Source Credibility Model

The source credibility model elaborates on the impact of the communicator's positive attributes on the receiver's acceptance of an information (Hovland et al, 1953). The model demonstrates that the performance of an information shared among consumers depends on the source's expertise, trust and attractiveness. The sender's expertise refers to the knowledge the individual possesses about the product. Trustworthiness refers to the extent the source is perceived as honest and is considered confident (Ohanian, 1990), while attractiveness refers to the extent to which the source is familiar to the information receiver (Kumar, 2010; Byrne et al., 2003). Source expertise enhances trust and confidence, while trustworthiness fosters belief in e-WOM. Familiarity with the source adds emotional appeal, collectively shaping consumers' perceptions and purchase decisions.

2.2.4 The Theory of Planned Behavior

The TPB is a theoretical framework developed based on the earlier Theory of Reasoned Action, to predict and explain human behavior in certain specific context (Ajzen, 1991, 2021). It reasons that an individual's intention to perform a behavior is a critical predictor of whether the action will be actually performed (Ajzen and Fishbein, 1980; Bosnjak et al., 2020). Accordingly, this human behavior is impacted by three components; attitude towards behavior, subjective Norm, perceived behavioral control. This model assists in portraying the subject matter where it elaborates on how positive e-WOM enhances customers' perceptions of power over the purchasing decision, modifies their attitudes, and affects subjective norms through endorsements from reliable sources. These elements work together to raise the probability that buyers will intend to buy based on the received e-WOM.

3 Methodology

3.1 Conceptual Framework

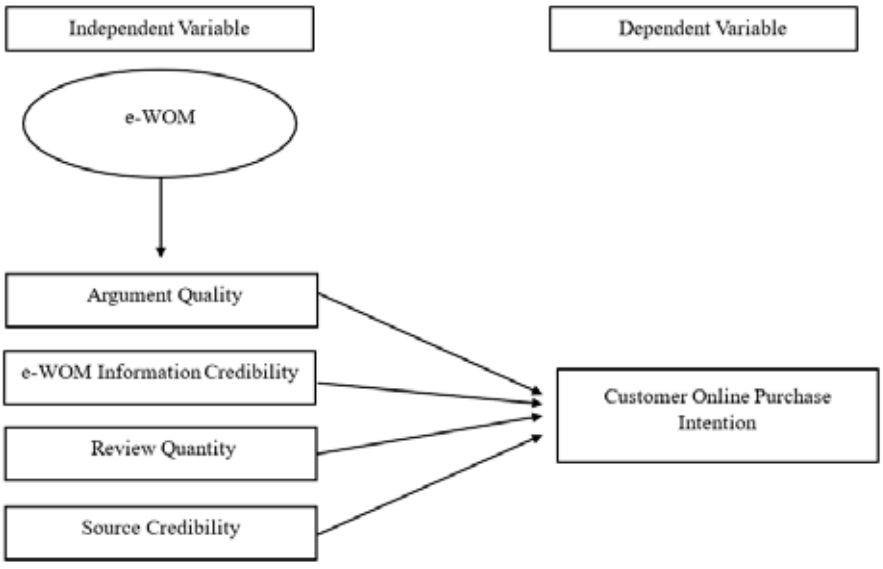


Figure 1: Conceptual Framework (Author Developed)

3.2 Independent and Dependent Variables Rationality

3.2.1 Argument Quality

The e-WOM argument quality is referred to as the convincing or the persuading strength of the information, which are shared in the digital platforms to customers by other customers, assisting them to evaluate a product or brand to make a purchasing decision (Ismagilova et al., 2019). In a study regarding the impact of e-WOM on purchase intention by Luo et al, (2014), and Teng et al, (2017), relevance, clarity, accessibility and timeliness of information were found to be significant predictors of argument quality.

3.2.2 E-WOM information Credibility

The credibility of information demonstrates how much the receiver of the information learns and adopts to the received information. For a message to be credible, it should be accurate, logical, believable and honest to the recipients (Thomas, Wirtz & Weyerer, 2019; Bataineh, 2015). Thus, when consumers perceive a credible message, they are more likely to engage in it and believe it (Abedi et al., 2020; Erkan, 2016; Yusuf et al., 2018).

3.2.3 E-WOM Review Quantity

The e-WOM review quantity refers to the total number of followers, reviews, likes, and comments that are posted about a particular product or service (Lim, 2016). In a study by Nushra and Mubarak (2022), it was identified that the volume of e-WOM indicates the products' popularity and value showing that it has persuaded a large number of audiences to consider buying the product. Hence, review quantity has been considered as a dimension influencing purchasing intention.

3.2.4 Source Credibility

Source credibility refers to the degree to which a source is trusted and considered to have reliable expertise. The evaluation of a source is analyzed in terms of expertise, trustworthiness, credential and familiarity (Ismagilova et al., 2019). The customer's belief and attitude, and the level of trust that is put towards the source will affect the message receiver's level of belief about what the source claims, impacting the customer's purchase intention (Sharif et al., 2016).

3.2.5 Purchase Intention

The purchasing intention is a focal theory in industrial psychology that is used to define and measure multiple aspects of customer's purchasing behavior. Multiple researchers have conducted studies to recognize the factors that influences the purchasing intentions. Kian et al. (2017) illustrated that demographic factors such as age, gender, professional status, education-level to be considered as certain factors influencing the purchasing intention. Yet, in aspect to SM, elements such as trust, social-influence, recognition, and customer preference were acknowledged as few most referred influencing predictors.

3.3 Research Design and Methodology

Research in the field of understanding the impact of e-WOM on customer purchasing intentions in the Sri Lankan fashion industry is guided by a positivist philosophy, emphasizing the acquisition of knowledge through empirical research and observation rather than speculation. This research study further utilized a deductive approach to test the relationship between independent and dependent variables, using quantitative methodology and standardized instruments, including a Likert scale survey (Bougie and Sekaran, 2019).

As for the sampling technique, the convenience sampling method was employed in which the participants were selected based on the availability and the willingness to participate in the study (Etikan et al., 2016). The research aims to gather data from individuals above 18 years of age in Colombo, due to the city's prominence in fashion and large number of social-media users (Bandara, 2021). The population of Colombo as of 2023, is 633,000 (MacroTrends, 2023). Accordingly, using the Anderson and Saunder's method (2012) of calculating the sample size, the sample size for this study was set as 384.

3.4 Hypotheses Development

Hypothesis 01	$H1_0$ - Argument Quality has no relationship on customer purchase intention $H1_1$ - Argument Quality has a positive relationship on customer purchase intentions	Senanayake and Mudiyanse (2022), Ismagilova et al. (2019),
Hypothesis 02	$H2_0$ –e-WOM Information Credibility has no relationship on customer purchase intention $H2_1$ - e-WOM Information Credibility has a positive relationship on customer purchase intentions	Ismagilova et al. (2019), Teng et al. (2017)
Hypothesis 03	$H3_0$ - Review Quantity has no relationship on customer purchase intention $H3_1$ - Review Quantity has a positive relationship on customer purchase intentions	Utami et al. (2020), Mehyar et al. (2020), Rachibini (2018)
Hypothesis 04	$H4_0$ –Source credibility has no relationship on customer purchase intention $H4_1$ - Source credibility has a positive relationship on customer purchase intentions	Mubarak & Ishar Ali (2018), Ismagilova et al. (2019)

Table 1: Hypothesis Development (Author Developed)

By conducting a comprehensive systematic literature review and meticulous analysis, the research hypotheses were formulated based on the synthesized arguments presented in the aforementioned sources, driven by a desire to address a crucial gap in the current understanding of the subject matter.

4 Data Analyzation and Representation

4.1 Data Analysis, Reliability and Validity

Prior to commencing the data analysis of all data collected (250 responses), it is essential and crucial to verify the reliability and validity of the data and questionnaire, in-order to conduct a comprehensive analysis aimed at exploring the relationship between variables. The validity and reliability tests were conducted using SPSS for each variable and the results, were compared against the threshold values depicted in table 1.

Measurement	Threshold Value
Reliability	
Cronbach's Alpha	>0.7
Validity	
KMO	>0.5
Bartlett's test	<0.05
AVE	>0.5

Table 2: Threshold Values for reliability and Validity (Author Developed)

Variable	Cronbach's Alpha	KMO	Bartlett's Test	AVE
Argument Quality	0.777	0.760	.000	0.60
e-WOM information credibility	0.860	0.822	.000	0.70
Review Quantity	0.797	0.776	.000	0.62
Source Credibility	0.801	0.770	.000	0.63
Purchase intention	0.806	0.836	.000	0.58

Table 3: Reliability and Validity results based on SPSS (Author Developed)

The SPSS Analysis revealed high internal consistency (Cronbach's Alpha: 0.777 to 0.860) and adequate convergent validity (AVE: 0.58 to 0.70) for the variables. Further, the Kaiser-Meyer-Olkin (KMO) values ranged between 0.760 to 0.836 indicating the data's suitability for factor analysis, while the variables'' significance were tested through the Bartlett's test and were at 0.000 ($p < 0.05$), supporting the validity of the measures.

4.2 Hypotheses Testing

The deductive approach was chosen to test the relationship between the independent variables and dependent variable based on the hypotheses that were formed. Accordingly, the correlation standards are shown as below,

Correlation	Interpretation
<0.4	Low
0.4 – 0.7	Moderate
>0.7	High

Table 4: Correlation Standards (Dayanand, 2018)

Through the application of correlation and regression tests on SPSS, it was revealed that each correlation value was significant at level 0.01 (Appendix). Therefore, in conclusion all 4 variables have significant and positive relationship with purchasing Intention, hence all four alternate hypotheses were accepted.

Hypotheses	Correlation	Significance	Sig. Regression coefficient	Conclusion
$H1_0$ – Argument Quality has no relationship with customer purchase intention $H1_1$ - Argument Quality has a positive relationship with customer purchase intentions	.368	.000	.428	$H1_0$ – Rejected $H1_1$ - Accepted
$H2_0$ – e-WOM Information Credibility has no relationship with customer purchase intentions $H2_1$ - e-WOM Information Credibility has a positive relationship with customer purchase intention	.421	.000	.398	$H2_0$ – Rejected $H2_1$ - Accepted
$H3_0$ - e-WOM Review Quantity has no relationship with customer purchase intentions $H3_1$ - e-WOM Review Quantity has a positive relationship with customer purchase intention	.264	.000	.295	$H3_0$ – Rejected $H3_1$ - Accepted
$H4_0$ – Source credibility has no relationship with customer purchase intentions $H4_1$ - Source credibility has a positive relationship with customer purchase intention	.621	.000	.622	$H4_0$ – Rejected $H4_1$ - Accepted

Table 5: Hypothesis Test Results (Author Developed*)*

4.3 Multiple Regression Analysis

Multiple regression analysis is an advanced version of the linear regression which allows the examination of the collective impact of different independent variables on the dependent variable. This is because in reality independent variables' effect does not act in isolation and their combined effect needs to be studied and demonstrated. In-order to correlate with each other, the combined significance of the variable should be less than 0.05.

Value of R2	Interpretation
$0.25<R^2<0.50$	Weak
$0.50<R^2<0.75$	Moderate
$R^2>0.75$	Substantial

Table 6: Interpretation of regression values (Hair, Ringle and Sarstedt, 2013)

As calculated through the multiple regression test, the R2 value of .414 indicates that 41.4% of the variation in the purchase intention can be accounted by the specified e-WOM variables. It was proved, that even though each independent variable had a positive and significant relationship with the dependent variable, combined together the effect of e-WOM variables show a weak effect. The weak effect indicates that apart from the mentioned variables in the study, there are several other variables that influences the dependent variable. The multicollinearity effect (Shrestha, 2020) between the mentioned independent variables were tested and was identified that there is no strong relationship

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of Estimate
1	.643	.414	.404	.50257
2	.621	.386	.383	.51147

Table 7: Model Summary developed based on SPSS (Author Developed)

.

5 Conclusion and Recommendation

5.1 Discussion of the research objectives

Objective 01: The different dimensions of e-WOM (argument quality, information credibility, review quantity and source credibility) were identified through empirical study based on previous researches, frameworks and models in section 2.2 and 2.3. A more detailed analyzation was also provided through the operationalization table along with the conceptual framework in section 3.1 and 3.2 respectively.

Objective 02: The relationship of each independent variable selected for this study was tested and analyzed against the dependent variable "customer purchase intention". Each variable showed a positive and significant relationship with purchasing intention and the hypotheses were tested and accepted accordingly as illustrated in section 4.4. It was also derived to a conclusion that source credibility is considered as the most influential and significant factor in changing the customer's intention

Objective 03: The significance of this study was achieved as the relationship between the dimensions of e-WOM were tested against the dependent variable and was identified that each of them have a positive and significant relationship. Thus, through this it is advised for the fashion industry of Sri Lanka to consider these variables and understand customer's intention to build their marketing strategies are provided under section 5.2

5.2 Recommendation

Recommendation for fashion brands

To maintain the argument quality of e-WOM information, fashion brands could encourage honest and authentic reviews from their customers by creating a community (ex: Facebook customer groups), a comfortable environment for the customers to share their thoughts within the social media platforms. Fashion brands could assist in improving the information credibility through paid advertisement and by creating strong value-added contents on social media, which will trigger the curiosity within the customers to engage about it on their own feeds. To increase the quantity in aspect of followers, likes and comments, fashion corporations could leverage user generated contents, use social listening tools and use resources in creating high value engaging social-media marketing contents. Finally, to increase the source credibility, fashion

brands could partner with social media influencers and celebrities who are well-known by the customers to promote their products, and invest on social media optimization to increase brand awareness.

Recommendation for future research

Future researchers should explore various e-WOM dimensions and collect data from diverse locations of Sri Lanka to enhance validity and reliability of the future researches. Conducting qualitative research with a focus group and using a mixed methodology can further enhance the study's scope and accuracy.

5.3 Limitation of Study

Limited researches have been conducted in Sri Lanka for the fashion industry in this context. As the questionnaire was sent out through different online platforms and hence there is a possibility for the respondents to have misinterpreted the question as they had no face-to-face discussion with the author. Furthermore, the study was solely focused on the quantitative method, hence supporting this research with a mixed method by including the qualitative aspects could have strengthened the research further. The survey was conducted within a certain period of time as the time was restricted and the data was collected from a target group of respondents from Colombo and could have been tested from different location of the country. Further to the above, the 12-page limitation restricted the author from providing a detailed in-depth analysis.

5.4 Conclusion

The study aimed to explore the impact of e-WOM on customer purchasing intention in the Sri Lankan Fashion industry by identifying various dimensions of e-WOM. Thus, e-WOM dimensions argument quality, information credibility, review quantity and source credibility were identified and indicators for each of the variables were recognized to build the questionnaire. The questionnaire was distributed to collect data to test the relationship between e-WOM and purchase intention and a total of 250 valid responses were received. Data was analyzed through SPSS and the hypotheses were tested and confirmed. Accordingly, it was identified that all 4 selected independent variables had a positive and significant relationship with customer purchase intention, with “source credibility” being the most impacting factor. The study concluded with discussing the findings, limitations, recommendation to the fashion industry and serving as a reference for future studies.

REFERENCES

[1] Ajzen, I. (1991). The theory of planned behavior. Organizational Behavior and Human Decision Processes, 50 (2), 179–211.

[2] [Badir, M. and Andjarwati, A.L. (2020). The effect of E-WOM, ease of use and trust on purchase decisions (study on Tokopedia application users). Journal Minds: Management Ide dan Inspirasi, 7 (1), 39.

[3] Bandara, D.M. (2021). Impact of social media advertising on consumer buying behavior: With special reference to Fast Fashion Industry. Sri Lanka Journal of Marketing, 7 (2), 80.

[4] Bataineh, A.Q. (2015). The impact of perceived e-WOM on purchase intention: The mediating role of Corporate Image. International Journal of Marketing Studies, 7 (1).

[5] Bhat, N.Y. and Bhat, S.A. (2020). The influence of electronic word of mouth (EWOM) on Consumers Purchase Intention: A review and analysis of the existing literature. ResearchGate. www.iosrjen.org.

[6] Bosnjak, M., Ajzen, I. and Schmidt, P. (2020). The theory of planned behavior: Selected recent advances and applications. Europe's Journal of Psychology, 16 (3), 352–356.

[7] Dayanand, A. (2018). Hypothesis types and research. International Journal of Nursing Science Practice and Research, 4 (2).

[8] Erkan, I. and Evans, C. (2016). The influence of ewom in social media on consumers' purchase intentions: An extended approach to information adoption. Computers in Human Behavior, 61 47–55.

[9] Etikan, I. (2016). Comparison of convenience sampling and purposive sampling. American Journal of Theoretical and Applied Statistics, 5 (1), 1.

[10] Hair, J.F., Ringle, C.M. and Sarstedt, M. (2013). Partial least squares structural equation modeling: Rigorous applications, better results and higher acceptance. Long Range Planning, 46 (1-2), 1–12.

[11] Hennig-Thurau, T. et al. (2004). Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the internet? Journal of Interactive Marketing, 18 (1), 38–52.

[12] Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). Communication and persuasion: Psychological studies of opinion change. American Sociological Review, 19 (3), 355.

[13] Ismagilova, E. et al. (2019). The effect of electronic word of mouth communications on intention to buy: A meta-analysis. Information Systems Frontiers, 22 (5), 1203–1226.

[14] Jalilvand, M. and Samiei, N. (2012). The effect of electronic word of mouth on brand image and purchase intention. Marketing Intelligence & Planning, 30 (4), 460–476.

[15] Jayarathne, D.N. (2018). Investigation on the impact of EWOM on social media towards the purchase intention of consumers in retail industry in Sri Lanka. The International Journal of Business & Management, 6 (12).

[16] Macrotrends (2023). Colombo, Sri Lanka Metro Area Population 1950-2023. MacroTrends.

[17] Nushra, M.N. and Mubarak, K.M. (2022). Influence of customer generated e-WOM on purchase intention of cosmetic products in Sri Lanka. Sri Lanka Journal of Marketing, 8 92.

[18] Petty, R.E., Cacioppo, J.T. and Schumann, D. (1983). Central and peripheral routes to advertising effectiveness: The moderating role of involvement. Journal of Consumer Research, 10 (2), 135.

[19] Rathnayake, Y. and Jayasuriya, N.A. (2021). The impact of electronic word of mouth on brand image and purchase intention – a study on hospitality industry in Sri Lanka. SLIIT Business Review, 01 (01), 59–77.

[20] Saunders, M.N.K. et al. (2019). "Research methods for business students" Chapter 4: Understanding ... Research Methods for Business Students (pp.128-171) Edition: 8Chapter: 4Publisher: Pearson Education.

[21] Wang, Y. (2016). Information adoption model, a review of the literature. Journal of Economics, Business and Management, Vol. 4, No. 11.

[22] Utami, S., Setyowati, N. and Mandasari, P. (2020). Celebrity brand ambassador and e-WOM as determinants of purchase intention: A survey of Indonesian celebrity cake. E3S Web of Conferences, 142 05001

A APPENDICES

Correlation and Regression Tests for "Argument Quality"

Correlations

```
CORRELATIONS
  /VARIABLES=AQ PI
  /PRINT=TWOTAIL NOSIG
  /MISSING=PAIRWISE.
```

Regression

```
REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT PI
  /METHOD=ENTER AQ.
```

Correlations

		AQ	PI
AQ	Pearson Correlation	1	.368**
	Sig. (2-tailed)		.000
	N	250	250
PI	Pearson Correlation	.368**	1
	Sig. (2-tailed)	.000	
	N	250	250

** Correlation is significant at the 0.01 level (2-tailed).

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.368a	.135	.132	.60679

a. Predictors: (Constant), AQ

ANOVAa

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	14.285	1	14.285	38.798	.000b
	Residual	91.311	248	.368		
	Total	105.596	249			

a. Dependent Variable: PI

b. Predictors: (Constant), AQ

Coefficientsa

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2.151	.257		8.522	.000
	AQ	.428	.069	.368	6.229	.000

a. Dependent Variable: PI

Correlation and Regression Tests for "Information Credibility"

Correlations

```
CORRELATIONS
  /VARIABLES=IC PI
  /PRINT=TWOTAIL NOSIG
  /MISSING=PAIRWISE.
```

Regression

```
REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT PI
  /METHOD=ENTER IC.
```

Correlations

		IC	PI
IC	Pearson Correlation	1	.421**
	Sig. (2-tailed)		.000
	N	250	250
PI	Pearson Correlation	.421**	1
	Sig. (2-tailed)	.000	
	N	250	250

** Correlation is significant at the 0.01 level (2-tailed).

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.421a	.177	.174	.59200

a. Predictors: (Constant), IC

ANOVAa

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	18.680	1	18.680	53.300	.000b
	Residual	86.916	248	.350		
	Total	105.596	249			

a. Dependent Variable: PI

b. Predictors: (Constant), IC

Coefficientsa

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2.438	.187		13.050	.000
	IC	.398	.055	.421	7.301	.000

a. Dependent Variable: PI

Correlation and Regression Tests for “Review Quantity”

Correlations

```
CORRELATIONS
  /VARIABLES=RQ PI
  /PRINT=TWOTAIL NOSIG
  /MISSING=PAIRWISE.
```

Correlations

		RQ	PI
RQ	Pearson Correlation	1	.264**
	Sig. (2-tailed)		.000
	N	250	250
PI	Pearson Correlation	.264**	1
	Sig. (2-tailed)	.000	
	N	250	250

**. Correlation is significant at the 0.01 level (2-tailed).

Regression

```
REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT PI
  /METHOD=ENTER RQ.
```

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.264[a]	.070	.066	.62939

a. Predictors: (Constant), RQ

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	7.355	1	7.355	18.566	.000[b]
	Residual	98.242	248	.396		
	Total	105.596	249			

a. Dependent Variable: PI

b. Predictors: (Constant), RQ

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2.773	.236		11.760	.000
	RQ	.295	.069	.264	4.309	.000

a. Dependent Variable: PI

Correlation and Regression Tests for “Source Credibility”

Correlations

```
CORRELATIONS
  /VARIABLES=SC PI
  /PRINT=TWOTAIL NOSIG
  /MISSING=PAIRWISE.
```

Correlations

		SC	PI
SC	Pearson Correlation	1	.621**
	Sig. (2-tailed)		.000
	N	250	250
PI	Pearson Correlation	.621**	1
	Sig. (2-tailed)	.000	
	N	250	250

**. Correlation is significant at the 0.01 level (2-tailed).

Regression

```
REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT PI
  /METHOD=ENTER SC.
```

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.621[a]	.386	.383	.51147

a. Predictors: (Constant), SC

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	40.718	1	40.718	155.649	.000[b]
	Residual	64.878	248	.262		
	Total	105.596	249			

a. Dependent Variable: PI

b. Predictors: (Constant), SC

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.576	.179		8.793	.000
	SC	.622	.050	.621	12.476	.000

a. Dependent Variable: PI

REMADDICT: A SOLUTION PUT FORWARD TO PREDICT WITHDRAWAL SYMPTOMS AND ACT AS A MAINTENANCE TOOL TO ASSIST ADOLESCENTS CHALLENGED WITH SUBSTANCE USE DISORDER

First A. Author

Avisha de Mel, Informatics Institute of Technology, 528 R.A De Mel Mawatha, Colombo, avisha.2018398@iit.ac.lk

Second B. Author

Janice Abeykoon, Informatics Institute of Technology, 528 R.A De Mel Mawatha, Colombo, janice.a@iit.ac.lk

ABSTRACT

This article is based on the research conducted to implement an application catering towards the youth diagnosed with Substance abuse disorder in Sri Lanka. The subject of 'war on drugs' has only worsened over time in the world, even so that the spread of drugs is rapidly increasing amongst the youth. It's an overlooked issue even in Sri Lanka, and is currently in the need of imminent attention as the country has been witnessing an alarming rise in numbers amongst adolescents in the recent past. Additionally, Individuals struggling with substance use disorder are most often stigmatized and criminalized in Sri Lanka, with little to no proper support in most parts of the island. The main aim of this study is to identify the factors contributing to the increase of drug addiction, and to develop a mean that would contribute as a withdrawal symptom predictor and a maintenance application to assist drawing patients closer to medical professionals in the field with features such as a knowledge hub, diagnosis based on DSM-5 criteria, conducting online therapy, goal trackers, searching nearby therapists, etc. With the world rapidly evolving in technology, the e-health sector of Sri Lanka too needs major contribution in terms of supplying to the uncontrollable menace of the substance abuse disorder amongst the younger population. Having little to no application built in this domain, REMADDICT aims to cater as an application that positively serves as a tool to aid adolescents burdened with the condition.

KEYWORDS

Substance use disorder (SUD), Adolescents, drug addiction, withdrawal symptoms

1 Introduction

The term 'substance use disorder' refers to a chronic, relapsing brain disorder because of obsessive substance use, despite the adversative consequences it may bring upon an individual. This condition involves the brain circuits associated with rewarding, stress and self-control whilst bringing functional changes to the brain. These changes have the potential of lasting a long time even after a person stops intaking drugs [1].

Drug addiction most often begins with consuming as an experiment in social situations or with prescribed medication. When the user starts consuming more frequently, the user becomes an addict. The risk of addiction and the rate of getting addicted depends on the drug. It is stated that some drugs even have a higher risk of addiction than others, and with passing time the dosage consumed increases as well. Addicts find it extremely difficult to go about without the drug and when they attempt to stop consuming, it may lead to withdrawal symptoms which are intense cravings that can make the user physically weak [2].

2 Problem Domain

2.1 Problem statement

Sri Lanka has been gravely burdened with drug addiction cases over the years. According to the statistics of the national dangerous drugs control board (NDDCB) in 2020, there are over 97,000 people who are addicted to Heroin alone across the country [3]. SUD is currently a severe issue due to its catastrophic impact on public health, cause of diseases, poverty, destruction of family life and increased rates of crime. Drug addicts get their supplies from the underground drug market, and although laws and regulations have been implemented to reduce drug addiction in Sri Lanka, the desired outcome is not met. Hence, the need for rehabilitation and support centers for drug addicts have increased [4].

It is stated that 16% of adolescents have used at least one illegal substance during their lifetime, particularly including drugs [5]. The willingness to experiment, and increased autonomy are most often swayed by peer pressure to make high-risk decisions which leads to substance abuse disorder. Moreover, due to constant usage, they tend to get addicted, and use drugs as an escape for all their problems until it becomes their primary focus, even above family life and studies [6].

The problem worsens as it's proven to say that there's a lack of resources and quality education to address this issue in some parts of Sri Lanka, which results in children consuming these substances, unknowing the consequences [7]. Additionally, even though some may attempt to get better, due to the inability of coping up with painful withdrawal symptoms, most users experience relapsing [8]. The following diagram identifies the main reasons for the prevalence of drug addiction in Sri Lanka and the causing effects to those problems,

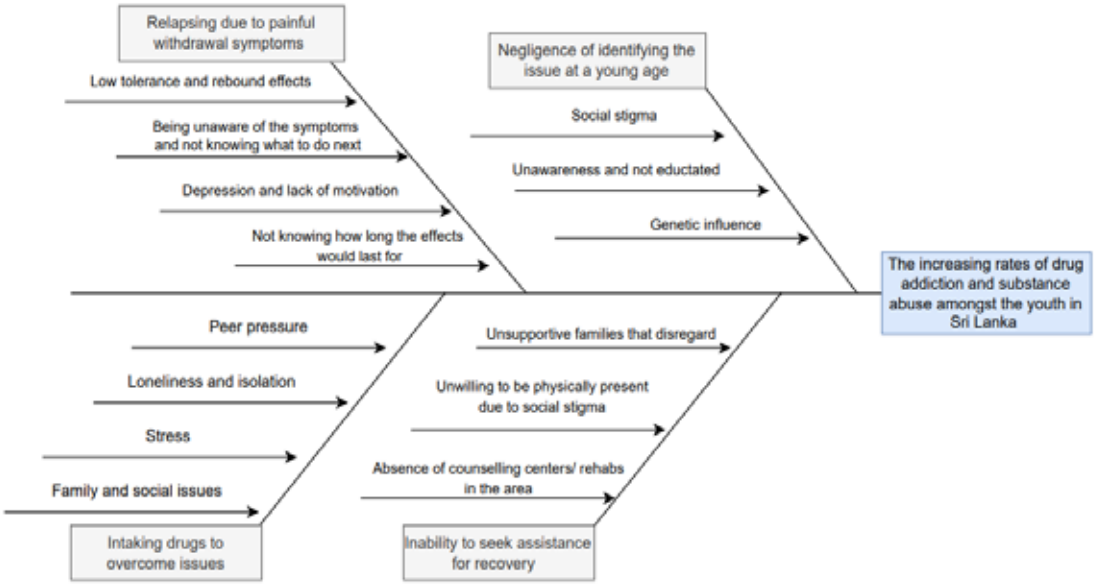


Figure 1: Fish bone diagram

2.2 Substance use disorder and withdrawal symptoms

One of the main contributors to the increasing rate of drug addiction amongst youth is relapsing. When a user refrains from substances for a certain period which is also known as abstinence, the body begins to exhibit withdrawal symptoms. The results obtained during research showed that isolation and lack of motivation are two supporting factors for relapsing as it takes an immense amount of willpower to battle against these symptoms. According to a recent study conducted in Sri Lanka, it's been concluded that nearly 40 per cent

of drug addicts who are rehabilitated, start using drugs again [9]. Likewise, it is evident that having low tolerance and rebound effects in the process of abstinence, most often leads to relapsing.
When it comes to withdrawal symptoms, it's important to be prepared and aware. Those who may be unaware of withdrawal symptoms and the next step to take, also tend to relapse under the assumption that the substance that kept them going is the only solution. To the diseased, these symptoms might be unbearable and might seem like forever. And thus, not knowing how long the symptoms will last for, also leads to relapsing.
Additionally, issues such as not identifying and addressing the condition at a young age as a result of the prevailing social stigma and lack of awareness on SUD, intaking drugs to overcome issues varying from peer pressure, stress depression, etc., and being unable to seek assistance for recovery due to unsupportive families, lack of resources in certain areas, are some of the other factors that were identified during research.

3 Literature Survey

3.1 The rapid increase of drug addiction in Sri Lanka

The total number of drug centered arrests from January to October 2022 was 67,900. Out of which 35,765 were for heroin, 25,114 for cannabis, 6,728 for methamphetamine, 232 for psychotropic substances, 21 for cocaine and 40 for hashish. In addition, there has been a significant rise in the use of methamphetamine or 'ice' amongst school children across the country and is even growing to be a potential national problem to the country in the future [10].
The rate of substance use in Sri Lanka has led to harmful effects on individuals, their families, and the community, with no understanding on the context and little to no observed evidence available. With the increasing rates of addiction and relapsing, the need for thoroughly assessed, effective rehabilitation treatment has also increased. Social inequities, post-traumatic stress disorder, poor social support, residing in urban areas, mental illnesses, family history, stress, mass media influence and peer pressure are the common risk factors for SUD in Sri Lanka. Adding on, research has also proved that most addicts start consuming at a very young age [11].

3.2 Drug addiction amongst the youth

Adolescence, defined to be the transitioning period from childhood to manhood, ranging from the years 10-19, is one of the crucial stages of brain development as the brain undergoes rapid changes. Furthermore, both individual and environmental characteristics most often influence changes that may take place during this period. Increased autonomy, the willingness to experiment and most significantly, peer pressure are factors that contribute especially towards drug addiction amongst the youth [12].
Drug addiction and substance abuse in Sri Lanka is a prominent issue and needs serious attention. Most cases have proven that the first use of these substances was at adolescence and some people keep continuing to consume. Whilst this issue is directly linked to peer pressure, it also depends on the individual's ability to deal with stress and frustration. Moreover, the lack of proper education, employment and educational opportunities also encourage them to turn to coping mechanisms that include SUD, leading them to prolonged long term effects both mentally and physically [13].

3.3 The impact of drug addiction

3.3.1 Impact to public health

Drugs are produced of chemicals that can bring about ill-effects to the brain. The effects depend on the type of the drug and the consequences can be long-lasting or perhaps even permanent so much so that it continues even after the person stops taking the substance [14]. Most individuals with drug addiction often have one or more associated health issues, including lung or heart disease, cancer, cardiovascular diseases, and mental health conditions. SUD's can also increase the chances of contracting infections such as HIV, or Hepatitis C [1]. It's also been found that the rates of deaths, illness and disorders are mainly associated with SUD, above any other preventable health condition.

3.3.2 Destruction to family life

According to psychology today, 1 in 5 children grow up in a home with an addict. Being a witness to SUD-related trauma at a young age may develop long-term effects on a child [15]. Likewise, parents who have affected teens in their homes also have a unique set of difficulties. They're constantly distressed about their children's well-being, or even blaming themselves for what went wrong [16]. Families and loved ones suffer at large with factors such as, difficulty in understanding the disease process and the addict's behavior, draining family resources, decrease in a sense of responsibility in the house, sickness and death, illicit relationships, distortion of interpersonal and intrapersonal family relationships, violence, etc. [1].

3.3.3 Increased rate of crime

Sri Lanka gets its drug supply from both internal and external sources. Whilst some drugs such as Cannabis are illegally cultivated in the country, pharmacies play the main role in the drug distribution network in the country as pharmacists use their personal license to buy and supply drugs over the limit. Moreover, these drugs are mainly imported to Sri Lanka via India or Pakistan. Websites in the USA, UK and Europe offer illegal drugs, enabling pharmacies to order online [17]. Most drug-related crimes are for possession and not for sale or distribution.

3.4 Withdrawal symptoms and its impacts

Substance abuse disorders most often lead to physiological dependence when an individual consumes in a large quantity. Once the consumption stops, individuals who are physiologically dependent, may experience unpleasant withdrawal symptoms. These include physical, mental, and emotional symptoms [18]

There are several factors that cause withdrawal symptoms. The type of drug the user was addicted to is the main factor, whilst factors like genetics and metabolism can be contributing factors as well. Withdrawal symptoms may appear from a few hours from last intake to a few days and can last from days to weeks and in severe addiction, long-term symptoms can last for months.

Withdrawal symptoms may also differ according to the drug that is used. Table 1 shows the differentiation of withdrawal symptoms according to the drug. (World Health Organization, 2009)

Table 1: Types of drugs and their withdrawal symptoms

Type of drug	Duration	Withdrawal symptoms
Opioids	8-24 hours (Short-acting) 10-20 hours (Long acting)	Nausea Anxiety Insomnia Hot/ cold flushes Sweating Muscle cramps Watery discharge from eyes and nose Diarrhea
Cannabis	1-2 weeks	Anxiety

Type of drug	Duration	Withdrawal symptoms
		Feeling fear and isolation Restlessness Irritability Poor appetite Disturbed sleep Gastrointestinal issues Tremor
Cocaine/ Meth	24 hours – 3/5 weeks	Depression Increased sleeping Agitation Muscle pains Increased eating

3.5 Withdrawal symptoms and relapsing

In the path to recovery, relapse is a stage that is met by all those affected by substance abuse and addiction. Relapse is so common that only a handful of people recover and stay clean on their first attempt.

Unfortunately, relapsing immensely affects the brain, as neurotransmitters go back to baseline levels after being significantly impacted in the phase of addiction. This is the main cause for withdrawal symptoms, and during the process of withdrawal, recovery, and relapse, the kindling effect takes place, where each progressive relapse is worse than the previous [19].

Sri Lanka has reported several cases of relapsing due to withdrawal symptoms mainly due to not being aware of these symptoms and how to proceed (17).

3.6 Diagnosis of drug addiction

The first step, or the hardest step in diagnosing usually depends on a friend, family or a person looking for treatment. It becomes harder if the individual doesn't know the extent of their problem and may even require group intervention. Next, the addict visits the support center or a psychiatrist, who screens and assesses the state of the condition and what treatment can be given [14]. Screening often requires a thorough evaluation through professionals in the industry such as psychologists, psychiatrists, and drug counsellors [2].

According to the Diagnostic and statistical manual of mental disorders (DSM-5) by the American psychiatric association, a person must have at least two signs in the symptoms section for more than 12 months to be diagnosed with SUD.

3.7 Existing applications

Despite the many systems that are developing with the rapid use of technology in Sri Lanka, currently there is no system to assist drug addicts specifically, implicated in the country (pilot study). As evident in the data gathered during research, the overall comparisons between a few widely used applications globally, indicates that REMADDICT has a unique feature to fit the Sri Lankan community, such as predicting withdrawal

symptoms with the due dates, getting medical help for emergencies, a knowledge hub for addicts and their loved ones, heart rate monitoring, staying anonymous, etc.

4 Methodology

Firstly, in order to validate the problem, a pilot study was conducted at Mel Madura; a support center in Colombo and the NMIH (National Institute of Mental Health) by interviewing psychiatrists and counselors. Once the problem was validated, the elicitation process was carried out by using mechanisms such as conducting interviews with the diseased and domain experts, observation, survey distribution, etc. The main target population of the survey consisted of both clients and their close circles (i.e., friends and family). It was distributed to a total of 83 people, out of which 44 were friends and family of the deceased, and the remaining 39 were patients. From the questionnaire distributed to the users, 51% belonged to the application's target market which is the age group of 18-30, 85% were Male and 15% Female. The survey that was distributed to friends and family of the affected denoted 97.7% of individuals who were closely related to the affected, from which 27.3% were Female and 65.9% were male. The survey was targeted mainly towards those belonging to the age group of 18-30 and was distributed to obtain results even from different areas of Sri Lanka.

5 Research findings

Amongst the individuals that the survey was passed on to, a count was obtained to measure the severity of the problem by questioning the age category of the affected. While the delegates (friends & family) survey indicated that 97.7% of the affected was from the years 18-30, the users survey depicted 51%. This signifies that the majority of the population belongs to the age category of 18-30. Thus, highlighting the severity of the spread of drugs amongst adolescents.

In order to get an understanding of the rate of consumption amongst the population, the users were prompted with the question that determined the overall consumption rate. While 53% of the users said daily, 24% said occasionally and 21% once a week. Likewise, it is evident that most users consume substances on a daily basis due to its compulsive effect, making it challenging to overcome. And there's still a population of individuals who consume on an occasional basis due to reasons like peer pressure, entertainment, etc.

Because the users survey was obtained from a support center, a majority of 59% said that they were under treatment. While 41% represented the minority that never obtained treatment. However, the delegates survey showed a significant majority of 77% individuals who don't receive treatment. Concluding, the fact that the condition is still prevalent, and effort is not taken for recovery.

Out of all the clients at Mel Madura, 23% are still unaware of what withdrawal symptoms are. However, the delegates survey demonstrated a higher percentage of 48%, which indicates how unaware most people are, when it comes to withdrawal symptoms. Most users return to drugs owing to being unable to bear the adverse effects of withdrawal symptoms. Likewise, 39% of the users stated that they return to drugs due to not being able to cope up with such symptoms or not knowing what to do next (refer figure 1).

As demonstrated in the chart (refer figure 2), 85% of users said that they have tried to stop, and relapsed due to varied reasons revolving around not having the willpower to abstain. Similarly, the results obtained from these questions demonstrate the importance of keeping one constantly motivated and in check, which can be achieved through an application.

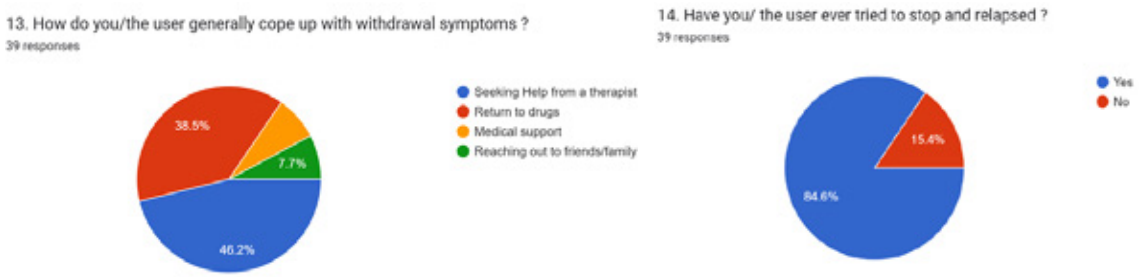


Figure 2: Survey results for coping mechanism and relapsing

6 The solution

To access the application, the user simply must create an account. With registration the user will be inserting data such as age, dosage, and type of drug to feed the Machine Learning (ML) model and undergo diagnosis that is based on the DSM-5 criteria. With completion of registration, they will be led on to the home screen which indicates all features of REMADDICT such as tracking goals, predicting withdrawal symptoms, providing online therapy, looking for nearby therapy resources, the knowledge hub, etc. Additionally, the therapist will also be provided with features such as knowledge hub, taking notes, scheduling appointments, and storing customer data after creating a profile.

The withdrawal symptom predictor is the main feature of the application (REMADDICT). The main aim of implementing this feature and making it the main functionality of the system was to overcome the issue of people relapsing due to unbearable withdrawal symptoms, as this domain area does not offer a solution that caters towards overcoming withdrawal symptoms as of yet. The functionality of this feature is to predict the due time a user may experience withdrawal symptoms in addition to also predicting the associated symptoms that are likely to occur. The user simply gives the inputs (age, dosage, and type of drug) at registration, and the model uses this input to determine the closest outcome by utilizing a machine learning prediction model based on a neural network. Machine learning was incorporated into the technology behind this feature due to the accuracy of its algorithms and growing popularity in this technologically proficient world. The model was trained by utilizing a dataset with standardized data of the inputs. This feature was taken a step further by directing the user to either reach out to one of their emergency contacts, by using the ‘panic button’ feature through the application itself, in the event of finding it difficult to cope with symptoms. Additionally, users also have the option to search and contact a nearby therapist with the aid of the feature ‘find nearby therapist’.

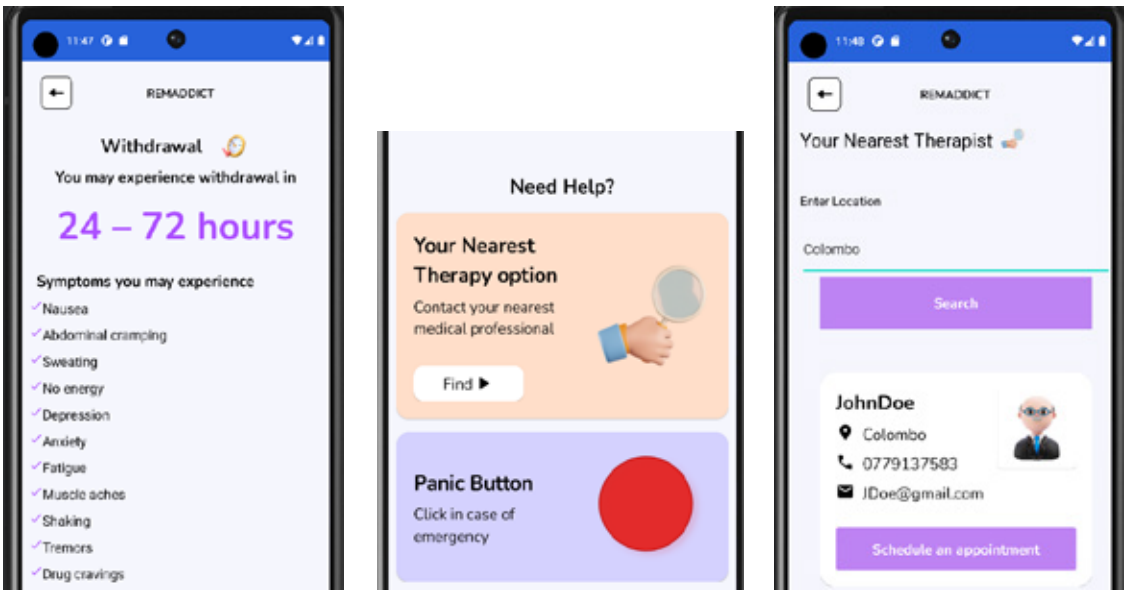


Figure 3: Withdrawal symptom predictor

Apart from the main feature, the application also included a diagnosis feature which was based on the DSM-5 (Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition) is a widely used classification system developed by the American Psychiatric Association (APA) for diagnosing mental disorders. This system was integrated to REMADDICT to diagnose a patient, and was incorporated by using a rating scale consisting of six questions that determine the users' state of condition (where 1-2 criteria matched = mild ,2- 4 criteria matched = medium, and 4-6 criteria matched = extreme).

The goal tracker and the sober tracker, which were integrated to boost the user's morale, have the same functionality. The user simply must create a goal and the system keeps count of how long that goal has been maintained. However, when it comes to the sober tracker, the user will be prompted to update the tracker daily by prompting a notification. With each milestone, users will be entitled to receive real benefits from organizations that may have collaborated with the application. For instance, if an individual reaches a milestone of staying sober for 30 days, he will be entitled to a free therapy trial and a support center that is partnered up with REMADDICT.

7 Future work and Recommendations

The application was reviewed by domain experts and IT experts in the industry after implementation, and the overall conclusion drawn through the review obtained by domain experts demonstrated that REMADDICT received a general positive response in terms of the concept it was built upon, giving the guarantee that this application is of benefit to both the industry and the community.

Although, a few professionals suggested adding a community feature as it can be of great benefit to the age group. IT experts also suggested that this application has reached a fairly good level in terms of apprehending the idea and development, with space to grow in terms of expanding the available resources to bring in more value to the application. However, based on the results of this evaluation, it can be concluded that REMADDICT has great potential to grow with a few enhancements that can be utilized to be developed even further in the future.

Since the target audience is for those between the ages of 18-30, it's an age group that requires novel means to get their attention. Likewise, implementing gamification to increase interactivity, and incorporating a community feature as suggested by professionals is one-way REMADDICT can be improved in the future.

Additionally, high-end technology such as AI, block-chain, and IOT devices can be utilized to diagnose the condition and predict withdrawal symptoms for more accurate results. Incorporating an advanced, personalized bot may also be beneficial to help the user through the process of sobriety.

8 Concluding Remarks

The menace of this condition in Sri Lanka grows to increase in number daily. This stirs up a need for more resources that provide treatment facilities. REMADDICT serves to act as a bridge between those who are reluctant to receive treatment, not privileged with having nearby treatment facilities, and treatment centers. During interviews with medical professionals stated that there has been no system that's been implemented thus far in Sri Lanka to cater towards supporting this arising struggle amongst the youth. Even though the withdrawal symptom predictor was integrated by collaborating one of today's most cutting-edge technologies, it can be further enhanced to be more accurate and advanced. Likewise, even though REMADDICT will be addressing a critical need with features that may be beneficial to its target audience, it takes its place as a milestone application in the domain with the capacity to be improved in the future.

ACKNOWLEDGMENTS

I would like to extend my heartfelt gratitude to my supervisor Ms. Janice Abeykoon for the tremendous support given to me throughout this project, and also to all who contributed towards making this project successful, and a very special mention to the staff of Informatics Institute of Technology.

REFERENCES

[1] national institute on drug abuse. 2020. preventing drug misuse and addiction: the best strategy. *national institute on drug abuse*. retrieved january 24, 2023 from https://nida.nih.gov/publications/drugs-brains-behavior-science-addiction/preventing-drug-misuse-addiction-best-strategy

[2] mayo clinic. 2022. drug addiction (substance use disorder) - symptoms and causes. *mayo clinic*. retrieved january 24, 2023 from https://www.mayoclinic.org/diseases-conditions/drug-addiction/symptoms-causes/syc-20365112

[3] 2022. nddcb publication | national dangerous drugs control board. *www.nddcb.gov.lk*. retrieved march 12, 2022 from http://www.nddcb.gov.lk/nddcb-publication.php

[4] isma lebbe mohamed mahir and thaseem mohamed fathima wazeema. 2020. social aspects of drug addiction in sri lanka. *journal of politics and law* 13, 2 (may 2020), 54. doi:https://doi.org/10.5539/jpl.v13n2p54

[5] nishadi d dharmarathna. 2022. the pattern of substance use and characteristics of the individuals enrolled in residential treatment at selected rehabilitation centers in sri lanka: a descriptive cross-sectional study. *substance abuse: research and treatment* 16, (january 2022), 117822182211008. doi:https://doi.org/10.1177/11782218221100823

[7] chamila siriwardene and a. h. i. sanjeewa. 2017. social work perspective on undue usage of psychoactive substances among school children. *e-journal of social work* 1, 1 (2017), 18–25.

[8] steven m melemis. 2015. relapse prevention and the five rules of recovery. *the yale journal of biology and medicine* 88, 3 (2015), 325–332. retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/pmc4553654/

[9] h. p. g. niwarthana, p. w. p. w. k. pallewaththa, and s. nanayakkara. 2023. explore factors effecting elevate the relapse rate among drug users in sri lanka: qualitative study. *repository.kln.ac.lk* (january 2023). retrieved june 24, 2023 from http://repository.kln.ac.lk/handle/123456789/11193

[10] jonathan frank. 2022. sri lankan youth at risk amid major drug wave. *sunday observer*. retrieved december 25, 2022 from https://www.sundayobserver.lk/2022/12/11/news-features/sri-lankan-youth-risk-amid-major-drug-wave

[11] akila r jayamaha, nishadi d dharmarathna, nimesha dm herath, and thamara darshana. 2022. the pattern of substance use and characteristics of the individuals enrolled in residential treatment at selected rehabilitation centers in sri lanka: a descriptive cross-sectional study. *substance abuse: research and treatment* 16, (january 2022), 117822182211008. doi:https://doi.org/10.1177/11782218221100823

[12] sameera senanayake. 2018. smoking, alcohol consumption, and illegal substance abuse among adolescents in sri lanka: results from sri lankan global school-based health survey 2016. *advances in public health* 2018, (november 2018), 1–7. doi:https://doi.org/10.1155/2018/9724176

[13] chatura rodrigo. 2014. talkingeconomics - a closer look at youth mental health in sri lanka and some consequences. *talkingeconomics*. retrieved november 29, 2022 from https://www.ips.lk/talkingeconomics/2014/08/12/a-closer-look-at-youth-mental-health-in-sri-lanka-and-some-consequences/

[14] alyssa peckham. 2022. physical and mental effects of drug abuse — medical news today. *www.medicalnewstoday.com*. retrieved december 21, 2022 from https://www.medicalnewstoday.com/articles/effects-of-drug-abuse#mental-effects

[15] krystina murray. 2018. how addiction affects the family. *addiction center*. retrieved january 16, 2023 from https://www.addictioncenter.com/addiction/how-addiction-affects-the-family/

[16] elliott redwine. 2020. the impact and effects of drug addiction on the entire family. *peace valley recovery*. retrieved january 17, 2023 from https://www.peacevalleyrecovery.com/blog/how-drug-addiction-affects-the-entire-family/

[17] 2011. *misuse of prescription drugs: a south asia perspective*. un. retrieved march 24, 2023 from https://www.unodc.org/documents/southasia/reports/misuse_of_prescription_drugs_-_a_south_asia_perspective_unodc_2011.pdf

[18] amelia sharp. 2015. drug withdrawal symptoms, timelines, and treatment. *american addiction centers*. retrieved march 13, 2023 from https://americanaddictioncenters.org/withdrawal-timelines-treatments

[20] 2018. what is the kindling effect in alcohol and drug addiction? *windward way*. retrieved march 25, 2023 from https://windwardway.com/alcohol-rehab/kindling-effect/

[21] 2018. diagnosing addiction: process and criteria. *www.medicalnewstoday.com*. retrieved march 25, 2023 from https://www.medicalnewstoday.com/articles/323487#the-diagnostic-process

The Influence of Online Shopping Experience and Attitudinal Loyalty on Online Impulsive Buying in Sri Lankan Online Fashion Retail Industry

Chethana Gallage

Business School, Informatics Institute of Technology, Wellawatta, Colombo, Sri Lanka, chethana.2019569@iit.ac.lk

Gayashini Ratnayake

Business School, Informatics Institute of Technology, Wellawatta, Colombo, Sri Lanka, gayashini.r@iit.ac.lk

Abstract

The remarkable growth in digitalization had successfully reshaped traditional shopping paradigms and introduced online shopping practices across retail industries worldwide including Sri Lanka (Athapaththu and Kulathunga, 2018). In the online retail industry, online shopping experience has resulted a strong relationship towards attitudinal loyalty and online impulsive buying (Gulfraz et al., 2021). This relationship has gained a significant interest from online retailers including fashion retailers due to the substantial influence on their business operations. The purpose of this study is to identify the impact of functional and psychological dimensions of online shopping experience and attitudinal loyalty on online impulsive buying in the Sri Lankan online fashion retail industry in a Business to Customer (B2C) context. The various dimensions and theoretical perspectives of the independent, mediator and dependent variables were inspected through theories and frameworks. The research uses positivism philosophy and deductive approach to test hypotheses. A validated questionnaire was randomly distributed amongst the population to acquire data. The SPSS software was used to carry out quantitative analysis. The study concluded with evidence that online shopping experience has a significant positive impact on online impulsive buying, online shopping experience has a significant positive impact on attitudinal loyalty, attitudinal loyalty has a significant positive impact on online impulsive buying and attitudinal loyalty significantly mediates the relationship between online shopping experience and online impulsive buying. Thus, all hypotheses were accepted. The study also provides valuable recommendation for online shopping experience and attitudinal loyalty backed by research findings and empirical investigations.

Keywords

Online Shopping Experience, Attitudinal Loyalty, Online Impulsive Buying, Online Fashion Retail Industry

1 Introduction

1.1 Background

The widespread use of Information and Communication Technologies (ICT) had caused huge influence on consumer behaviours (Athapaththu and Kulathunga, 2018). The expansion of e-commerce is an excellent example to understand the influence of ICT on consumers. E-commerce is identified as the modern trending art of buying and selling goods and services online (Fonseka et al, 2020). Its expansion is one of the most significant changes in business world in recent years. E-commerce had enabled businesses to expand their operations beyond geographical limits (Fonseka et al, 2022). The global ecommerce retail industry amounted to 5.2 trillion U.S. dollars in 2021 and is expected to rise to 8.1 trillion dollars by 2026 (Statista, 2022).

With time, e-commerce has gained significant fame in Sri Lanka and is accelerating at a high pace (Aboobucker, 2019) with many industries offering goods and services online. In 2023, the e-commerce user penetration of Sri Lanka is expected to be 38.2% and is predicted to grow to 46.5% by 2027 (Statista, 2022). Industries such as fashion and electronics are predominant ecommerce retail industries in Sri Lanka (Raina

et al, 2020). It is projected that in 2023, revenue generated from online fashion retail would reach US$714.60m in Sri Lanka (Statista, 2022). Furthermore, by 2027 it is expected that user penetration for online fashion shopping in Sri Lanka would to hit 33.5% (Statista, 2022).

1.2 Research Gap

The interrelation that exists between various dimensions of shopping experience are not analyzed in most studies (Gulfraz et al., 2021). There had been various studies conducted on online shopping experience overtime yet few studies had focused on functional and psychological factors which affects online shopping experience (Pandey and Chawla, 2018; Gulfraz et al., 2021). Unplanned behaviours, defined as impulsive buying when shopping had only been explored in few studies (Wu et al., 2020). Another dimension that had not received much attention is attitudinal loyalty of online shoppers (Sundstrom et al, 2019; Anshu et al., 2022; Gulfraz et al., 2022). Also, studies lack investigation on how attitudinal loyalty comprise customer experiences despite importance in effective management of customer experience and emotions (Iyer et al., 2020; Cachero-Martínez and Vazquez-Casielles, 2021). Nevertheless, no study had focused on online shopping experience and attitudinal loyalty on online impulsive buying in Sri Lankan fashion industry nor in any other industry in Sri Lanka.
Thereby, this study would focus on attitudinal loyalty and online impulsive buying towards online shopping. The scope of this study is focused on online shopping experience with relation to online fashion retail brands in Sri Lanka. Sri Lankan fashion ecommerce industry had a rise by 18.8% in 2022 and expected to grow in future (Statista, 2022).

1.3 Research Questions

The following research questions were identified.

RQ1: What is the influence of online shopping experience on impulsive buying online?
RQ2: What is the influence of online shopping experience on attitudinal loyalty?
RQ3: What is the influence of attitudinal loyalty on impulsive buying online?
RQ4: What are the factors that influence impulsive buying online?

1.4 Research Objectives

The following research objectives were identified based on the above research questions.

RO1: To identify the influence of online shopping experience on impulsive buying online
RO2: To identify the influence of online shopping experience on attitudinal loyalty
RO3: To identify the influence of attitudinal loyalty on impulsive buying online
RO4: To identify the factors that influence impulsive buying online

2 Literature Review

2.1 Theories and Models

2.1.1 Theory of Planned Behaviour (TPB)

Theory of Planned Behaviour is an extended version of Theory of Reasoned Actions by Ajzen and Fishbein (1980). TPB identifies three factors which influence human intentions and behaviours (Kaye et al, 2020; Ajzen, 1991). First, attitude towards behavior refers to learned tendencies that may respond either favourably or unfavourably. Attitudes are behavioural beliefs which are the expected outcomes. Second, subjective norms refers to opinions of reference groups and stands for one's perception on what majority thinks. Third, perceived behavioural control refers to subjective evaluation on level of hardships faced by an individual when engaged in a certain behavior. This looks into behaviours which are beyond boundaries of a person's will. (Rejali et al, 2023; Moon et al, 2021).
In the context of this study, favourable behavioural outcomes which are perceived advantages of online shopping experience determine positive attitudes and intentions to shop online, high approval of online fashion shopping determine strong subjective norms and favourable factors perceived by individuals

determine perceived behavioural control which influence intentions and behaviours favourably for online fashion shopping in Sri Lanka (Bezirgani and Lachapelle, 2021).

2.1.2 Technology Acceptance Model (TAM)

Technology Acceptance Model recognizes three factors, namely people's attitudes towards using technology, behavioural intentions to use technology and actual usage of technology which are impacted by perceived usefulness and perceived ease of use (Bashir et al, 2022; Davis, 1989). According to this model, it can be identified that behavioural intentions are influenced by attitude, which is influenced by perceived usefulness and perceived ease of use, and perceived usefulness (Rejali et al, 2023). Attitude refers to an individual's emotional state on using technology whilst perceived usefulness refers to rise in productivity through the utilization of technology (Bashir et al, 2022). It is identified that positive attitudes and high perceived usefulness results in high behavioural intention to utilize technology (Rejali et al, 2023).

Researchers have used this model frequently to evaluate online consumer behavior (Sen, 2019). In the context of this study, it is understood how attitude and perceived usefulness impacts an online context such as website in this situation. The model was used by Hausman and Siekpe (2009) to examine the impact of website interface on consumer online intention and found that website interface influenced online fashion shopping behavior (Sen, 2019).

2.1.3 Unified Theory of Acceptance and Use of Technology (UTAUT)

Unified Theory of Acceptance and Use of Technology is a framework used to examine psychological factors which predict behavioural intentions to use technology (Rejali et al, 2023; Piarna et al, 2020; Venkatesh et al, 2003). The model identifies how users wish to use information systems against how they actually use it (Rehman et al, 2022). UTAUT consists of four factors of behavioural intentions, namely effort expectancy, performance expectancy, facilitating conditions and social influence (Rehman et al, 2022).

UTAUT is a valid framework which helps to analyse purchase intention and actual purchase in the online shopping context (Piarna et al, 2020; Celik, 2016). In the context of this study, the four factors can be discussed. First, performance expectancy in online shopping discusses the extent of customer's perceptions that online fashion shopping would allow them to attain gains in performance. Second, effort expectancy indicates the ease of use where many customers see online fashion shopping as effortless and easy to use. Third, social influence described customer's perception on online fashion shopping driven by important others. Fourth, facilitating conditions regards the extent by which customers perceive that they have the means to support online fashion shopping (Rejali et al, 2023; Rehman et al, 2022).

2.1.4 Stimulus-Organism-Response Model (S-O-R)

The Stimulus-Organism-Response model studies relationships between environmental stimuli, cognitive and affective organism and consumers' behavioral responses (Kimiagari and Asadi Malaf, 2021). The Stimulus-Organism-Response model discuss three aspects, namely environmental or stimulus, which triggered following consumer behaviours and responses, the organism that responds and the actual responses. (Chen and Yao, 2018). The model is widely used to examine online impulsive behaviours of an individual (Kimiagari and Asadi Malaf, 2021; Chen and Yao, 2018).

In the context of this study, online shopping experience can be examined. The model propose that features of an online environment provoke individuals to react in an online environment by all explicit actions such as desire to discover and feel the desire. The model has been used by many researchers (Rahman et al, 2023; Liu et al, 2021) to investigate on how ecommerce platforms affect consumer behaviours. In the context of online impulsive buying, the response aspect of the model helps to understand results of consumers' reaction to drivers of online impulsive buying when engaged in online fashion shopping (Kimiagari and Asadi Malaf, 2021). Furthermore, the model speaks that perceived value gained through online shopping experience (stimulus), such aesthetic appeal and functionality has the tendency to promote customer loyalty in the form of attitudinal loyalty and behavioural loyalty (Liu et al, 2021).

3 Conceptual Framework

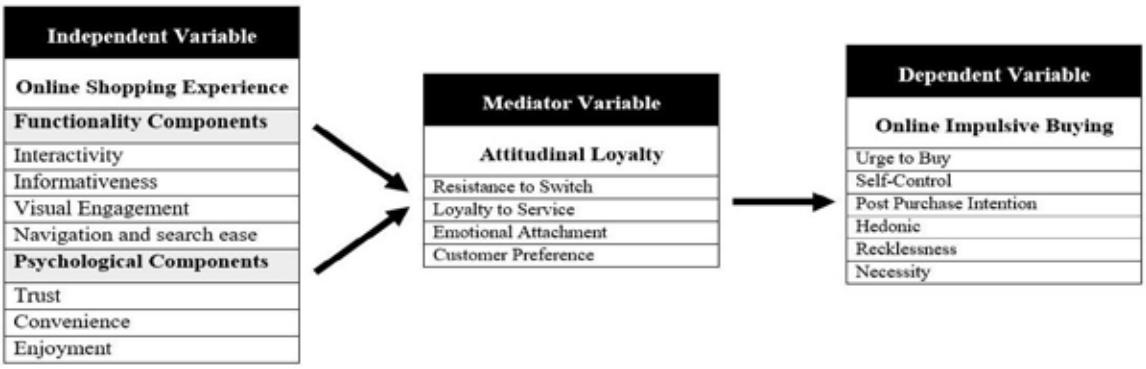


Figure 1: Conceptual Framework (Author developed)

3.1 Online Customer Shopping Experience

Meyer and Schwager (2007) defined customer shopping experience as "customer internal and subjective resonance to any direct or indirect interaction with the company's products, service or brands" which is a definition followed by many researchers (Barari et al, 2020).

Online customer shopping experience is defined as the cumulative psychological effect of customer interactions on a virtual environment (Bleier et al., 2019). It is important for retailers to deliver a good and complete online customer shopping experience that offers customers both frictionless and pleasurable shopping experiences (Kahn, 2018). There are various dimensions for online custosmer shopping experience identified and categorized as either functional or psychological (Gulfraz et al, 2022). Information, time, beyond locational barriers and convenience experienced when shopping online empowers customers (Barari et al, 2020).

In previous studies by Gao and Yee (2022) and Logianti Diani et al (2021) it was reavealed that online shopping experience has a significant positive impact on online impulsive buying. Furthermore Gulfraz et al (2022) discuss in depth how online shopping experience in terms of functional and psychological showcase a significant positive impact on online impulsive buying. Furthermore, theories and models such as Technology Acceptance Model, Unified Theory of Acceptance and Use of Technology, Theory of Planned Behaviours and Stimulus Organism Response Model justifies the positive impact of online shopping experience on online impulsive buying in both functional and psychological dimensions.

H1: Online shopping experience has a significant positive impact on online impulsive buying.

3.2 Attitudinal Loyalty

Attitudinal loyalty is defined as psychological disposition one has towards a particular brand (Kaur et al, 2019; Husain, 2018). Attitudinal loyalty is the strength of consumers' attitudes towards a particular brand that resist them from switching to rival brands (Choi and Kim, 2020) but pay higher process despite the lower prices offered by competitors (Gulfraz et al, 2022, Casper Ferm and Thaichon, 2021). Hence, acts as brand advocates (Nyagadza at al., 2022, Casper Ferm and Thaichon, 2021). This is due to the vital role played by trust (Alalwan et al., 2018) which drives customers to continue their relationship with the particular business (Gulfraz et al, 2022).

Gulfraz et al (2022) and Cachero-Martinez et al (2021) it was proved that online shopping experience has a significant positive impact on online impulsive buying.

Thereby, attitudinal loyalty indicates cognitive and affective components of relationships with high emotional attachment (Gulfraz et al, 2022) which results in stronger long-term relationships, positive word-of-mouth and recommendations to others (Casper Ferm and Thaichon, 2021). Furthermore, theories and models such as Technology Acceptance Model, Unified Theory of Acceptance and Use of Technology, Theory of Planned Behaviours and Stimulus Organism Response Model justifies the positive impact of online shopping experience on attitudinal loyalty in both functional and psychological dimensions.

H2: Online shopping experience has a significant positive impact on attitudinal loyalty.

3.3 Impulsive buying

Piron (1991) defined impulsive buying as "a purchase that is unplanned, the result of an exposure to a stimulus, and decided on the spot". Impulsive buying is identified as a behavior driven by unexpected, powerful and persistent desire to make a purchase (Wu et al, 2020).

Online impulsive buying can be described as a sudden purchases made online with no pre-planned intentions (Ampadu et al, 2022). Online impulsive buying are hedonic purchase practices which does not take into account adequate information about products, possible consequences and evaluation of alternatives (Gulfraz et al, 2022; Hashmi et al, 2019).

Buying decisions are identified as impulsive if they meet three criteria. First, decision must be spontaneous and unplanned. Second, buyers have little to no concern on consequences of purchase decision. Third, purchase decisions are evoked by irresistible temptations like dispositional, situational or socio-demographic (Gulfraz et al, 2022) which cause strong emotional attachments (Gulfraz et al, 2022; Spiteri Cornish, 2020).

It is important for online retailers to be aware on the importance of well-designed and managed websites in order to provide a pleasant online shopping experience for customers which affects customer shopping behavior which in turn would result in online impulsive buying (Gulfraz et al, 2022).

In previous studies by Gulfraz et al (2022) and Lazim et al (2020) it was reavealed that attitudinal loyalty has a significant positive impact on online impulsive buying. Furthermore, theories and models such as Technology Acceptance Model, Unified Theory of Acceptance and Use of Technology, Theory of Planned Behaviours and Stimulus Organism Response Model justifies the positive impact of attitudinal loyalty on online impulsive buying.

H3: Attitudinal loyalty has a significant positive impact on online impulsive buying.

It is taken into account that due to the functional and psychological aspect, attitudinal loyalty would mediate between the influences of online shopping experience on online impulsive buying. This is because online shopping experience tend to have an impact of customers' attitude towards loyalty for a particular brand which in turn would influence shopping behaviours like online impulsive buying (Gulfraz et al, 2022).

Gulfraz et al (2022) Lazim et al (2020) and Cachero-Martinez et al (2021), in their studies proved that attitudinal loyalty successfully mediates the relationship between online shopping experience and online impulsive buying. These empirical studies emphasize the significance of the mediation of attitudinal loyalty in the relationship between online shopping experience and online impulsive buying.

H4: Attitudinal loyalty mediates the relationship between online shopping experience and online impulsive buying.

4 Research Methodology

4.1 Research Philosophy

The study was based on philosophy of positivism since the research was carried and verified with the support of empirical observations and logical analysis to draw conclusions (Bell et al, 2022) on the impact of online shopping experience and attitudinal loyalty on online impulsive buying (Kankam, 2019).

4.2 Research Approach

This study was based on the deductive approach since the study focuses on testing the relationships between main variables (Bougie and Sekaran, 2019; Bell et al, 2022).

4.3 Research Instrument

This study was based on empirical investigations of proposed concepts presented in the conceptual framework (Figure 1) and was conducted based on quantitative research technique influenced by positivistic paradigm which allowed data to be collected in numerical form and be categorized, ranked and measured in different units of measurement (Bell et al, 2022). A cross sectional survey was used to collect data from the respondents. Online shopping experience was operationalized using 9 items using the measures used by Gulfraz et al (2022) on a 5 point Likert Scale ranging from Strongly Disagree (1) to Strongly Agree (5). Attitudinal Loyalty was operationalized using 4 items using measures adopted by Parasuramen et al., (1988)

& Kaur et al., (2019) on a 5 Point Likert Scale ranging from Strongly Disagree (1) to Strongly Agree (5). Online impulsive buying was operationalized using 6 items adopted from the measures used by Cornish, (2020) on a 5 point Likert Scale ranging from Strongly Disagree (1) to Strongly Agree (5).

4.4 Sample and Data

The simple random sampling method was used to conduct the research due to its ability to create a representative sample and to eliminate biasness in the research which in turn would enable more generalized findings which reflect the diverse characteristics and attributes of the targeted population of this study (Bell et al., 2022).

Data was gathered from Sri Lankan population who engage in online shopping. A pilot study was initially carried out with 30 respondents. Thereafter, the questionnaire was distributed amongst 397 participants, where 242 valid responses were gathered. The response rate was 61%, hence is acceptable.

Thereafter, SPSS version 26 was used to analyze data which were gathered to draw conclusions through tests such as normality, regression, mediation and hypothesis testing.

5 Data Representation and Analysis

5.1 Data Cleaning

There were no missing values.

5.2 Normality of Data

5.2.1 5.2.1 Skewness and Kurtosis

The data collection is surmised as normal when skewness is between -1 and +1 and kurtosis values are between -3 and +3 (Bougie and Sekaran, 2019). Thereby, it can be identified that data distribution is normal since skewness of variables falls between -1 to +1 and kurtosis values of variables range between -3 to +3.

5.2.2 Kolmogorov - Smirnov (K-S) and Shapiro-Wilk Test

Furthermore, the data collection is surmised as normal if p value of K-S and Shapiro-Wilk tests, is greater than 0.05. The p value of K-S and Shapiro-Wilk tests, is greater than 0.05. Thereby, data distribution is as normal.

5.3 Multicollinearity Testing

The Variance Inflation Factor (VIF) is 2.057 which is less than the threshold value of 5 and Tolerance values is 0.486 greater than the threshold value of 0.2 (Seneviratne and Cooray, 2019). Hence, represents the absence of multicollinearity as it exists when there is high intercorrelations among independent variables (Shrestha, 2020). This absence determine normal distribution.

5.4 Validity

5.4.1 Sample Adequacy

Table 1: KMO and Bartlett's Test (Author developed)

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.904
Bartlett's Test of Sphericity	Approx. Chi-Square	8564.086
	Df	820
	Sig.	.000

Kaiser-Meyer-Olkin is 0.904 which is above the threshold value of 0.5 and the p value of Bartletts test of Sphericity is less than 0.05. Therefore, the sample size can be regarded as adequate (Goni et al., 2020).

5.4.2 Convergent Validity

Inter-item correlations is expected to be greater than 0.3 to ensure convergent validity of sample (Campbell and Fiske, 1959) and main variables satisfy the requirement. Factor analysis was performed and three main factors were formed followed by rotation of all factors which proved that convergent validity exists.

5.4.3 Construct Validity

A pilot study was conducted with 30 respondents and Cronbach Alpha value was 0.912. Therefore, construct validity exists since the value was greater than threshold value of 0.7.

5.4.4 Discriminant Validity

Wilks Lamda test indicates significant relationship between independent and dependent variables as p value is less than 0.05. Discriminant validity is high as Eigen values are less than 1 yet smaller than Eigen value. Discriminant validity exists since Eigen values are less than 1 (Park and Kim, 2021).

5.5 Reliability Analysis

Cronbach alpha was used to determine reliability of the research instrument. A cronbach alpha value greater than threshold value of 0.7 proves that reliability exists. The cronbach alpha for the online shopping experience is 0.908, for attitudinal loyalty is 0.933 and for online impulsive buying is 0.779. Thus, the instrument is reliable since cronbach alpha value exceeds 0.7.

5.6 Correlation Analysis

Correlation analysis revealed that the p values for the relationships between: online shopping experience and attitudinal loyalty; online shopping experience and impulsive buying; and attitudinal loyalty and impulsive buying, were all 0.000, indicating a statistically significant relationship. Furthermore, Pearson correlation between online shopping experience and attitudinal loyalty is 0.717 which is greater than 0.6 which is the threshold value (Cohen, 1988). Thereby, indicates a strong positive correlation. Pearson correlation between online shopping experience and impulsive buying is 0.534 which indicates a positive correlation. Pearson correlation between impulsive buying and attitudinal loyalty is 0.566 which indicates positive correlation.

Table 2: Model Summary (Author developed)

Relationship	Pearson Correlation (R Value)	Strength of the Relationship
Online Shopping Experience and Attitudinal Loyalty	0.717	High
Attitudinal Loyalty and Impulsive Buying	0.566	High
Online Shopping Experience and Impulsive Buying	0.534	High

5.7 Regression Analysis

Regression analysis was conducted to identify the significance of main variables. R square value indicated a 35.4% variance in dependent variable (Table 3). The P value is less than 0.05 which shows that impact of attitudinal loyalty and online shopping experience on online impulsive buying is significant. Attitudinal loyalty makes highest contribution to online impulsive buying since beta value is 0.377 (Table 5). Thus reveals the strength of the relationship. This agrees with empirical findings by Molinillo et al., (2020) and Gulfraz, et al., (2022).

Table 3: Model Summary (Author developed)

Model Summary[b]				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.595[a]	.354	.349	.56831
a. Predictors: (Constant), Online Shopping Experience, Attitudinal Loyalty				
b. Dependent Variable: Online Impulsive Buying				

Table 4: ANOVA (Author developed)

ANOVA[a]						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	42.291	2	21.146	65.471	.000[b]
	Residual	77.192	239	.323		
	Total	119.483	241			
a. Dependent Variable: Impulsive Buying						
b. Predictors: (Constant), Online Shopping Experience, Attitudinal Loyalty						

Table 5: Coefficients (Author developed)

Coefficients[a]								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	.817	.286		2.857	.005		
	Attitudinal Loyalty	.358	.071	.377	5.061	.000	.486	2.057
	Online Shopping Experience	.384	.109	.263	3.528	.001	.486	2.057
a. Dependent Variable: Impulsive Buying								

5.8 Mediation Analysis

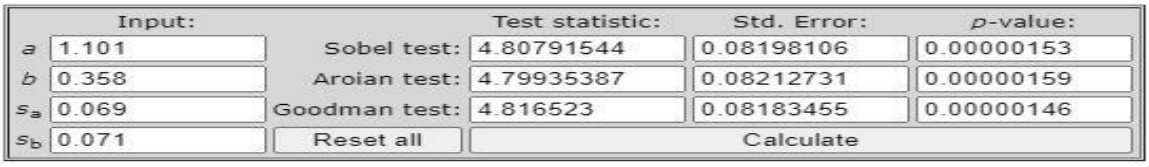


Figure 2: Sobel Test (Author developed)

The p value is significant and null hypothesis is rejected. Thereby, attitudinal loyalty has a significant mediator effect. The sobel's test revealed that attitudinal loyalty mediates the relationship between online shopping experience and online impulsive buying,

6 Conclusion, Limitations and Recommendations

6.1 Research Findings

This study was initiated with the aim of understanding the impact of online shopping experience and attitudinal loyalty on online impulsive buying in the online fashion retail industry in Sri Lanka. The study was based on four hypotheses. The data was collected and analyzed through various test which were conducted using the SPSS software. Following are the findings of this study.

H1: Online shopping experience has a significant positive impact on online impulsive buying.

The correlation analysis revealed that online shopping experience and online impulsive buying have a relationship based on the p value of 0.000 and positive relationship with a correlation value of 0.534. The coefficient table displays a p value of 0.000 between online shopping experience and online impulsive buying. This shows that the relationship between online shopping experience and online impulsive buying is significant. This, supports empirical findings by Gulfraz, et al., (2022) and Anshu et al., (2022); Pandey and Chawla, (2018). Therefore, this hypothesis is accepted.

H2: Online shopping experience has a significant positive impact on attitudinal loyalty.

The correlation analysis revealed that online shopping experience and attitudinal loyalty have a relationship based on the p value of 0.000 and a strong and positive relationship with a correlation value of 0.717. The coefficient table displays a p value of 0.000 between online shopping experience and attitudinal loyalty. This shows that the relationship between online shopping experience and attitudinal loyalty is significant. This supports empirical findings by Cachero-Martínez and Vazquez-Casielles, (2021); and Bleier et al., (2019). Therefore, this hypothesis is accepted.

H3: Attitudinal loyalty has a significant positive impact on online impulsive buying.

The correlation analysis revealed that attitudinal loyalty and online impulsive buying have a relationship based on the p value of 0.000 and positive relationship with a correlation value of 0.566. The coefficient table displays a p value of 0.000 between attitudinal loyalty and online impulsive buying. This shows that the relationship between attitudinal loyalty and online impulsive buying is significant. This supports empirical findings by Gulfraz, et al., (2022); Cachero-Martínez and Vazquez-Casielles, (2021); Iyer et al., (2020). Therefore, this hypothesis is accepted.

H4: Attitudinal loyalty mediates the relationship between online shopping experience and online impulsive buying.

The Sobels test as presented in figure 2, disclosed that attitudinal loyalty significantly mediates the relationship between online shopping experience and online impulsive buying. Hence, agrees with empirical findings by Gulfraz, et al., (2022) and Anshu et al., (2022). Therefore, this hypothesis is accepted.

The findings disclosed that attitudinal loyalty makes highest contribution to dependent variable, online impulsive buying because beta value is 0.377 (Table 5). It was also revealed through regression analysis that online impulsive buying predicts attitudinal loyalty and online shopping experience by 34.9%. The factors which contribute more for attitudinal loyalty are loyalty to service, emotional attachment and customer preference. This agrees with empirical findings by Gulfraz, et al., (2022); Cachero-Martínez and Vazquez-Casielles, (2021); Iyer et al., (2020).

The following analysis was proposed by the study based on the regression equation $Y = a + b x$ to identify how changes in the independent variables are related with changes in the dependent variable.

Online Impulsive Buying = 0.817 + 0.384 (Mean of Online Shopping Experience) * 0.358 (Mean of Attitudinal Loyalty).

The study also revealed the existence of a significant difference between online impulsive buying behaviours of men and women through the Levene's test which resulted in a p value of 0.041 which is less than 0.05. This, supports empirical findings by Fenton-O'Creevy and Furnham, (2020) and Santini et al., (2018). Furthermore, online impulsive buying is also influenced by other factors such as age and income (Fenton-O'Creevy and Furnham, 2020).

Furthermore, the demographic analysis conducted through crosstabulations resulted in the following. The results disclosed that out of the sample most respondents spend between 1 to 3 hours online on a daily basis out of which women spend the most compared to male respondents. It was also disclosed that the respondents between the age group of 18 to 25 spent the most time online and also it was mostly those who fit this age group who spent 1 to 3 hours online followed by those who are between 26 to 40 years of age. Furthermore it was revealed that out of the sample most respondents who shop online earned a monthly income between Rs. 25,000 to Rs. 50,000.

6.2 Limitations

Despite, disclosure of regression analysis indicating that online shopping experience and attitudinal loyalty predicts online impulsive buying by 35.4% it cannot be denied that there are other factors such as time, customer's mood, customer's situation and marketing by retailers which influence online impulsive buying. Therefore, there are other drivers of online impulsive buying and not just online shopping experience and attitudinal. Furthermore, this was a quantitative study which was based on a sample and does not represent the entire target population. The study utilized simple random sampling technique and most respondents were from Colombo and nearby cities and most respondents were female. Hence the results are less generalizable.

6.3 Recommendations

This section presents recommendations for online shopping experience and attitudinal loyalty which were suggested based on data gathered during the survey process of this study which are supported by previous studies conducted by McLean and Wilson, 2016; Anshu et al., 2022; Zhang et al., 2018.

This study recommends online fashion retail platforms to minimize customer inconveniences by enhancing navigation and search facilities, ensure faster online purchase processes, to provide rich information on products for easier and successful transactions and to utilize cookie data for a more personalized customer experience in order to improve online shopping experiences for customers.

This study recommends online fashion retail platforms to enhance brand image to make the ecommerce platform always the first choice for purchases, to continue good service practices while further improving service quality and to provide loyalty benefits in order to prevent customer switch and rise attitudinal loyalty.

6.4 Direction for Future Research

This study focused only on B2C ecommerce platforms yet future studies could examine online shopping experience on C2C and B2B platforms. Also, future studies could examine areas such as online shopping experience and self-control on online impulsive buying or online shopping experience and attitudinal loyalty on repurchase intention. Furthermore, it is encouraged for future research studies to use a different sampling technique in order to ensure the collection of more generalized data.

References

Alexander Bleier, Colleen M. Harmeling, and Robert W. Palmatier. 2018. Creating Effective Online Customer Experiences. *Journal of Marketing*, 83, 2 (December 2018). DOI: https://doi.org/10.1177/0022242918809930.

Amjad Ur Rehman et al. 2022. Does e-shopping service quality enhance customers' e-shopping adoption? An extended perspective of unified theory of acceptance and use of technology. PLoS One, 17, 2 (February, 2022), 1-4. DOI: https://doi.org/10.1371%2Fjournal.pone.0263652.

Ana Bezirgani and Ugo Lachapelle. 2021. Online grocery shopping for the elderly in Quebec, Canada: The role of mobility impediments and past online shopping experience. *Travel Behaviour and Society*, 25 (October 2021), 134-135. DOI: https://doi.org/10.1016/j.tbs.2021.07.001.

Angela V. Hausman and Jeffrey S. Siekpe. 2009. The effect of web interface features on consumer online purchase intentions. *Journal of Business Research*, 62, 1 (January 2009), 6-8. DOI: https://doi.org/10.1016/j.jbusres.2008.01.018.

Beomjoon Choi and Hyun Sik Kim. 2020. Customer-to-customer interaction quality, promotion emotion, prevention emotion and attitudinal loyalty in mass services. *Journal of Service Theory and Practice*, 30, 3 (June 2020), 258-264. DOI: https://doi.org/10.1108/JSTP-08-2019-0172.

Brighton Nyagadza et al. 2022. Effect of hotel overall service quality on customers' attitudinal and behavioural loyalty: perspectives from Zimbabwe. *Tourism Critiques: Practice and Theory*, 3, 1 (May 2022), 43-47. DOI: http://dx.doi.org/10.1108/TRC-12-2021-0026.

Chia-Chen Chen and Jun-You Yao. 2018. What drives impulse buying behaviors in a mobile auction? The perspective of the Stimulus-Organism-Response model. *Telematics and Informatics*, 35, 5 (August 2018), 1250-1251. DOI: https://doi.org/10.1016/j.tele.2018.02.007.

Christopher Meyer and Andre Schwager. (2007). Understanding Customer Experience. *Harvard Business Review.* 85, 2 (March 2007), 116-126. DOI: https://www.researchgate.net/publication/6460601_Understanding_Customer_Experience.

Donald T. Campbell and Donald W. Fiske. 1959. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 2 (March 1959). DOI: https://doi.org/10.1037/h0046016.

Emma Bell, Alan Bryman, and Bill Harley. 2022. *Business Research Methods*, (6th ed.). Oxford: Oxford University Press.

Fernando D. O. Santini et al. 2018. Antecedents and consequences of impulse buying: a meta-analytic study. *RAUSP Management Journal*, 54, 2 (September 2018), 179-185. DOI: http://dx.doi.org/10.1108/RAUSP-07-2018-0037.

Francis Piron. 1991. Defining Impulse Purchasing. Advances in Consumer Research, 18, 509-514. https://www.acrwebsite.org/volumes/7206/volumes/v18/NA-18/full.

Fred D. Davis. 1989. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *IT Usefulness and Ease of Use*, 13, 3 (September 1989), 319-325. DOI: http://dx.doi.org/10.2307/249008.

Fuzhen Liu et al. 2021. Listening to online reviews: A mixed-methods investigation of customer experience in the sharing economy. *Decision Support Systems*, 149 (October 2021), 2-5. DOI: https://doi.org/10.1016/j.dss.2021.113609.

Gopalkrishnan R. Iyer et al. 2019. Impulse buying: a meta-analytic review. *Journal of the Academy of Marketing Science*, 48 (July 2019), 385-397. DOI: https://doi.org/10.1007/s11747-019-00670-w.

Graeme McLean and Alan Wilson. 2016. Evolving the online customer experience ... is there a role for online customer support? *Computers in Human Behavior*, 60 (July 2016), 607-609. DOI: https://doi.org/10.1016/j.chb.2016.02.084.

Hakan Celik. 2016. Customer online shopping anxiety within the Unified Theory of Acceptance and Use Technology (UTAUT) framework. *Asia Pacific Journal of Marketing and Logistics*, 28, 2 (March 2016). DOI: https://doi.org/10.1108/APJML-05-2015-0077.

Hong Zhang, Ling Zhao and Sumeet Gupta. (2018). The role of online product recommendations on customer decision making and loyalty in social shopping communities. *International Journal of Information Management*, 38 (February 2018), 159-161. DOI: https://www.sciencedirect.com/science/article/pii/S0268401216306971.

Hong Zhang, Ling Zhao, and Sumeet Gupta. 2018. The role of online product recommendations on customer decision making and loyalty in social shopping communities. *International Journal of Information Management*, 38 (February 2018), 159-161. DOI: https://doi.org/10.1016/j.ijinfomgt.2017.07.006.

Icek Ajzen and Martin Fishbein. 1977. Attitude-Behavior Relations: A Theoretical Analysis and Review of Empirical Research. *Psychological Bulletin*, 84, 5 (September 1977). DOI: https://doi.org/10.1037/0033-2909.84.5.888.

Icek Ajzen. 1991. The Theory of Planned Behavior. *Organizational Behavior and Human Decision Processes*, 50 (December 1991), 179-184. DOI: https://doi.org/10.1016/0749-5978(91)90020-t.

Ilmudeen Aboobucker. 2019. Factors Influencing Consumers' Trust on E-commerce Adoption in Sri Lanka. (March 2019). DOI: http://dx.doi.org/10.2139/ssrn.3344618.

Ing-Long Wu, Mai-Lun Chiu, and Kuei-Wan Chen. 2020. Defining the determinants of online impulse buying through a shopping process of integrating perceived risk, expectation-confirmation model, and flow theory issues. *International Journal of Information Management*, 52 (June 2020), 1-5. DOI: https://doi.org/10.1016/j.ijinfomgt.2020.102099.

Jacob Cohen. (1992). A power primer. *Psychological Bulletin*, 112, 1 (September 2007), 155–159. DOI: https://doi.org/10.1037/0033-2909.112.1.155.

Jayani C. Athapaththu and D. Kulathunga. 2018. Factors Affecting Online Purchase Intention: Effects of Technology and Social Commerce. *International Business Research,* 11, 10 (September 2018), 111-112. DOI: https://doi.org/10.5539/ibr.v11n10p111.

JiHyo Moon, Yunseon Choe, and HakJun Song. 2021. Determinants of Consumers' Online/Offline Shopping Behaviours during the COVID-19 Pandemic. *International Journal of Environmental Research and Public Healt*h, 18, 4 (February 2021), 5. DOI: https://doi.org/10.3390%2Fijerph18041593.

Kapila Fonseka et al. 2020. Literature Review of Technology Adoption Models at Firm Level; Special Reference to E-Commerce Adoption. *Global Journal of Management and Business Research: B Economics and Commerce*, 20, 6 (July 2020), 1-3. DOI: http://dx.doi.org/10.34257/GJMBRBVOL20IS6PG1.

Kapila Fonseka, Adam A. Jaharadak, and Murali Raman. 2022. Impact of E-commerce adoption on business performance of SMEs in Sri Lanka; moderating role of artificial intelligence. *International Journal of Social Economics*, 49, 10 (May 2022), 1518-1521. DOI: https://doi.org/10.1108/IJSE-12-2021-0752.

Kumari Anshu, Loveleen Gaur, and Gurmeet Singh. 2022. Impact of customer experience on attitude and repurchase intention in online grocery retailing: A moderation mechanism of value Co-creation. *Journal of Retailing and Consumer Services*, 64 (January 2022), 1-5. DOI: https://doi.org/10.1016/j.jretconser.2021.102798.

Lara S. Cornish. 2020. Why did I buy this? Consumers' post-impulse-consumption experience and its impact on the propensity for future impulse buying behaviour. *Journal of Consumer Behaviour*, 19, 1 (January 2020), 36-46. DOI: https://doi.org/10.1002/cb.1792.

Lars-Erik C. Ferm and Park Thaichon. 2021. Customer pre-participatory social media drivers and their influence on attitudinal loyalty within the retail banking industry: A multi-group analysis utilizing social exchange theory. *Journal of Retailing and Consumer Services*, 61 (July 2021), 2-4. DOI: https://doi.org/10.1016/j.jretconser.2021.102584.

Malin Sundström, Sara Hjelm-Lidholm and Anita Radon. (2019). Clicking the boredom away – Exploring impulse fashion buying behavior online. *Journal of Retailing and Consumer Services*, 47, (March 2019), 150-152. DOI: https://www.sciencedirect.com/science/article/pii/S096969891830167X?via%3Dihub.

Mark Fenton-O'Creevy and Adrian Furnham. 2019. Money Attitudes, Personality and Chronic Impulse Buying. *Applied Psychology*, 69, 4 (August 2019). DOI: https://doi.org/10.1111/apps.12215.

Mohammed D. Goni et al. 2020. Development and validation of knowledge, attitude and practice questionnaire for prevention of respiratory tract infections among Malaysian Hajj pilgrims. *BMC Public Health*, 20 (March 2020), 2-8. DOI: https://doi.org/10.1186/s12889-020-8269-9.

Mojtaba Barari, Mitchell Ross, and Jiraporn Surachartkumtonkun. (2020). Journal of Retailing and Consumer Services. *Journal of Retailing and Consumer Services*, 53 (March 2020), 1-5. DOI: https://www.sciencedirect.com/science/article/pii/S0969698919307039.

Mojtaba Barari, Mitchell Ross, and Jiraporn Surachartkumtonkun. 2020. Negative and positive customer shopping experience in an online context. *Journal of Retailing and Consumer Services*, 53 (March 2020), 1-5. DOI: https://doi.org/10.1016/j.jretconser.2019.101985.

Muhammad B. Gulfraz et al. 2022. Understanding the impact of online customers' shopping experience on online impulsive buying: A study on two leading E-commerce platforms. *Journal of Retailing and Consumer Services*, 68 (September 2022), 3-11 DOI: https://doi.org/10.1016/j.jretconser.2022.103000.

Muhammad S. Rahman et al. 2023. The new wave of AI-powered luxury brands online shopping experience: The role of digital multisensory cues and customers' engagement. *Journal of Retailing and Consumer Services*, 72 (February 2023), 1-5. DOI: https://doi.org/10.1016/j.jretconser.2023.103273.

Noora Shrestha. 2020. Detecting Multicollinearity in Regression Analysis. *American Journal of Applied Mathematics and Statistics*, 8, 2 (June 2020), 39-42. DOI: http://dx.doi.org/10.12691/ajams-8-2-1.

Puneet Kaur et al. 2019. Attitudinal and Behavioral Loyalty Toward Virtual Goods. *Journal of Computer Information Systems*, 61, 2 (February 2019). DOI: https://doi.org/10.1080/08874417.2019.1568844.

Ravi Raina et al. 2020. *E-commerce Readiness Assessment Report Sri Lanka: March 2020*. London: Commonwealth Secretariat. https://www.srilankabusiness.com/ebooks/sri-lanka-e-commerce-readiness.pdf.

Rian Piarna, F. Fathurohman and Nunu N. Purnawan. (2020). Understanding online shopping adoption: The unified theory of acceptance and the use of technology with perceived risk in millennial consumers context. *Scientific Journal in Accounting and Management*, 17, 1 (March 2020) 51-54. DOI: https://www.semanticscholar.org/paper/Understanding-online-shopping-adoption%3A-The-unified-Piarna-Fathurohman/e26d6dfdcaa8c407a6c8d488c9056a3944d6160f.

Rian Piarna, Ferdi Fathurohman, and Nunu N. Purnawan. 2020. Understanding online shopping adoption: The unified theory of acceptance and the use of technology with perceived risk in millennial consumers context. *Scientific Journal in Accounting and Management*, 17, 1 (January 2020), 51-54. DOI: http://dx.doi.org/10.31106/jema.v17i1.5050.

Roger Bougie and Uma Sekaran. 2019. *Research Methods For Business: A Skill Building Approach*, (8th ed.). U.S. Wiley.

Salman Kimiagari and Neda S. A. Malafe. 2021. The role of cognitive and affective responses in the relationship between internal and external stimuli on online impulse buying behavior. *Journal of Retailing and Consumer Services*, 61 (July 2021), 2-7. DOI: https://doi.org/10.1016/j.jretconser.2021.102567.

Sebastian Molinillo et al. (2020). The impact of affective and cognitive app experiences on loyalty towards retailers. *Journal of Retailing and Consumer Services*, 54 (May 2020). DOI: https://www.sciencedirect.com/science/article/abs/pii/S0969698919304710.

Seth Ampadu et al. 2022. Online personalized recommended product quality and e-impulse buying: A conditional mediation analysis. *Journal of Retailing and Consumer Services*, 64 (January 2022), 2-4. DOI: https://doi.org/10.1016/j.jretconser.2021.102789.

Sherrie-Anne Kaye et al. 2020. Assessing the feasibility of the theory of planned behaviour in predicting drivers' intentions to operate conditional and full automated vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour*, 74 (October 2020), 174. DOI: https://doi.org/10.1016/j.trf.2020.08.015.

Shweta Pandey and Deepak Chawla. 2018. Online customer experience (OCE) in clothing e-retail: Exploring OCE dimensions and their impact on satisfaction and loyalty – Does gender matter? *International Journal of Retail & Distribution Management*, 46, 3 (April 2018), 323-331. DOI: https://doi.org/10.1108/IJRDM-01-2017-0005.

Silvia Cachero-Martínez and Rodolfo Vázquez-Casielles. 2021. Building consumer loyalty through e-shopping experiences: The mediating role of emotions. *Journal of Retailing and Consumer Services*, 60 (May 2021), 1-5. DOI: https://doi.org/10.1016/j.jretconser.2021.102481.

Sina Rejali et al. 2023. Comparison of technology acceptance model, theory of planned behavior, and unified theory of acceptance and use of technology to assess a priori acceptance of fully automated vehicles. *Transportation Research Part A: Policy and Practice*, 168 (February 2023), 3-6. DOI: https://doi.org/10.1016/j.tra.2022.103565.

Tayyeba Bashir et al. 2022. An assessment of influential factors developing the intention to use social media sites: A technology acceptance model based approach. *Frontiers in Psychology*, 13 (October 2022), 2-3. DOI: https://doi.org/10.3389%2Ffpsyg.2022.983930.

Viswanath Venkatesh et al. 2003. User Acceptance of Information Technology: Toward a Unified View. MIS Quarterly, 27, 3 (September 2003). DOI: https://doi.org/10.2307/30036540.

Xian Gao and Choy-Leong Yee. (2022). Factors Affecting Consumers' Impulse Buying Behaviour In Online Shopping:A Systematic Literature Review. *Journal of Positive School Psychology*, 6, 10, 4075-4097. DOI: https://journalppw.com/index.php/jpsp/article/view/14144/9169.

Impact of Social Media Communication and Brand Equity on Consumer Perceptions of Brands in the Fast-Food Industry

Ashweka Vasanthakumar, IIT Business School, Informatics Institute of Technology, Colombo, Sri Lanka, ashweka.20191313@iit.ac.lk

Gayashini Ratnayake, IIT Business School, Informatics Institute of Technology, Colombo, Sri Lanka, gayashini.r@iit.ac.lk

ABSTRACT

Traditional media have experienced an immense transformation by moving into social media platforms over the past decade which has become a major platform for brands to have the opportunity to interact with consumers. Social media has now taken a place in impacting consumer perceptions and decisions. This study examines how social media communication content influences customer decisions, providing valuable insights for brand and marketing managers. Additionally, it offers strategies to optimize content, reduce advertising costs, and attract potential customers effectively. Hence, this study is to explore the influence of social media communications and brand equity on consumer perceptions of brands in the fast food industry in Sri Lanka. This paper is assessed through a mixed method and deductive approach, having the primary data collected through data collection questionnaire form using quantitative analysis with the use of the SPSS tool and deriving secondary data from online publications and journals to analyze the hypothesis for this study. With 244 responses through the primary data collection and the findings from the secondary data, the study reveals a positive outcome in the relationship on the impact of social media communication and brand equity on consumer perceptions of the brands.

KEYWORDS

Social media communication, Brand equity, Consumer perceptions, Brand attitude, Purchase intention, Fast-food industry.

1 Introduction

1.1 Background

The fast-paced evolution in the media industry has a significant impact on users and brands (Chadwick 2021). Social media plays an important role in how customers pursue a brand and it also shows how businesses and brands reach out to their customers (Cheng 2019). Social media has opened up several pathways for both local and global brand to lead their way to their potential customers (Chen et al 2019). Social media is becoming a preferred source for consumers to seek information instead of relying on traditional media channels such as television, radio, and magazines. This shift in consumer behavior is due to the emergence of social media platforms, which have transformed one-way communication into multi-dimensional, peer-to-peer interaction. This change is altering the traditional marketing communication landscape as consumers now play a more significant role in shaping brand communications, which were previously controlled by brand and marketing managers.

1.2 Research Gap

Several researchers have assessed the impact of electronic word of mouth on customer purchase intention (Majeed et al 2021; Zhao et al 2020). Further have assessed how user-generated social media and firm-created social media impact brand associations (Kaldeen 2019). There is a need to understand the specific social media feature on how that has impacted the decision on the customer purchase intention. There is a lack of understanding of how user-generated social media and firm-created social media as separate elements support the customer purchase intention (Majeed et al 2021). Brand equity factors are explored to determine their effect on customer purchase intention (Hasan et al 2021). Brand loyalty and brand awareness are also used as separate elements to understand the behavior of the customer purchase intention. Limited research has been identified in the context of brand equity acting as an indirect factor in accessing the decision-making process directly influenced by social media content. Most of the research conducted in Sri Lanka analyzes consumer behavior toward brand performance and how it could be used to improvise the franchising of the brand (Hanaysha 2022). The identified literature review has indicated that there is a positive relationship between social media communication and customer purchase intention. Predictively brand equity and customer purchase intention have also been concluded as positively

correlated variables (Majeed et al 2021). Lastly, this research aims to understand the effect of social media on two different aspects as user-generated social media and firm-created social media, and how that directly influences the customer purchase intention and keeps brand equity as a mediating factor which on the other hand gets influenced by social media and enhances the customer purchase intention benefits to the brand (Kaldeen 2019).

1.3 Scope of the Study

The scope of the study is limited to the choice of social media, which is Facebook and the influence of it on the brand McDonald's as a representation of the brand influence. As discussed above Facebook users are around 2/5 of the Sri Lankan population (Department of Census and Statistics Sri Lanka 2023). And a widely used site by companies to promote and communicate with their consumers. Further, the chosen brand Mcdonalds is one of the leading fast food chains in Sri Lanka which uses Facebook as the main medium of its communication.

1.4 Research Questions

RQ1: What is the impact of social media communication on consumer purchase intention in the Fast Food Industry?
RQ2: What is the influence of social media communication on brand equity in the Fast Food Industry?
RQ3: What is the impact of brand equity on consumer purchase intention in the Fast Food Industry?
RQ4: What is the mediating influence of brand equity on the relationship between social media communication and consumer purchase intentions in the Fast Food Industry?

1.5 Research Objectives

RO1: To identify the impact of social media communication on consumer purchase intention in the Fast Food Industry.
RO2: To understand the influence of social media communication on brand equity in the Fast Food Industry.
RO3: To analyze the impact of brand equity on consumer purchase intention in the Fast Food Industry.
RO4: To identify the mediating influence of brand equity on the relationship between social media communication and consumer purchase intentions in the Fast Food Industry.

2 Literature Review

This section will discuss the definitions, theories, and models that will be employed to examine key terms and elucidate the main aspects of the research.

2.1 Theoretical Frameworks and Models

2.1.1 Aaker's Brand Equity Model

Aaker's Brand Equity Model is a framework that measures and maintains brand equity. The model has five key dimensions: brand loyalty, brand awareness, perceived quality, brand associations, and proprietary brand assets. According to the brand equity theory defined by Aaker (1991), it is "a set of brand assets and liabilities linked to a brand, its name and symbol that add to or subtract from the value provided by a product or service to a firm and/or to that firm's customers". The five elements identified by Aaker (1991) show how each contributes toward strong brand equity (Lang & Guzmán 2022).

2.1.2 Keller's Brand Equity Model

According to Keller (1993), brand equity is the difference between how consumers respond to the marketing of the brand and how they would respond to a similar offering from another company. He's also discussed how it adds value to the brand apart from its name and the functional benefits it provides. He has identified four key components that create and contribute toward brand equity. Brand salience, brand performance, brand imagery, and customer feelings and judgments about the brand (Tasci 2021).

2.1.3 Theory of Reasoned Action

The theory of Reasoned Action is a remarkably robust theory that analyzes attitudes and behaviors. It also discusses how attitudes and behaviors are interconnected in human behavior. This theory is used widely to understand consumer behavior, including purchase intentions, product usage, and brand loyalty. This theory has also facilitated understanding of human behavior in the context of health promotion. Both attitude and behavior

are subjective standards that influence the intentions of purchase to understand and forecast people's behavior (Raut & Das, 2020). This theory is an excellent facilitator to understanding the attitude toward behavior, subjective norms and goals (Ajzen and Fishbein, 1975).

2.1.4 The Theory of Planned Behavior

The Theory of Planned Behavior is an extended theory developed after the Theory of Reasoned Action. Perceived behavior control predicts and explains a significant amount of variance beyond attitudes and subjective norms in terms of behavioral intention. This model outlays the relationship between attitudes, intentions, and behavior and how perceived behavior control may predict or explain those relationships (Ajzen and Fishbein, 1980; Sarah, 2010; La Barbera & Ajzen, 2021).

2.1.5 Brackett and Carr Framework

It is a useful model to understand consumer attitudes toward web advertisements that impact brand value (Bracket and Carr, 2001). He has suggested that advertising value is the main determinant that transmutes the attitude towards advertisement (Anubha & Shome, 2021). Further, the advertising value depends on the information provided, entertainment, credibility, and the lack of irritation. Regardless of the advertising value demographic variables also determine the attitude towards the web advertisement (Islam et al., 2021).

2.1.6 Social Exchange Theory

This theoretical framework is used to understand and explain social interaction and the relationship between gains and expenses (Ferm & Thaichon, 2021). It gives an idea of people's engagement in social media interaction and returns the gains received as rewards by the interaction (Homans, 1958). This relationship is not easily measurable. This theory is related to social media communication as customers are willing to sacrifice their time commitments to gain benefits like looking for particular products and services and in return, it impacts the purchase intention (Homans, 1961; Ahmad et al., 2023).

3 Conceptual Model

Table 1: Conceptual Framework

Independent Variables	Mediator	Dependent Variable
Social Media Communication	Brand Equity	Consumer purchase Intentions
1. Firm-created Communication 2. User-generated Communication	1. Perceived Quality 2. Brand Loyalty 3. Brand Awareness 4. Brand Associations	1. Customer feedback 2. Purchase Intention 3. Customer Preference

3.1.1 Social Media Communication

Social media communication is defined as “a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of user-generated content” (Kaplan & Haenlein 2010). Digital platforms include social networking sites, blogs, forums and messaging apps, etc. Brands use social media primarily for brand promotion, customer engagement, crisis management, and advocacy, etc (Cheung et al 2019). Further, it is a tool for modern communication and marketing-related strategies. The term “Social Media” has turned out to be a crucial aspect of human lives (Bruhn et al 2012). Where businesses utilize it as a potent tool to engage their potential customers. In the field of the fast food industry, communicating through social media has led to an impact on consumer buying decisions and expectations (Guha et al 2021). Social media links brand equity such that it positively creates a way to brand awareness and brand image. According to the study (Riaz et al 2019), the relationship between social media communication and brand equity was assessed and it had a positively co-related outcome (Crespo-Almendros et al 2022). (Styles and Ambler 1995) stated that the relationship between the firm and the public creates long-term buying behavior. A well-understood brand behavior raises the competitive barrier and leads to brand growth (Ahn et al 2018). Brackett and carr framework and the theory of planned behavior support this hypothesis.

H1: There is a significant positive relationship between social media communication on consumer purchase intentions.

Social media communication can be differentiated into two parts which are firm-created social media communication and user-generated social media (Saura et al 2021). The firm-created information is initiated, circulated, and consumed by users such that they educate products and services. The firm-created social media will also include word of mouth and electronic word of mouth (Seyyedamiri & Tajrobehkar 2021). User-generated social media communication includes social networking sites including Facebook, Twitter and YouTube, etc (Kar 2021). Internet and Web 2.0 are two proactive elements that enhance two-way communication and help online branding (Li and Bernoff 2011). Particularly this helps the brand equity in creating awareness of the brand. Aaker's brand equity model and social exchange theory support this hypothesis.

H2: There is a significant positive relationship between social media communication and brand equity.

3.1.2 Brand Equity

Brand equity is a term used to describe the value and strength that the brand has created in its marketplace. According to Keller, Parameswaran, & Jacob (2011), Brand Equity is the customer's beliefs related to the positive attributes of a brand and constructive consequences based on brand consumption. Rosenbaum-Elliott (2015) stated, "It is the simple difference between the value of a branded product, and the value of that product without that brand name attached to it ". Both user-generated social media and firm-generated social media communication are vital when building up a strong customer purchase intention. Brand equity plays a mediating part in the behavior of consumers. They have also found that clients with directly obtained messages from enterprises highly tend towards purchasing more than clients without direct enterprise messages. Another research has shown that pleasing content will provoke the attention of customers even if the reason behind using social media differs for individuals. This has been explored in a prior study where it was revealed that entertainment is a motivating aspect in encouraging the involvement of participants with social media. Along with analyzing social media, the impact of social media on clients' buying patterns online before launching marketing campaigns by companies that are active in this field.

H4: Brand equity mediates the relationship between social media communication and consumer purchase intentions.

3.1.3 Consumer Purchase Intention

This refers to the individual's inclination to lead towards purchase a product or service in the near future. This is concluded as the key factor in consumer behavior research which reflects the intentions of an individual's behavior, for example, will the individual purchase a particular brand's product. According to Moon (2008), the factors that influence purchase intentions include demographic attributes such as gender, age, professional and qualification background, product attributes, brand reputation, price, promotion, social and cultural norms, and personal values and beliefs. Schivinski and Dabrowski, (2016) have also stated that the common elements that contribute to customer purchase intention are consumer preferences, customer feedback and purchase intent (Nimantha 2019).

According to the study conducted by Aaker (1991), Brand equity is positively correlated to purchase intentions and he has also listed four key drivers that support the claim. Martini (2020) stated that strong brand equity attracts emotions and perceptions to precede buyer interests. Brand loyalty is another key function that retains existing customers. Makussi (2020) stated that a brand in relation to the consumer is either favorable or not connecting to the brand within its cognitive point. The brand association should be assessed in two different paths namely consumer behavior and mind brand association. Brand attributes benefit consumers to form a correlation to brand loyalty, and purchase patterns and in return brand identification can provide a purchase motive for the consumers. Aaker's brand equity model, Keller's brand equity model and theory of reasoned action model support this hypothesis.

H3: There is a significant positive relationship between brand equity and consumer purchase intentions.

4 Methodology

4.1 Research Philosophy

The research aims to understand social media communication's influence on customer purchase intention. A positive correlation between the two variables is expected and brand equity is the mediating effect point (Bell 2022). It is assumed that the statistical results obtained correlate with the previous results obtained from the empirical research. A positive attitude towards the research is made through dependable knowledge which is

derived from empirical and experimental means of literature reviewed by observed samples and which can be verified to the chosen industry and the geographical location.

4.2 Research Approach

The deductive approach is used to identify the relationship between the key concept theories and the empirical test. Based on these hypotheses were developed (Bougie and Sekaran 2019). The research aims and objectives were developed from the previous generalized theoretical research carried out, and a logical reasonably explainable conclusion is expected.

4.3 Research Method

A quantitative research method is used to identify the relationship between the stated variables. To understand the relationship statistical tools and models are used (Bougie and Sekaran 2019). Initially, a literature review was conducted to understand the performance of the variables on how customer purchase intention is affected. To get a detailed in-depth understanding the variables were chosen. A structured questionnaire was developed based on the theories and empirical research related to the chosen variables. The questionnaire was separated into 4 distinct parts. A 5-point Likert scale was used to rate different paths of the questions. 244 responses were collected. Initially, a pilot test was used before circulating the questionnaire.

4.4 Sampling Strategy

4.4.1 Sampling Technique

The sampling technique used was the 'simple random sampling method'. This method was chosen because it helps to prevent biased results and is easier to administer the results. This method also helps the researcher to draw conclusions for a generalized group (Bell 2022).

4.4.2 Sample Size

The samples are targeted within the Colombo district, the reason for the choice of location is to prevent samples from not knowing the brand. To understand the aim of the research and the results obtained a pilot test was conducted for a sample size of 25 randomly chosen. The questionnaire was thereafter distributed among 400 participants aiming to at least reach 260 responses. Usable responses obtained were only 244 sample results 51 results were rejected due to incomplete responses. A response rate of 61% was obtained which was fairly acceptable.

4.4.3 Research Instrument Development

Social media communication was operationalized using 2 indicators (firm-created communication and user-generated communication using measures adopted by Li and Bernoff, (2011) and Keller (1993) using a 5-point Likert Scale ranging from Strongly Agree (5) to Strongly Disagree (1). Brand equity was operationalized using measures adopted by Riaz et al (2019) and used measures such as brand loyalty, brand awareness, perceived quality, brand associations and brand salience. A 5-point Likert scale ranging from Strongly Agree (5) to Strongly Disagree (1). Consumer purchase intention was operationalized using measures adopted by Makussi (2020) using a 5-point Likert Scale ranging from Strongly Agree (5) to Strongly Disagree (1).

4.4.4 Data Collection and Data Analysis

Data for this study was collected through surveys and interviews. The respondents were a diverse group of consumers, representing various demographics and preferences. A pilot study was conducted with a small sample to test the survey's effectiveness, refine questions, and ensure data quality. Based on the result obtained from the pilot test some of the questions were altered, and a reliability test was carried out to identify the suitability of the responses. Thereafter, to analyze the data SPSS software was used. Normality, Correlation, Regression, Reliability, Validity and some other tests were carried out to examine the relationship of the variable with the real-time data set obtained.

5 Data Analysis and Representation

5.1 Normality of Data

5.1.1 Skewness and Kurtosis

To verify that the data is normally distributed Kurtosis and Skewness are used. The data collected on all three variables seem to be evenly distributed. Data on the brand equity has a slight variation and the Kurtosis is also a bit higher than the given range. This concludes that there are positive outliers present in the obtained real-time data. The histogram for all three variables for the collected data set resulted in the shape of the bell curve. Although outliers are present a bell curve of Mesokurtic (Normal Kurtosis) for the social media communication and customer purchase intention is obtained. Brand equity is Leptokurtic (Positive Kurtosis). Overall a linear data distribution is observed in the shown histogram and a proper bell curve is obtained. The data obtained gives a positive Skewness value ranging closer to 1 for all three variables and Kurtosis also ranges closer to positive 3.

5.1.2 Kolmogorov - Smirnov (K-S) and Shapiro-Wilk Test

The Kolmogorov-Smirnov test and Shapiro-Wilk test were carried out to understand the normality of data. The p-value was more than 0.05 and hence the data distribution can be considered normally distributed (Mishra et al., 2019).

5.1.3 Missing Values

During the data screening process, missing values were detected in the dataset and subsequently removed.

5.2 Justification of the Research Instrument

5.2.1 Reliability Analysis

The assessment of the research instrument's reliability involves analyzing the Cronbach Alpha value, which is 0.794. This value surpasses the threshold of 0.7, as noted by Mishra et al (2019) and Tavakol and Dennick (2011), indicating that the research instrument is indeed reliable.

Table 2: Reliability Statistics

Variables	Number of Items	Cronbach Alpha
Social Media Communication	8	0.725
Brand Equity	4	0.685
Customer Purchase Intentions	3	0.674
Whole Instrument	15	0.794

5.2.2 Validity

5.2.2.1 Sample Adequacy

The adequacy of the sample is assessed using the Kaiser-Meyer-Olkin measure and is 0.645 below the threshold of 0.5 (Budaev 2010; Goni et al 2020) and hence the sample size is adequate. Bartlett's test of sphericity is highly significant since the p-value is less than 0.05 (Goni et al 2020) and this further concludes that the sample size is adequate.

5.2.2.2 Convergent Validity

To establish convergent validity in the sample, it is expected that the inter-item correlations should surpass 0.3 (Campbell and Fiske 1959). However, it is worth noting that while the firm-generated social media communication and brand equity variables did not meet this threshold, all other variables did.

5.2.3 Construct Validity

To establish construct validity, a pilot study was conducted among 25 respondents, resulting in a Cronbach Alpha value of 0.843. This value exceeds the threshold value of 0.7, indicating that construct validity has been achieved.

5.2.4 Discriminant Validity

Discriminant validity evaluates the extent to which each construct differs from the others. The Wilks Lamda test indicated that there is a significant relationship between the independent variables and the dependent variable, as the p-value is below 0.05. The discriminant validity is greater when the eigenvalues are lower, as values less than 1 suggest that the constructed graph shows that discriminant validity exists.

5.3 Correlation Analysis

The Pearson correlation analysis was used to identify the significant relationship between the independent and dependent variables. The Pearson correlation between social media communication and brand equity is 0.468. This shows a positive correlation. The Pearson correlation between brand equity and customer purchase intention is 0.612. This shows a moderate but positive correlation. The Pearson correlation between customer purchase intention and social media communication is 0.457.

The relationship between independent and dependent variables is shown by the p-value. Since the p-value is less than 0.05, social media communication has a statistically significant relationship with brand equity. Since the p-value is less than 0.05, brand equity has a statistically significant relationship with customer purchase intention. Further, Since the p-value is less than 0.05, social media communication has a statistically significant relationship with customer purchase intention. This concludes that there is a positive relationship between the chosen variables.

5.4 Regression Analysis

In the regression analysis conducted the R2 value is 0.412 and the adjusted R2 value is 0.407. This determines that 41% variation of the dependent variable, customer purchase intention can be explained by the independent variables, brand equity and social media communication. However, the obtained value is considered to be a weak impact (Hair et al 2013). According to the ANOVA table and coefficient table, social media communication and brand equity are significant predictors of customer purchase intention as their p-values are below the threshold of 0.05. Brand Equity makes the largest contribution to the dependent variable, the beta value of brand equity is 0.510 and social media communication is 0.218 (Astivia, and Zumbo, 2019).

5.5 Mediator Analysis

The Sobels test for social media communication, as an independent variable to customer purchase intention as a dependent variable and brand equity as a mediator. An indirect effect analysis for the variable shown above using the Sobel test concludes that the p-value is less than or equal to 0.05 which means there is a satisfactory mediating relationship. To find out the point of the indirect variable effect at which the Sobel test is statistically significant unstandardized Coefficients and the beta value is multiplied by 0.36. There is a satisfactory relationship when considering the p-value.

Table 3: Mediator Coefficients 1 (Author Developed)

Coefficients[a]						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.689	.231		7.321	.000
	SocialMediaCommunication	.509	.062	.468	8.239	.000

a. Dependent Variable: Brand Equity

Table 4: Mediator Coefficients 2 (Author Developed)

Coefficients[a]						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-.319	.315		-1.015	.311
	SocialMediaCommunication	.337	.086	.218	3.899	.000
	Brand Equity	.724	.079	.510	9.124	.000

a. Dependent Variable: Consumer Purchase Intention

Table 5: Sobels Test (Author Developed)

Input:			Test statistic:	Std. Error:	*p*-value:
a	0.509	Sobel test:	6.11493682	0.06026489	0
b	0.724	Aroian test:	6.09484008	0.0604636	0
s_a	0.062	Goodman test:	6.13523369	0.06006552	0
s_b	0.079	Reset all	Calculate		

Sobel's test output revealed that the p-value is 0.000 and hence less than the threshold level of 0.05. Hence, brand equity mediates the relationship between social media communication and consumer purchase intention.

5.6 Hypothesis Testing

H1: There is a significant positive relationship between social media communication on consumer purchase intentions.

H1 was confirmed by the results obtained (A positive effect on social media communication and customer purchase intention= ($p < 0.001$; β 0.218; t-value 3.899). The correlation analysis revealed that social media communication has a positive relationship with consumer purchase intention. The regression analysis revealed that since the p-value is significant, social media communication has a significant influence on consumer purchase intention and hence hypothesis 1 is accepted (Majeed et al, 2021).

H2: There is a significant positive relationship between social media communication and brand equity.

H2 was confirmed by the results obtained (A positive effect on social media communication and brand equity= ($p < 0.001$; β 0.468; t-value 8.239).

The correlation analysis revealed that social media communication has a positive relationship with brand equity. The regression analysis revealed that since the p-value is significant, social media communication has a significant influence on brand equity and hence hypothesis 2 is accepted (Goh & Candy, 2022).

H3: There is a significant positive relationship between brand equity and consumer purchase intentions

H3 was confirmed by the results obtained (A positive effect on brand equity and customer purchase intention= ($p < 0.001$; β 0.510; t-value 9.124).

The correlation analysis revealed that brand equity has a positive relationship with consumer purchase intention. The regression analysis revealed that since the p-value is significant, brand equity has a significant influence on consumer purchase intention and hence hypothesis 3 is accepted (Hanaysha, 2022).

H4: Brand equity mediates the relationship between social media communication and consumer purchase intentions.

Sobel's Test output revealed that the p-value is 0.000 and hence less than the threshold level of 0.05. Hence, brand equity mediates the relationship between social media communication and consumer purchase intention and this hypothesis was accepted.

6 Conclusion and Recommendation

6.1 Main Conclusions

The purpose and the aim of this study is, to identify and explore the impact of social media communication on customer purchase intention, and assess the influence of brand equity on the relationship between social media and customer purchase intention in the fast food industry.

The following section will discuss the results obtained, and a comparison of empirical research data will be carried out. Statistical analysis was used to ensure the reliability, dimensionality, and validity of the measures. The data is distributed normally confirmed by the normality test as such multicollinearity does not exist confirming that the obtained data set is sufficient for further analysis. The KMO value is greater than 0.5 confirming that the collected sample is adequate to perform the rest of the statistical analysis. Convergent, construct, and discriminate validity were assessed and concluding that they do exist in validating the data. The research instrument was found to be internally consistent across all three variables and hence reliable.

The demographic findings concluded that the majority of the responses were collected from the male gender between the age of 18 – 25. The independent sample T-test determines the relationship between gender and customer purchase intention. Since the significance level for Levene's test is 0.015, which is below 0.05, we can conclude that there is no assumption of equal variances between the male and female groups. Additionally, the significance level of 0.012 is also below 0.05, indicating that there is significant evidence against equal variances. As a result, there is no discernible impact on the customer purchase intention and gender differences.

Further, social media communication had a positive correlation with customer purchase intention. Regression analysis revealed social media communication, brand equity predict customer purchase intention decisions by 41%. customer purchase intention is not only affected by social media communication and brand equity but also impacted by the price, culture context and other individual motivations. Pearson (R) value of 0.468 indicates a positive medium-strength relationship between brand equity versus social media communication (Majeed et al, 2021). Pearson (R) value of 0.612 indicates a positive large relationship between brand equity versus customer purchase intention. Pearson (R) value of 0.457 indicates a positive medium-strength relationship for customer purchase intention vs social media communication.

Brand equity has a significant and positive influence on consumer purchase intentions. Both regressed data and the correlation indicate a moderate strength relationship. This is similar to the results obtained by (He et al., 2021 & Goh and Candy, 2022).

Social media communication has a significant and positive impact on the brand equity of fast food brands in Sri Lanka. However, these findings differ from Majeed et al (2021) as he argued that social media communication can negatively impact brand equity in the fashion industry. This is because the content created in social media communication makes the customers positively reflect on the brand image and the brand awareness by how well they engage with the customer, Coca-Cola is an excellent example to verify the context.

Overall the independent variable has a significant impact on the dependent variable.

The analysis suggests the formulation of the following regression equation.

Customer purchase intention= 0.337 (social media communication) + 0.724 (brand equity) -0.319

In summary, findings demonstrate that there is a direct influence from social media communication to customer purchase intention and brand equity also indirectly affects the social media performance in contributing to the customer purchase intention.

Verified that the Hanaysha et al (2021) study presents a positive relationship between social media communication and customer purchase intention in the fast food industry. The results obtained for this aspect of the study are similar to those reported by Hanaysha (2022).

6.2 Recommendations

In support of the above findings and testing proved, it is observed that organizations in the fast-food industry are highly recommended to adopt and communicate via social media platforms which helps them to reach a wider range of audiences effectively which influences the customer purchase intention positively. Together with social media communication, the organization needs to focus on brand equity by developing brand awareness, brand loyalty, and brand association which leads to enhancing a wealthier brand personality and brand image. This development of brand equity would also impact the customer purchase intention positively.

6.2.1 Social Media Communication

Social media communication can help fast food companies increase customer purchase intention by using high-quality visuals, collaborating with influencers, and offering exclusive deals and promotions. Personalized communication and social proof through customer reviews can also build customer loyalty. These tactics can help fast food companies create a comprehensive social media strategy, increase brand awareness, drive sales, and build lasting customer relationships (Cheung et al 2019).

6.2.2 Brand Equity

Marketers must recognize the potential of brand awareness as a crucial determinant of consumer response. Firms should focus on implementing positive brand association practices, such as enhancing brand strength and brand description, to positively influence consumer response. To improve their performance, managers of fast food firms should prioritize providing high-quality and up-to-date services, ensuring round-the-clock support, fostering responsible employee behavior, and adopting user-friendly technology.

REFERENCES

[1] Aaker, D. 1996. *Measuring brand equity across products and markets. California Management Review* 38 (3): 102–120.

[2] Ahn, J. A., & Seo, S. 2018. *Consumer responses to interactive restaurant self-service technology (IRSST): The role of gadget-loving propensity.* International Journal of Hospitality Management, 74, 109–121. DOI: https://doi.org/10.1016/j.ijhm.2018.02.020.

[3] Ali, H. 2019. *Purchase Decision and Repurchase Models: Product Quality and Process Analysis (Case Study of House Ownership Credit Financing in Permata Sharia Bank Jakarta).* Scholars Bulletin, 5(9). DOI: https://doi.org/10.36348/sb.v05i09.

[4] Ansari, S., Ansari, G., Ghori, M. U., & Kazi, A. G. 2019. *Impact of Brand Awareness and Social Media Content Marketing on Consumer Purchase Decision*. Journal of Public Value and Administrative Insight, 2(2), 5–10. DOI: https://doi.org/10.31580/jpvai.v2i2.896.

[5] Bell, E. 2022. *Business research methods. Oxford university press.*

[6] Bougie, R. and Sekaran, U. 2019. *Research methods for business: A skill building approach.* John Wiley and Sons.

[7] Bruhn, M., Schoenmueller, V., & Schäfer, D. B. 2012. *Are social media replacing traditional media in terms of brand equity creation?.* Management Research Review, 35(9), 770-790.

[8] Campbell, D.T. and Fiske, D.W. 1959. *Convergent and discriminant validation by the multitraitmultimethod matrix.* Psychological bulletin, 56 (2), 81.

[9] Chadwick, S. 2021. *How social media is shaping the future of public relations. Neal Schaffer*. Online. ACM https://nealschaffer.com/social-media-shaping-future-public-relations/ (April 2023).

[10] Chen, Y.-R. R., Cheng, Y., Hung-Baesecke, C.-J. F., & Jin, Y. 2019. Engaging international publics via mobile-enhanced CSR (mCSR): A cross-national study on stakeholder reactions to corporate disaster relief efforts. American Behavioral Scientist, 63(12), 1603–1623. Online. ACM (April 2023). DOI: https://doi.org/10.1177/0002764219835258.

[11] Cheng, Y. 2019. The social-mediated crisis communication research: Revisiting dialogue between organizations and publics in crises of China. Public Relations Review, 46(1), 101769. Online. ACM (April 2023) DOI: https://doi.org/10.1016/j.pubrev.2019.04.003.

[12] Cheung, M. L., Pires, G. D., & Rosenberger III, P. J. 2019. *Developing a conceptual model for examining social media marketing effects on brand awareness and brand image*. International Journal of Economics and Business Research, 17(3), 243- 261.

[13] Crespo-Almendros, E., Prados-Peña, M. B., Porcu, L., & Alcántara-Pilar, J. M. 2022. *Influence of the perceived benefit of a social network-based sales promotion on the perceived quality of a heritage complex.* Academia Revista Latinoamericana de Administración, 35(2), 437-450. DOI: https://doi.org/10.1108/ARLA-03-2022-1382.

[14] Davis, F.D., Bagozzi, R.P. and Warshaw, P.R. 1989. *User acceptance of computer technology: A comparison of two theoretical models*. Management Science, 35(8), pp.982-1003.

[15] Department of Census and Statistics Sri Lanka. 2023. *Department of Census and Statistics Sri Lanka*. Online. ACM (April 2023) http://www.statistics.gov.lk/.

[16] Elliott, R., Percy, L., & Pervan, S. 2011. *Strategic brand management.* Oxford: Oxford University Press.

[17] Fishbein, M., & Ajzen, I. 1975. *Belief, attitude, intention, and behavior: An introduction to theory and research*. Addison-Wesley.

[18] Goh, A., & Candy, C. (2022). *Purchase intention in Burger King fast food: the role of brand equity*. Journal Manajemen Maranatha, 21(2), 237-249.

[19] Goni, M.D., Naing, N.N., Hasan, H., Wan-Arfah, N., Deris, Z.Z., Arifin, W.N., Hussin, T.M.A.R., Abdulrahman, A.S., Baaba, A.A. and Arshad, M.R. 2020. *Development and validation of knowledge, attitude and practice questionnaire for prevention of respiratory tract infections among Malaysian Hajj pilgrims*. BMC Public Health, 20 (1), 1-10.

[19] Guha, S., Mandal, A., & Kujur, F. 2021. *The social media marketing strategies and its implementation in promoting handicrafts products: a study with special reference to Eastern India*. Journal of Research in Marketing and Entrepreneurship, 23(3), 436-458. ISSN: 1471-5201. DOI: https://doi.org/10.1108/JRME-03-2021-0022

[20] Hanaysha, J.R. 2022. *Exploring the Impact of Social Media Marketing Features on Purchase Intention in the Fast-Food Industry.* South Asian Journal of Management, 21(2), 1-16. DOI: https://doi.org/10.1177/09726225221098792.

[21] Hasan, M., Saeed, N., Munawar, S., Hussain, S., Afzal, H., & Saleem, Q. 2021. *Impact of Brand Equity on Purchase Intention Through Brand Preference: Macro and Micro Perspective from Online Fast Food Industry of Pakistan*. Journal of Marketing and Consumer Research, 68, 18-31.

[22] Iyer, P., Davari, A., Srivastava, S., & Paswan, A. K. 2020. *Market orientation, brand management processes, and brand performance.* Journal of Product & Brand Management, 30(2), 197-214. DOI: https://doi.org/10.1108/jpbm-08-2019-2530

[23] Kaldeen, M. 2019. *Impact of firm-created and user-generated social media communication on brand associations*. Journal of Information Systems & Information Technology (JISIT), 4(2), 1-13. ISSN: 2478-0677

[24] Kaplan A.M., Haenlein. M. 2010. *Users of the world, unite! The challenges and opportunities of social media.* Bus Horiz 53(1):59–68.

[25] Kar, A. K. 2021. *What Affects Usage Satisfaction in Mobile Payments? Modelling User Generated Content to Develop the "Digital Service Usage Satisfaction Model".* Information Systems Frontiers, 23(6), 1341-1361. DOI: https://doi.org/10.1007/s10796-020-10047-x.

[26] Keller, K. L., Parameswaran, M. G., & Jacob, I. 2011. *Strategic brand management: Building, measuring, and managing brand equity*. Pearson Education India.

[27] Keller, K.L. 1993. *Conceptualizing, measuring and managing customer-based brand equity*. Journal of Marketing 57 (1): 1–22.

[28] Majeed, M., Owusu-Ansah, M. and Ashmond, A.A. 2021. *The influence of social media on purchase intention: The mediating role of brand equity. Cogent Business & Management*, 8(1), 1944008. DOI: https://doi.org/10.1080/23311975.2021.1944008.

[29] Makussi, M. 2020. *Influence of Online Advertising on Enhancing Brand Equity: A Case Study of Vodacom Tanzania.* Doctoral dissertation, Mzumbe University.

[30] Martini, E. 2020. *The brand equity model: a comparison between private label and national brand*.

[31] Mishra, P., Pandey, C.M., Singh, U., Gupta, A., Sahu, C. and Keshri, A. 2019. *Descriptive statistics and normality tests for statistical data.* Annals of cardiac anaesthesia, 22 (1), 67.

[32] Riaz, H., Ahmed, H., & Akhtar, S. 2019. *The role of social media marketing on building brand equity: An insight of fast food industry of Pakistan.* New Media and Mass Communication, 83, 22. DOI: https://doi.org/10.7176/NMMC.

[33] Saura, J. R., Ribeiro-Soriano, D., & Palacios-Marqués, D. 2021. *From user-generated data to data-driven innovation: A research agenda to understand user privacy in digital markets*. Journal of Business Research, 60, 212-219.

[34] Schivinski, B., & Dabrowski, D. 2014. *The effect of social media communication on consumer perceptions of brands. Journal of Marketing Communications*, 1–26. DOI: https://doi.org/10.1080/13527266.2013.871323.

[35] Seyyedamiri, N., & Tajrobehkar, L. 2021. *Social content marketing, social media and product development process effectiveness in high-tech companies. International Journal of Emerging Markets*, 16(1), 38-58. DOI: https://doi.org/10.1108/IJoEM-02-2018-0059.

[36] Styles, C., and T. Ambler. 1995. *Brand Management*. Pitman, London: Financial Times Handbook of Management.

[37] Tavakol, M. and Dennick, R. 2011. *Making sense of Cronbach's alpha. International journal of medical education*, 2, 53.

[38] Zhao, Y., Wang, L., Tang, H. and Zhang, Y. 2020. *Electronic word-of-mouth and consumer purchase intentions in social e-commerce*. Electronic Commerce Research and Applications, 41.

Influence of social media marketing using Instagram and consumer trust on consumer purchase intentions in the skincare industry in Colombo District in Sri Lanka

Hansini Navodya Somaratna, IIT Business School, Colombo, Sri Lanka, hansini.2019472@iit.ac.lk

Gayashini Ratnayake, IIT Business School, Colombo, Sri Lanka. gayashini.r@iit.ac.lk

Abstract

In Colombo District, Sri Lanka, the skincare industry is influenced by social media marketing and consumer trust, shaping purchase intentions significantly. Social media serves as a powerful tool for brands to engage potential customers through influencers and beauty bloggers (Rathnayake & Lakshika, 2023). Consumer trust in brands and influencers plays a pivotal role in purchasing decisions(Kim et al., 2019). Understanding this relationship is crucial for businesses to develop effective strategies and for consumers to make informed choices in the skincare market(Narang & Sharma, 2022). Aim of this study is to fill this knowledge gap by investigating how consumer trust and social media marketing influence consumer purchase intentions in the Sri Lankan cosmetics industry. A deductive research approach based on hypothesis testing was adopted. A questionnaire was distributed to 450 individuals, and 252 viable replies were received with a response rate of 63%., simple random sampling was selected as the sampling approach. The study revealed a significant positive relationship between social media marketing through Instagram and consumer trust, suggesting that social media marketing through Instagram has a significant influence on consumer trust in the Sri Lanka skincare industry. However, social media marketing on Instagram was not seen as sufficient to encourage customers to make a purchase decisions, particularly in the Skin care industry in Sri Lanka.Normality tests, independent sample tests, correlation and regression analyses, and hypothesis testing were all performed quantitatively using SPSS version 23. The Sobels test was used for mediation.

Keywords: Influence, social media marketing, consumer trust, consumer purchase intentions, skincare industry, Colombo District, Sri Lanka.

1 Introduction

The personal care segment of Sri Lanka's skincare market has witnessed substantial expansion in recent times, with an estimated revenue of US$92.69 million expected by 2027 (Statista, 2023). To capitalize on this growth, marketers in the industry are leveraging social media platforms to connect and engage with customers in novel ways. Social media marketing has been found to significantly influence consumer behavior, particularly purchase intentions (Huang et al., 2020; Kim et al., 2019), through interactive advertising, user-generated content, and influencer marketing (Munsch, 2021).

1.1 Research Gap

Social media marketing and consumer trust have been understudied in the Sri Lankan skincare industry, with previous studies focusing on other countries (Alalwan et al., 2018; Kim et al., 2019). Empirical research is needed to examine the connection between social media marketing, customer trust, and purchase intentions in the Sri Lankan skincare sector. Trust as a mediator and direct effects of social media marketing on consumer buy intentions are still missing.

1.2 Research Questions

The following research questions were identified.

RQ1: What is the influence of social media marketing on consumer purchase intentions?

RQ2: What is the influence of Consumer trust on consumer purchase intentions?

RQ3: What is the influence of social media marketing on consumer trust?

RQ4: How does Consumer trust mediate the relationship between social media marketing and consumer purchase intentions?

1.3 Research Objectives

Based on the above research questions, the following research objectives were identified
RO1: To identify the influence of social media marketing on consumer purchase intentions?

RO2: To assess the influence of Consumer trust on consumer purchase intentions?

RO3: To understand the influence of social media marketing on consumer trust?

RO4: To identify how does Consumer trust mediate the relationship between social media marketing and consumer purchase intentions?

1.4 Scope of Study

This study examines the Sri Lankan skincare industry, focusing on consumer trust and social media marketing. It aims to close the research gap by examining how these factors affect consumer behaviour and purchase intentions in the market. Marketers can create efficient strategies and policies to cater to customer preferences and requirements. This study aims to investigate the influence of Instagram on consumer behavior within the skincare industry in Colombo District, Sri Lanka.

2 Literature Review

2.1 Theoretical Background and Models

2.1.1 Social Exchange Theory

Alamanos(2023), Ganesan, Shi & Prentice (2019), Wang, Chen, & Liang, and Davlembayeva & Alamanos(2023) emphasize the importance of social exchange theory in determining consumer purchase intentions. Consumers are more inclined to make purchases when they perceive a higher benefit-cost ratio in the exchange. Consumer trust is also a crucial element in determining purchase intentions, as customers are more inclined to buy a brand's goods or services when they believe in it. Open and sincere interactions with customers in social media marketing can improve customer trust, ultimately improving their propensity to make purchases.

2.1.2 Social Learning Theory

Rumjaun & Narod's (2020) theory suggests that people learn knowledge through watching others' behaviour and results. Social media marketing campaigns featuring positive reviews and endorsements can influence the perception of skincare products, increasing public buy inclinations in the sector.

2.1.3 Social Influence Theory

Liu et al. (2019) highlight the Social Influence Theory, which suggests that consumers' attitudes, beliefs, and subsequent actions are influenced by their social network. Consumers are more likely to buy a product in social media marketing if they believe it is well-liked by others in their network. This social proof can be a powerful tool for influencing consumer purchase intentions, as it fosters a sense of social proof and influences the behavior of consumers.

2.1.4 Technology Acceptance Model

Denaputri & Usman, Cheung et al (2020)., and Jin & Phua (2018) have all contributed to the development of the Technology Acceptance Model (TAM), a modification of the Theory of Reasoned Action (TRA). TAM focuses on perceived usefulness and ease of use, which are crucial components of online buying. Perceived usefulness refers to a person's belief that using a system will enhance their job performance, while perceived ease of use refers to the degree to which a person believes using a system is free of effort. Consumers are more likely to interact with a company on social media if they find the platform user-friendly and the brand's content educational and helpful. Phua'(2018)s research shows that customers who perceive social media platforms as practical and easy to use have a more favourable opinion of the company and are more inclined to buy goods.

2.1.5 The Theory of Planned Behaviour

Ajzen's (2020) Theory of Planned Behaviour highlights the importance of consumer attitudes, subjective standards, and perceived behavioural control in shaping behaviour intention. Social media marketing can

influence consumer purchase intentions in the skincare sector by influencing consumer perceptions of the skincare line and enhancing their ability to control behaviour by providing information about accessibility and availability of products. Giampietri et al(2018). also highlight the potential impact of social media marketing on consumer behaviour.

2.1.6 Expectancy Value Model

Ajzen and Fishbein's (1980) theory of reasoned action (TRA) provides a comprehensive explanation of social conduct, based on the Expectancy-Value Model. The model suggests that consumer behaviour is influenced by their value for a product and their expectations for its performance. Barjolle highlights the importance of high expectations for product value and performance in shaping buy intentions. This comprehensive explanation highlights the importance of understanding consumer behaviour and its impact on decision-making.

2.1.7 The Expectation Disconfirmation Theory
Choi et al. (2019) suggest that a company's performance should match or surpass customer expectations, as customers are more likely to trust a firm when they fulfil promises and use their goods or services again.
2.1.8 The Theory of Reasoned Action

The TRA states that behavioural intention, influenced by a person's attitude and subjective norms, determines consumer behaviour. Ajzen & Fishbein (1980) argue that attitudes are favourable or negative assessments of actions, while subjective norms are how individuals feel under social pressure to engage in or refrain from behaviour. A study on online customers used an adapted TRA approach to analyse the impact of perceived risk and individual trust on online purchase attitudes. The findings suggest that consumers' perceptions of risk and trust significantly influence their opinions, making them more likely to shop online. Consumer trust can affect subjective norms, while social media marketing can affect attitudes towards skincare products. The findings imply that consumers' perceptions of risk and trust while making purchases online have a significant impact on people's opinions (Jadil et al., 2022). According to the study consumers' attitudes made them more likely to shop online(Jadil et al., 2022).

3. Conceptual Framework
Social Media marketing is the independent variable. Social media marketing is the use of social media platforms to engage with target audience in order to promote the brand, improve sales, and drive traffic to the website. (Zahay et al., 2023). This entails posting high-quality material on social media profiles, listening to and engaging with followers, assessing results, and conducting social media marketing. (Zahay et al., 2023). Because of social media platforms such as Facebook, Instagram, YouTube, and WhatsApp, businesses all over the world now have access to a broader range of commercial options. These platforms enable marketers to interact with potential customers, build relationships with them, and increase the legitimacy and dependability of their companies(McClure and Seock, 2020). According to a study, social media content marketing has a moderately positive and significant relationship with consumer purchase decisions, which suggests that it is important in this technologically advanced era. Powerful content will influence consumers to take the desired action, which is to make a purchase decision. (Ansari et al., 2019).

The Theory of Planned Behavior expands this by taking into account perceived behavioral control, demonstrating how social media marketing can help consumers feel more in control of their purchasing decisions. The Elaboration Likelihood Model lends support to the notion that engaging social media content might influence purchase intentions via active information processing. Furthermore, social identity theory, informational social influence, the AIDA model, and the customer engagement model all underline the importance of social media marketing in generating trust, creating a sense of belonging, and encouraging consumers to make purchases.(Confente & Vigolo, 2018) (Giampietri et al., 2018) (Then & Johan, 2021).
H1: There is a strong positive relationship of social media marketing on consumer purchase intention

Consumer trust is the mediator variable. Trust is the belief in another party, such as a trusted third party, that increases customers' intent to shop online and encourages greater shopping behaviours. Zhao et al. (2019) emphasize that buyers base their decisions on merchants' claims and deeds. Trust is crucial for businesses to build relationships with consumers, as demonstrated by Abror et al. (2019) and Bhalla (2020). Benevolence and trust based on credibility are essential for building trust in transactions, as they ensure consistent ties between seller and buyer. Sanny et al. (2020) found that social media marketing has a positive impact on brand trust. This finding is consistent with Bernardo et al. (2020), who also found that this impact is positive. An increase in social media marketing variables will lead to an increase in brand trust variables. (Haudi et al., 2022). Online groups that generate a sense of belonging and trust are highlighted in Social Identity Theory.The Source Credibility Theory

emphasizes the need of recognized influencers in fostering trust. By exceeding expectations, the Expectancy-Disconfirmation Model strengthens confidence. Positive contacts generate trust in Social Exchange Theory. Overall, these ideas support the importance of social media marketing in increasing consumer trust in brands and products.(Wang, Chen, & Liang, 2018; Davlembayeva& Alamanos2023)..
H2: There is a strong positive relationship between social media marketing and consumer trust.

Consumer purchase intention is the dependent variable. Purchase intention in social commerce (SC) is influenced by social factors, technical factors, and motivational factors. Zhang et al. (2022) highlight the importance of trust in building consumer trust, which influences views and intent to buy. These factors, including utility and hedonic motivation, act as mediators for consumers' decisions, influencing their views and intent to buy in SC. However, trust does not directly affect purchase intentions, as noted by Rahman et al. (2020) and Sharma et al. (2019).).
When consumers trust a brand, the Theory of Reasoned Action and Theory of Planned Behavior emphasize the importance of attitudes, subjective standards, and perceived behavioral control on purchase intentions. The Expectancy-Value Theory explains how consumer trust effects expected outcomes and perceived value of a product, resulting in increased purchase intentions. Furthermore, the Social Exchange Theory, Brand Trust Model, Attribution Theory, Cognitive Dissonance Theory, and Affect-Reflexivity Theory all help to understand how trust in a brand influences customers' emotions, beliefs, and perceptions. These ideas emphasize the importance of trust in establishing and keeping strong ties with consumers, which leads to greater buy intentions. (Confente & Vigolo, 2018) (Giampietri et al., 2018) (Then & Johan, 2021)
H3: There is a strong positive relationship between consumer trust and consumer purchase intention

In the context of social media marketing, a study looked at the mediating role that trust plays in the relationship between customers' social motivation (SM) and their intention to make an online purchase. The findings showed that social media significantly increases consumer trust and online purchase intent. They also demonstrated how customers' levels of trust had a big impact on their online PIs. In addition, the results showed that customers' online purchase intentions are partially mediated by SM impacts by their trust in fashion merchants. (Karunasingha & Abeysekera, 2022).
H4: Consumer trust mediates the relationship between social media marketing and consumer purchase intention

4 Research Methodology

The research philosophy, research approach ad research tools employed in this study are discussed below.

4.1 Research Philosophy
The study's research philosophy is positivism, prioritizing empirical evidence and objective observation in explaining social processes (Bryman, 2016).

4.2 Research Approach
A deductive approach was used in the study since the research objective focused on testing relationships between the main variables.

4.3 Research Method
This study employs a survey research design, gathering data from a sample using a questionnaire or survey instrument. Quantitative procedures are less biased and involve statistical tools and models, as demonstrated by Bougie and Sekaran (2019). The survey was created using validated scales and indicators, and the research instrument underwent pilot testing before use.

4.4 Sampling Strategy
The sampling technique, sample size, research instrument development, operationalization of variables and data collection is discussed below.

4.4.1 Sampling Technique

Because it minimizes bias and is straightforward to apply, simple random sampling was selected as the sampling approach (Bell, 2022). Every member of the population has an equal probability of getting chosen using simple random sampling. This facilitates the development of a representative sample that precisely captures the traits of the entire population (Berndt, 2020). It lessens bias and makes sure that the results can be confidently generalized to a bigger group. The possibility of bias entering the sample is reduced with the use of random selection (Berndt, 2020).

4.4.2 Sample Size & Respondents

The goal of this study is to collect data from people over the age of 18 who live in Sri Lanka's Colombo District. Colombo's current metro area population in 2023 is 633,000, a 1.12% growth from 2022. 2023 (MacroTrends, Year?) A pilot study with 30 participants was carried out. The Questionnaire was then distributed to 403 people, with 252 viable replies obtained. Because they were incomplete, 148 responses were removed. The response rate was 63%, which was considered adequate.

The main city in Sri Lanka, Colombo, has a diversified population that comes from different socioeconomic, cultural, and racial backgrounds. A number of well-known supermarkets, including Keells and Glomark, as well as a number of retail centers, including Colombo City Center, Liberty Plaza, and One Galle Face, provided the data for this study. Due to their large presence and popularity within the Colombo metropolitan area, these establishments were chosen as data gathering stations.

4.4.3 Research Instrument Development

A cross sectional survey was used to collect data. Social media marketing was operationalized using three items used by Yousafzai et al (2019 using a 5 point Likert scale ranging from Strongly Disagree (1) to Strongly Agree (5).Consumer trust was operationalized using two items applied by Al-Swidi & Al-Matari (2019) using a 5 point Likert scale ranging from Strongly Disagree (1) to Strongly Agree (5).Consumer purchase intention was operationalized using seven items based on the measures used by Wimalasiri et al (2020) using a 5 point Likert scale ranging from Strongly Disagree (1) to Strongly Agree (5). The research instrument consisted of screening questions, a demographic profile, and primary factors.

5 Data Analysis

5.1 Normality of the Data

Sekaran and Bono (2019) discuss the primary characteristics of normality tests, focusing on skewness and kurtosis. Skewness, a proportion of an informational collection's evenness, is assumed normal when values range from -1 to +1. Mishra et al. (2019) found that data distribution is normal with kurtosis values between -3 and +3, indicating a normal distribution.

5.1.1 Skewness and Kurtosis & Kolmogorov – Smirnov (K-S) and Shapiro Wilk-test

Histograms showed a bell curve and linear distribution, with information dispersion being direct and near the line, with no anomalies. Sekaran's Kolmogorov-Smirnov test evaluates data set normality, ensuring P-value greater than 0.05 for normal distribution in a normally distributed data set (Sekaran, 2003). Investigated missing values in dataset through screening, removing them. The data distribution is normal because all kurtosis values fall within this range. The histograms showed a bell curve, and the data distribution was linear (Appendix). The p plot showed that the information dispersion was direct and nearer to the line and anomalies were absent. The histogram graphs showed a ringer bend, it was ordinary to affirm that the information conveyance.

Statistics

		Purchase Intention	Trust	Social Media Marketing
N	Valid	252	252	252
	Missing	0	0	0
Skewness		-.440	-.441	-.496
Std. Error of Skewness		.153	.153	.153
Kurtosis		-.906	.223	.113
Std. Error of Kurtosis		.306	.306	.306

5.2 Justification of the Research Instrument

The research instrument was assessed for reliability and validity. Reliability is assessed to ascertain whether the same results will be achieved if the same instrument is used multiple times. Validity is assessed to ascertain whether a method accurately measures what it is supposed to measure.

5.2.1 Reliability Analysis

Sekaran and Saunders (2009) used Cronbach's coefficient alpha to assess the consistency of measurement in data collection and analysis. The reliability test ensures reliable results, with values below 0.60 considered poor, within 0.70-0.80 acceptable, and above 0.80 exceptional. Table 11 displays Cronbach's alpha values for three measurement scales, indicating questionnaire reliability.

Variable	Cronbach Alpha Value
Consumer Purchase Intention	0.754
Trust	0.887
Social Media Marketing	0.943

Table 1 Cronbach's Alpha Values of Variables (Author developed)

5.2.2 Validity

5.2.2.1 Sample Adequacy Test

Variable	**KMO Value**
Purchase Intention	**0.558**
Trust	**0.720**
Social Media Marketing	**0.764**

Table 2 KMO Value for Variables

Table 2 shows KMO and Bartlett's Test values above 0.5, indicating validated instruments with a 95% confidence level, indicating sufficient sample size.

5.2.2.2 Convergent Validity

Campbell and Fiske's(1959) inter-item correlations ensure convergent validity, as all main variables meet criterion, forming three main factors.

5.2.2.3 Construct Validity

Construct Validity of pilot study exceeded Cronbach Alpha, with 30 respondents analyzed. Discriminant validity tests reveal huge relationship between free factors and dependent variables, indicating discriminant validity, but eigen value is less than 1.

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.082	2	.541	3.257	.040[b]
	Residual	41.367	249	.166		
	Total	42.449	251			

a. Dependent Variable: Purchase Intention

b. Predictors: (Constant), Social Media Marketing, Trust

Table 3 ANOVA (Author developed)

5.3 Regression Analysis

R square values are used to examine the significance of the key variables using regression analysis. According to the table above's Adjusted R square value, social media marketing and trust predict consumer purchase intention 1.8 percent of the time. The ANOVA table reveals that the p value is less than 0.05 and hence trust and social media marketing have a significant relationship with consumer intent to buy. The trust makes the largest contribution to the dependent variable. As a result, the consumer's intention to buy was determined by the beta value of 0.169.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2.903	.130		22.389	.000
	Trust	.093	.049	.169	1.907	.058
	SocialMediaMarketing	-.006	.039	-.014	-.157	.875

a. Dependent Variable: PurchaseIntention

Table 4 Coefficients (Author developed)

5.4 Correlation Analysis

The correlation analysis reveals that there is a positive relationship between social media marketing and trust since the p value is statistically significant. Trust and consumer purchase intention have a positive relationship since the p value is significant since it is less than 0.05. However, since the p value exceeds 0.05, there does not seem to be a relationship between social media marketing and consumer purchase intention.

Correlations

		PurchaseIntention	Trust	SocialMediaMarketing
PurchaseIntention	Pearson Correlation	1	.159*	.106
	Sig. (2-tailed)		.011	.093
	N	252	252	252
Trust	Pearson Correlation	.159*	1	.709**
	Sig. (2-tailed)	.011		.000
	N	252	252	252
SocialMediaMarketing	Pearson Correlation	.106	.709**	1
	Sig. (2-tailed)	.093	.000	
	N	252	252	252

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Table 5 Correlation Analysis Output

5.5 Mediator Analysis

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.656	.132		12.566	.000
	SocialMediaMarketing	.565	.035	.709	15.912	.000

a. Dependent Variable: Trust

Table 6 Mediator Coefficients 1 (Author developed)

The regression results between trust and social media marketing, are shown in the table above to determine the mediator relationship between social media marketing (Independent variable) and trust (mediator).

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2.903	.130		22.389	.000
	SocialMediaMarketing	-.006	.039	-.014	-.157	.875
	Trust	.093	.049	.169	1.907	.058

a. Dependent Variable: PurchaseIntention

Table 7 Mediator Coefficients 2 (Author developed)

The regression analysis between the independent variable, mediator and dependent variable is shown above.

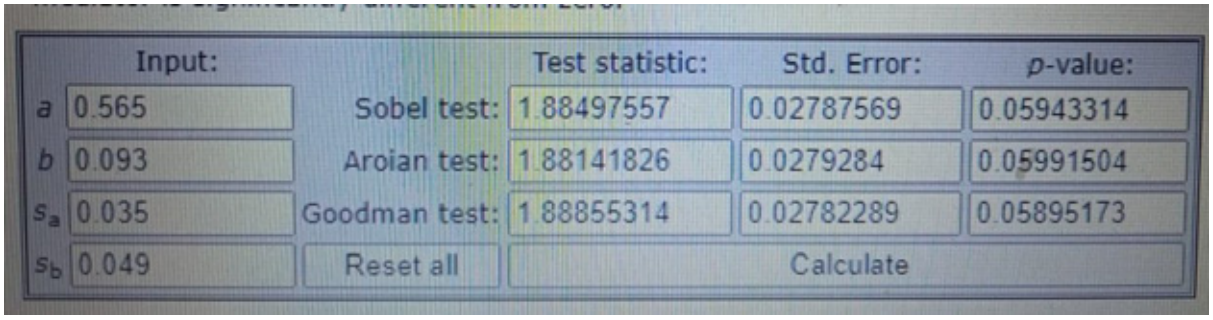


Figure 1 Sobel test (Author developed)

The Sobel's test is used to determine whether a mediator variable significantly carries the influence of an independent variable to a dependent variable. The Sobel's test output shows that that p value is slightly greater than 0.05 and hence trust does not mediate the relationship between social media marketing and customer purchase intention.

5.6 Hypothesis Test

H1: There is a strong positive relationship of social media marketing on consumer purchase intention
Based on the Coefficient Table, the p value is greater than 0.05 and hence this hypothesis is rejected. The correlation analysis also reveals that social media marketing does not have a direct relationship with the consumer purchase intentions. Hence this hypothesis is rejected.

H2: There is a strong positive relationship between social media marketing and consumer trust.
The correlation analysis reveals that there is a positive relationship between social media marketing and trust since the p value is statistically significant. The regression analysis shows that the p value is significant and hence this hypothesis is accepted.

H3: There is a strong positive relationship between consumer trust and consumer purchase intention
Based on the correlation analysis, trust and consumer purchase intention have a positive relationship since the p value is significant since it is less than 0.05. However,.Based on the Coefficient Table, the p value is greater than 0.05 and hence this hypothesis is rejected.

H4: Consumer trust mediates the relationship between social media marketing and consumer purchase intention
Based on Sobels test, trust does not mediates the relationship between social media marketing and consumer purchase intention. This hypothesis is rejected

5.7 Recommendations

To enhance customer trust, companies should focus on transparency by providing accurate information about their products and pricing. Encouraging customer reviews and addressing negative feedback promptly helps build credibility. Providing accurate and detailed information about products, ingredients, and manufacturing processes, as well as being transparent about pricing, promotions, and discounts, fosters a sense of openness and honesty. Prioritizing high-quality customer service and using influencer marketing also contribute to establishing trust with customers. Influencer marketing can also be a valuable tool in building trust and credibility with customers. Collaborating with influencers and bloggers who have a strong following and a positive reputation in the industry can enhance a company's brand image and credibility.

For social media marketing, companies should be responsive to customer inquiries and create engaging content to attract and retain followers. Encouraging user-generated content, such as customer reviews and photos, can be a powerful way to build trust and credibility. Encouraging user-generated content and monitoring social media metrics helps showcase the brand and optimize marketing efforts. Implementing these strategies can lead to stronger customer relationships, increased brand loyalty, and improved overall marketing performance.

5.7.1 Recommendations for Future Research

Sri Lanka's skincare market faces challenges in consumer trust and buy intentions due to various social media platforms like Facebook, Instagram, and Twitter. Influencers' impact on consumer trust and purchasing intentions is also crucial. Cultural influences, pricing, and customer attitudes towards social media marketing are also

important factors to consider. Future studies should use longitudinal data to monitor shifts in customer attitudes and understanding of these factors in the Sri Lankan skincare market.

6 Conclusion

This study's findings indicate that social media marketing has a considerable beneficial impact on consumer trust, lending credence to the idea that businesses may use social media platforms to create relationships with potential customers and boost the legitimacy of their brands. Furthermore, the study found a positive association between consumer trust and consumer buy intention, supporting the idea that trust is important in influencing customers' purchasing decisions.

However, the data did not support the hypothesis asserting a direct positive association between social media marketing and consumer purchase intention, implying that social media marketing may not have a direct impact on buy intentions. In future study, it is critical to recognize this restriction and investigate additional factors that may influence consumer purchase intentions. Furthermore, the results did not support the hypothesized hypothesis about consumer trust mediating the relationship between social media marketing and consumer purchase intention. While the study found evidence that trust influences buy intentions, it did not discover evidence that trust functions as a mediator in the relationship between social media marketing and purchase intention.

Furthermore, the study discovered that social media marketing does not influence consumer purchase intention in the Skin care industry in Sri Lanka. These findings support prior empirical research by Pathan et al. (2021) where they state that information from the social media marketing is not only enough but also information from the friends, peers and family also very crucial when purchasing the product.

Previous research has revealed a correlation between social media activity and customer views of brand social media sites and future purchase intentions (McClure & Seock, 2020). Another study found that social media marketing has a positive impact on consumers' purchase intentions, as one of the factors that influenced people to acquire products from such platforms (Kumar & Sharma, 2022). Among other factors, customers followed their favourite brands online and received up-to-date information from them (Kumar & Sharma, 2022). However, my findings do not support these findings since social media marketing through Instagram did not have a significant and positive impact on customer purchase intentions in the Skin care industry in Sri Lanka. In contrast to prior research in other contexts, the findings of the study done in the cosmetics business in Sri Lanka reveal a lack of significant and beneficial influence of social media marketing through Instagram on client purchase intentions. Unique sample characteristics, industry peculiarities, evolving market trends, cultural and socioeconomic aspects, content and strategy variances, competitive environment, external influences, and platform relevance are all possible explanations for this disparity ((Kumar & Sharma, 2022) (McClure & Seock, 2020)).

These empirical findings consistently highlight the positive impact of social media marketing on consumer trust. Social media platforms provide businesses with opportunities to engage with customers in real-time, offer personalized interactions, and demonstrate transparency and credibility (Lou and Yuan, 2019). Such activities contribute to building trust in the brand and its products, ultimately influencing consumer purchase intentions positively. A study showed that social media significantly increases consumer trust and online purchase intent. They also demonstrated how customers' levels of trust had a big impact on their online PIs.

The findings also revealed that consumer trust didn’t mediate the relationship between social media marketing through Instagram and consumer purchase intentions. These findings support previous empirical findings by Jain N.K el al. (2018) states that social media can attract a consumer and make him a customer but it won’t get the locality of that customer. According to the research data, social media marketing increases consumer trust, which leads to greater buy intentions. Businesses may establish trust in their brand and products by providing real-time engagement, individualized interactions, and transparency via social media platforms. In contrast to prior research, a study found that consumer trust did not buffer the association between social media marketing on Instagram and purchase intentions. Sample characteristics, assessment methodologies, trust-building efforts, brand reputation, cultural and societal impacts, and platform-specific effects are all potential factors impacting this discrepancy. More study is needed to fully comprehend the complicated dynamics of social media marketing, customer trust, and buy intentions across diverse settings and platforms ((Lou and Yuan, 2019) (Jain N.K el al. (2018)).

Females have long been the core market for the cosmetics industry. Recently, males have become more aware of the importance of beauty, creating a niche market for male consumers, and are increasingly concerned with their own appearance, as previously study has shown (Narang & Sharma, 2022).

REFERENCES

[1] Abror, A. et al. (2019). Service quality, religiosity, customer satisfaction, customer engagement and Islamic Bank's Customer Loyalty. Journal of Islamic Marketing, 11 (6), 1691–1705. Available from https://www.emerald.com/insight/content/doi/10.1108/JIMA-03-2019-0044/full/html [Accessed 2 April 2023].
[2] Ajzen, I. (2020) "The theory of planned behavior: Frequently asked questions," Human Behavior and Emerging Technologies, 2(4), pp. 314–324. Available at: https://doi.org/10.1002/hbe2.195.
[3] Alalwan, A. A., Rana, N. P., Dwivedi, Y. K., & Algharabat, R. (2018). Social media in marketing: A review and analysis of the existing literature. Telematics and Informatics, 35(1), 1-13. https://fardapaper.ir/mohavaha/uploads/2017/10/Social-Media-in-Marketing-A-Review-and-Analysis-of-the-Existing-Literature.pdf

[4] Author links open overlay panelYassine Jadil a *et al.* (2022) *Understanding the drivers of online trust and intention to buy on a website: An emerging market perspective*, *International Journal of Information Management Data Insights*. Elsevier. Available at: https://www.sciencedirect.com/science/article/pii/S266709682200009X (Accessed: May 1, 2023).

[5] Bock, G. W., Zmud, R. W., Kim, Y. G., & Lee, J. N. (2020). Behavioral intention formation in knowledge sharing: Examining the roles of extrinsic motivators, social-psychological forces, and organizational climate. MIS Quarterly, 44(1), 111-143. https://www.academia.edu/3239341/Behavioral_intention_formation_in_knowledge_sharing_examining_the_roles_of_extrinsic_motivators_social_psychological_forces_and_organizational_climate

[6] Chen C. D., Zhao Q., Wang J. L. (2020). How livestreaming increases product sales: Role of trust transfer and elaboration likelihood model. Behaviour and Information Technology. Advance online publication. https://doi.org/10.1080/0144929X.2020.1827457

[7] Chen, S.-C., & Lin, C.-P. (2019). Understanding the effect of social media marketing activities: The mediation of social identification, perceived value, and satisfaction. Technological Forecasting and Social Change, 140, 22–32. doi:10.1016/j.techfore.2018.11.025 https://www.sciencedirect.com/science/article/pii/S0040162518311570

[8] Chetioui, Y., Benlafqih, H. and Lebdaoui, H. (2020) "How fashion influencers contribute to consumers' purchase intention," *Journal of Fashion Marketing and Management: An International Journal*, 24(3), pp. 361–380. Available at: https://doi.org/10.1108/jfmm-08-2019-0157.

[9] Cheung, M.L., Pires, G. and Rosenberger, P.J. (2020) "The influence of perceived social media marketing elements on consumer–brand engagement and brand knowledge," *Asia Pacific Journal of Marketing and Logistics*, 32(3), pp. 695–720. Available at: https://doi.org/10.1108/apjml-04-2019-0262.

[10] Choi, I.Y., Moon, H.S. and Kim, J.K. (2019) "Assessing personalized recommendation services using expectancy disconfirmation theory," Asia Pacific Journal of Information Systems, 29(2), pp. 203–216. Available at: https://doi.org/10.14329/apjis.2019.29.2.203.

[11] Chong, S.C. et al. (2020) "Consumers' purchase intention toward ergonomic footwear in Malaysia," International Journal of Financial Research, 11(2), p. 88. Available at: https://doi.org/10.5430/ijfr.v11n2p88.

[12] CHUSNAINI, A.Z.M.I.L.C.H.U.S.N.A.I.N.I. and RASYID, R.E.I.Z.A.N.O.A.M.R.I. (2022) "Social Media Marketing: Social media content, Brand Image, Brand Awareness and purchase intention," Accounting and Management Journal, 6(2), pp. 46–53. Available at: https://doi.org/10.33086/amj.v6i2.3028.

[13] Confente, I. and Vigolo, V. (2018) "Online travel behaviour across cohorts: The impact of social influences and attitude on hotel booking intention," International Journal of Tourism Research, 20(5), pp. 660–670. Available at: https://doi.org/10.1002/jtr.2214.

[14] Davlembayeva, D. & Alamanos, E. (2023) Social Exchange Theory: A review. In S. Papagiannidis (Ed), TheoryHub Book. Available at http://open.ncl.ac.uk / ISBN: 9781739604400 https://open.ncl.ac.uk/theory-library/social-exchange-theory.pdf

[15] Denaputri, A. and Usman, O. (2020) *Customers' Trust and Technology Acceptance Model on online purchase intention*, *SSRN*. Available at: https://paperaas.ssrn.com/sol3/papers.cfm?abstract_id=3647193 (Accessed: May 1, 2023).

[16] Ebrahim, R.S. (2019) "The role of trust in understanding the impact of social media marketing on Brand Equity and brand loyalty," *Journal of Relationship Marketing*, 19(4), pp. 287–308. Available at: https://doi.org/10.1080/15332667.2019.1705742.

[17] Erkan, I. and Evans, C. (2018), "Social media or shopping websites? The influence of eWOM on consumers' online purchase intentions", Journal of Marketing Communications, Routledge, Vol. 24 No. 6, pp. 617-632.

[18] Giampietri, E. et al. (2018) "A theory of planned behaviour perspective for investigating the role of trust in consumer purchasing decision related to short food supply chains," Food Quality and Preference, 64, pp. 160–166. Available at: https://doi.org/10.1016/j.foodqual.2017.09.012.

[19] Hidayat, A. et al. (2021) 'Consumer Trust as the antecedent of online consumer purchase decision', Information, 12(4), p. 145. doi:10.3390/info12040145.

[20] Huang, Z., Benyoucef, M., & Dai, S. (2020). Social media marketing and brand loyalty: A meta-analytic review. Information Technology & People, 33(1), 205-228. https://e-tarjome.com/storage/panel/fileuploads/2022-03-16/1647420628_E16221.pdf

[21] Ibrahim, B. and Aljarah, A. (2023) "The role of social media marketing activities in driving self–brand connection and user engagement behavior on Instagram: A moderation–mediation approach," European Journal of Innovation Management [Preprint]. Available at: https://doi.org/10.1108/ejim-08-2022-0452.

[22] Jadil, Y., Rana, N.P. and Dwivedi, Y.K. (2022) "Understanding the drivers of online trust and intention to buy on a website: An emerging market perspective," *International Journal of Information Management Data Insights*, 2(1), p. 100065. Available at: https://doi.org/10.1016/j.jjimei.2022.100065.

[23] Jin, S.-A.A. and Phua, J. (2014) "Following celebrities' tweets about brands: The impact of Twitter-based electronic word-of-mouth on consumers' source credibility perception, buying intention, and social identification with celebrities," *Journal of Advertising*, 43(2), pp. 181–195. Available at: https://doi.org/10.1080/00913367.2013.827606.

[24] Julfikar Ali, M. et al. (2022) 'Determinants of consumers' purchase intention to buy smartphones online', Innovative Marketing, 18(2), pp. 109–119. doi:10.21511/im.18(2).2022.10.

[25] Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. Business Horizons, 53(1), 59–68. doi:10.1016/j.bushor.2009.09.003 https://www.sciencedirect.com/science/article/pii/S0007681309001232

[26] Kumar, A. and Sharma, N.K. (2022) "Impact of social media on consumer purchase intention," Research Anthology on Social Media Advertising and Building Consumer Relationships, pp. 1501–1518. Available at: https://doi.org/10.4018/978-1-6684-6287-4.ch081.

[27] Lim, W.M. (2022) "Toward a theory of social influence in the new normal," Activities, Adaptation & Aging, 46(1), pp. 1–8. Available at: https://doi.org/10.1080/01924788.2022.2031165.

[28] Liu, X., Min, Q. and Han, S. (2019) "Understanding users' continuous content contribution behaviours on microblogs: An integrated perspective of uses and gratification theory and social influence theory," Behaviour & Information Technology, 39(5), pp. 525–543. Available at: https://doi.org/10.1080/0144929x.2019.1603326.

[30] Manzoor, U. et al. (2020). Impact of social media marketing on consumer's purchase intentions: The mediating role of customer trust. International Journal of Entrepreneurial Research, 3 (2), 41–48. Available from https://readersinsight.net/IJER/article/view/1386 [Accessed 2 April 2023]

[31] McClure, C. and Seock, Y.K. (2020), "The role of involvement: investigating the effect of brand's social media pages on consumer purchase intention", Journal of Retailing and Consumer Services, Vol. 53 (March), p. 101975.

[32] Munsch, A. (2021) "Millennial and Generation Z Digital Marketing Communication and advertising effectiveness: A qualitative exploration," Journal of Global Scholars of Marketing Science, 31(1), pp. 10–29. Available at: https://doi.org/10.1080/21639159.2020.1808812.

[33] Mohan Raj, M.P., A., V. and S., A. (2018) “Measurement of customers’ brand choice and Brand Loyalty Expectancy Value & Colombo – Morrison Model Approach,” JOURNAL OF INTERNATIONAL BUSINESS RESEARCH AND MARKETING, 4(1), pp. 24–29. Available at: https://doi.org/10.18775/jibrm.1849-8558.2015.41.3003.

[34] Narang, R. and Sharma, R. (2022) “Impact of demographic factors on purchase intention of Organic Skin Care Products: A study in select cities of India,” Orissa Journal of Commerce, pp. 58–73. Available at: https://doi.org/10.54063/ojc.2021.v42i04.05.

[35] Oghazi, P. et al. (2018) ‘Online purchase return policy leniency and purchase decision: Mediating role of consumer trust’, Journal of Retailing and Consumer Services, 41, pp. 190–200. doi:10.1016/j.jretconser.2017.12.007.

[36] Peña-García, N. et al. (2020) ‘Purchase intention and purchase behavior online: A cross-cultural approach’, Heliyon, 6(6). doi:10.1016/j.heliyon.2020.e04284.

[37] Pradhana, D., Duraipandiana, I. and Sethi, D. (2016), “Celebrity endorsement: how celebrity–brand–user personality congruence affects brand attitude and purchase intention”, Journal of Marketing Communications, Vol. 22 No. 5, pp. 456-473.

[38] Rumjaun, A. and Narod, F. (2020) “Social Learning theory—Albert Bandura,” Springer Texts in Education, pp. 85–99. Available at: https://doi.org/10.1007/978-3-030-43620-9_7.

[39] Sreen, N., Chatterjee, S. and Sadarangani, P. (2021) ‘Eco-friendly products purchase intention: A comparison of theory of planned behaviour and social cognitive theory’, International Journal of Economics and Business Research, 22(2/3), p. 149. doi:10.1504/ijebr.2021.116324.

[40]Statista, 2023. Statista. [Online] Available at: https://www.statista.com/outlook/dmo/ecommerce/beauty-health-personal-household-care/personal-care/sri-lanka#revenue [Accessed 2023 January 2023].

[41]Usman, O. and Zillah, I.A. (2019) ‘The effect of recommendation, customer satisfaction, online shopping experience, trust and word-of-mouth affecting Consumer Online Shopping Decision’, SSRN Electronic Journal [Preprint]. doi:10.2139/ssrn.3511634.

[42] Valentini, C. et al. (2018) ‘Digital Visual engagement: Influencing purchase intentions on Instagram’, Journal of Communication Management, 22(4), pp. 362–381. doi:10.1108/jcom-01-2018-0005.

[43] Wang, Y. (2016) “Information adoption model, a review of the literature,” *Journal of Economics, Business and Management*, 4(11), pp. 618–622. Available at: https://doi.org/10.18178/joebm.2016.4.11.462.

[44] Yeon J., Park I., Lee D. (2019). What creates trust and who gets loyalty in social commerce? Journal of Retailing and Consumer Services, 50, 138–144. https://doi.org/10.1016/j.jretconser.2019.05.009\

[45] Zahay, D.L. et al. (2023) Social Media Marketing: A strategic approach, Google books. Singapore: Cengage Learning Asia Pte Ltd. Available at: https://books.google.lk/books?id=iv9UEAAAQBAJ&printsec=frontcover&dq=Social+media+marketing+2023&hl=en&newbks=1&newbks_redir=0&source=gb_mobile_search&ovdme=1&sa=X&redir_esc=y#v=onepage&q=Social%20media%20marketing%202023&f=false (Accessed: February 6, 2023).

[46] Zhou T. (2019). The effect of social interaction on users’ social commerce intention. International Journal of Mobile Communications, 17(4), 391. https://doi.org/10.1504/ijmc.2019.100501

Mama Aid - An application to identify postpartum complications and assist in re-establishing mother's health

Erandi Kavindya Liyanagamage
Business School, Informatics Institute of Technology, Colombo 06, Sri Lanka, erandi.20191082@iit.ac.lk

Abarnah Kirupananda
Business School, Informatics Institute of Technology, Colombo 06, Sri Lanka, abarnah.k@iit.ac.lk

ABSTRACT

This project focuses on addressing the challenges faced by mothers during the postpartum stage, specifically postpartum depression. Depression is a common mental health issue worldwide, and postpartum depression affects around 10-15% of new mothers. It can have both physical and mental effects on mothers due to factors like fatigue, malnutrition, weight gain, and delayed milk supply.

To support the recovery and health of postpartum mothers, an IT solution called "MamaAid" was developed. This mobile application aims to provide a comprehensive solution by addressing postpartum complications, evaluating postpartum depression levels, and offering personalized meal plans and exercise recommendations. It also includes features like postpartum depression screening, exercise suggestions, and medical appointment scheduling.

The project involved a literature review, interviews with medical professionals, and questionnaires with postpartum mothers to develop an effective solution. The results highlighted the importance of a holistic approach to self-management and received positive feedback from experts and stakeholders. MamaAid aims to guide postpartum mothers towards achieving their "healthy mom" goals.

KEYWORDS

Postpartum depression, Personalized meal plan recommendation, Postpartum complications, Depression Screening

1 Introduction

Postpartum depression (PPD) is a significant mental health concern that affects a substantial number of women globally. It is characterized by the presence of depressive symptoms experienced by mothers following childbirth, typically occurring within the first four weeks after delivery [1]. The prevalence of postpartum depression is estimated to be around 18%. Despite Sri Lanka's reputation for exemplary maternal and child health care, a recent study conducted in 2022 revealed that a concerning percentage of mothers, 9.4% and 5.6%, were at risk of developing postpartum depression, as indicated by scores of 9 or more and 10 or more on the Edinburgh Postpartum Depression scale (EPDS) [2]. This finding highlights the need for increased attention and support for postpartum mental health in the country.

It is important to recognize that postpartum depression is not solely a mental illness; its effects can also manifest physically and emotionally. Women experiencing PPD may exhibit a wide range of symptoms, including feelings of sadness, anxiety, irritability, and hopelessness. These emotional struggles can significantly impact a mother's overall well-being and her ability to care for herself and her newborn [2].

In addition to the emotional toll, postpartum mothers may face various other challenges that exacerbate the risk of developing PPD. One such challenge is the lack of proper nutrition during the postpartum period. Mothers often neglect their own dietary needs while prioritizing the care of their newborns, leading to nutrient deficiencies and imbalances that can contribute to physical and mental health issues. Adequate nutrition is essential for promoting maternal well-being and preventing postpartum depression [3].

Postpartum fatigue is another common issue faced by new mothers. The demanding nature of caring for a newborn, coupled with sleep deprivation, can lead to extreme exhaustion and overwhelming fatigue. This physical fatigue can further contribute to the emotional strain experienced by postpartum mothers and increase their vulnerability to PPD [4].

Weight loss and weight gain are additional concerns that postpartum mothers may grapple with. The pressure to regain pre-pregnancy weight and appearance, combined with societal expectations, can create a significant source of stress for many women. Rapid weight loss or struggles with losing weight can negatively impact body image and self-esteem, further exacerbating the risk of developing postpartum depression [5].

Unfortunately, due to the demanding nature of their routines and a lack of understanding about postpartum depression, many mothers tend to neglect these indicators and symptoms. The pressure to meet societal expectations and fulfill the role of a "perfect mother" often leads women to downplay or dismiss their feelings, which can make it even more challenging for them to achieve their "Healthy Mom" goals.

In conclusion, postpartum depression is a global concern that requires immediate attention, including in Sri Lanka. By increasing awareness, providing education, implementing screening procedures, and creating support networks, we can empower mothers to prioritize their mental health and well-being. This comprehensive approach should address not only the emotional aspects of postpartum depression but also the physical challenges faced by new mothers, such as lack of nutrition, postpartum fatigue, and weight-related issues. By acknowledging and addressing these factors, we can support postpartum mothers in their journey towards better mental health and help them achieve their "Healthy Mom" goals.

2 Background

2.1 An Overview of Postpartum Depression and Its Impact on Mother’s Health

Depression is a highly prevalent mental health disorder worldwide, ranking as the second most common condition in terms of prevalence in 2019. It affects a significant number of individuals and has a particularly noticeable impact on women compared to men [3]. Among the various forms of depression, postpartum depression (PPD) stands out as a specific type experienced by mothers after giving birth. It is characterized by symptoms including anxiety, stress, loss of energy, disorientation, feelings of guilt or unworthiness, sleep disturbances, appetite disruption, and even suicidal thoughts. These symptoms not only have a profound effect on mothers but also potentially influence the mother-infant bond and the child's growth and development [4].

The postpartum period poses the highest risk of severe depression that may require hospitalization for women. PPD can significantly impair a woman's ability to care for herself and fulfill her role as a mother [5]. The exhaustion, inadequacy, worthlessness, and helplessness experienced by mothers with PPD often lead to social isolation. The physiological, emotional, and sociological changes during pregnancy and the postpartum period contribute to emotional difficulties for women [6]. PPD not only impacts the mother's health but also has consequences for the baby's health and development. If left untreated, PPD can result in long-term effects and may become chronic or recurring. Furthermore, women who have experienced PPD have a higher risk of relapse and recurrence, even during subsequent pregnancies. It is therefore crucial to provide effective treatment and support to address PPD and mitigate its potential consequences [7].

Given the significance of PPD, it is essential to prioritize the well-being of mothers during the postpartum period. Healthcare professionals, including obstetricians, pediatricians, and mental health specialists, play a crucial role in identifying and addressing PPD. Early detection and intervention are vital in preventing the escalation of symptoms and reducing the long-term impact on both mothers and infants. Treatment options for PPD typically involve a combination of therapies, including counseling, support groups, and, in some cases, medication. Additionally, providing a supportive and understanding environment for mothers can greatly aid their recovery. Societal awareness and education about PPD are also crucial. By promoting open discussions and reducing the stigma surrounding mental health issues, we can encourage women to seek help without hesitation. Support networks, both online and offline, can offer a sense of community and provide valuable resources and guidance for women experiencing PPD. Moreover, involving partners and family

members in the care and support of mothers can contribute to a more comprehensive and effective recovery process.

In conclusion, PPD is a specific form of depression that affects mothers after childbirth, with significant implications for both maternal and infant well-being. Recognizing the signs and symptoms, providing timely intervention and support, and raising awareness about PPD are essential steps in mitigating its impact. By prioritizing the mental health of mothers and implementing comprehensive care strategies, we can promote healthier outcomes for both mothers and their children.

2.2 PPD General Issues

Healthcare providers, including nurses, doctors, and social workers, often lack awareness of the various issues that women may face during pregnancy. Symptoms of postpartum depression typically arise within three months after childbirth, but they can also emerge unexpectedly. Unfortunately, postpartum depression hampers women's ability to care for themselves and their newborns. It is crucial to improve the understanding and recognition of postpartum depression among healthcare professionals to provide appropriate support and intervention for affected women [8].

Table 2. DSM-IV Criteria for a Major Depressive Episode

At least 5 of the following symptoms are present during the same period. Either depressed mood or loss of interest or pleasure must be present. Symptoms are present most of the day, nearly daily for at least 2 weeks:

1. Depressed mood (sometimes irritability in children and adolescents) most of the day, nearly every day
2. Markedly diminished interest or pleasure in all or almost all activities most of the day, nearly every day
3. Significant weight loss/gain or decrease/increase in appetite
4. Insomnia/hypersomnia
5. Psychomotor agitation/retardation
6. Fatigue or loss of energy
7. Feelings of worthlessness or excessive inappropriate guilt
8. Impaired concentration or indecisiveness
9. Recurrent thoughts of death, suicidal ideation with or without a specific plan, or suicide attempt

[a]Adapted from reference 4.

Figure 1: DSM-IV Criteria for Major Depressive Episode [8]

2.2.1 Appetite Change

Stress, anxiety, and depression have been shown to have a significant impact on the hormones and neurotransmitters that regulate appetite and food reward. These psychological factors can lead to changes in eating behaviors, resulting in either increased or decreased food intake, as supported by various studies [9]. The postpartum period, in particular, poses a vulnerable time for women when it comes to eating and body image issues. The distress and dissatisfaction with their body shape and size during this phase can contribute to the development of eating issues and disordered eating patterns. Negative feelings and concerns about their postpartum bodies can have detrimental effects on both the physical and mental health of women, creating challenges with eating and feeding leading up to postpartum depression [10].

It is crucial to recognize and address these psychological factors and body image concerns in order to support the well-being of postpartum women. By providing appropriate support and resources, healthcare professionals can help women navigate this sensitive period and promote healthy eating behaviors and positive body image, which are vital for their overall health and happiness.

2.2.2 Weight Loss and Gain

Obesity is identified as a contributing factor to depression, particularly in pregnant women who experience weight gain. Concerns about changes in appearance and perceived body changes during pregnancy can trigger anxiety and sadness [11]. Pregnancy typically involves a weight gain of around 12 kg, and the body mass index (BMI) is used to assess weight gain in pregnant women. Approximately 46% of women experience weight changes outside the recommended range during pregnancy [12]. Despite the widespread occurrence

of postpartum depression (PPD) and its long-term impact on families, the condition is often overlooked by both patients and healthcare professionals, resulting in underdiagnosis and undertreatment. Studies have also shown a correlation between weight gain and depression [13].

The impact of postpartum depression extends beyond the individual and can affect the entire family dynamic. It can strain relationships, impair maternal-infant bonding, and hinder the overall well-being of the mother and child. Early detection and intervention are crucial for mitigating the negative consequences associated with postpartum depression. Healthcare providers play a vital role in assessing the mental health of pregnant women and monitoring their emotional well-being throughout pregnancy and the postpartum period. By prioritizing mental health screening and education, healthcare professionals can better support women in managing their weight, minimizing the risk of depression, and promoting overall maternal and infant health. Additionally, increased awareness and open dialogue about postpartum depression can help reduce the stigma associated with seeking help and encourage affected individuals to seek appropriate support and treatment.

Table 2

Odds ratios for type of weight gain and BMI by depression.

Variable		Non depressed n (%)	Depressed n (%)	OR	CI	p-value
Type of weight gain (kg)	Inadequate	42 (24%)	13 (7.4%)	0.55	0.24-1.60	0.326
	Adequate	59 (33.9%)	33 (19%)	1.05	0.21-1.57	0.281
	Excessive	17 (9.7%)	10 (5.7%)	Ref.		
Type of BMI (kg/m^2)	Normal	59 (33.9%)	29 (16.6%)	0.82	0.56-2.03	0.826
	Over weight	59 (33.9%)	27 (15.5%)	Ref.		

Figure 2: Odd ratio for types of Weight Gain in BMI with Depression [13]

3 Self-Assessment and Self-Management Interventions

3.1 Importance of Self-assessment and Self-management Interventions in Postpartum Complications

Self-management is a proactive approach that involves employing self-control techniques to mitigate the effects of health disorders or their risk factors. It encompasses a range of tasks, including setting goals, self-monitoring symptoms, making informed decisions, problem-solving, and effectively managing physical, emotional, and cognitive reactions [14]. Self-management strategies are often used in conjunction with traditional treatments and play a vital role in empowering individuals to control their symptoms and prevent relapses. Technology, such as manuals, videos, and websites, frequently serves as a platform for delivering self-management assistance. While it is important to note that self-management techniques cannot replace professional diagnosis and treatment, they provide mothers with valuable tools to assess their well-being and actively manage their recovery process [15]. Mama Aid, an innovative mHealth application, is designed with the aim of supporting mothers in self-assessment and guiding them towards achieving their "healthy mom" goals. With its user-friendly interface and comprehensive features, Mama Aid strives to empower mothers in their journey towards optimal health and well-being.

3.2 Methods of Self-assessment and Self-management Interventions in Postpartum Complications

Postpartum depression (PPD) is categorized as a major depressive disorder and a significant mood disorder. As such, self-management techniques commonly used for depression can also be effective in addressing PPD and its associated complications. Multiple treatment methods are available for managing depression and the postpartum challenges it brings. The objective of these techniques is to offer support, alleviate symptoms, and prevent relapses for mothers navigating through PPD [16]. By utilizing self-management strategies, women experiencing PPD can access valuable tools to cope with their condition, enhance their well-being, and promote a healthier postpartum journey.

Table 1: Self-Management Techniques in Postpartum

Component	Self-Management Technique
Information	• Educating family and friends about PPD and its complications
Symptom Management	• Self-monitoring of symptoms • Relaxation • Being aware of postpartum complications and its symptoms • Using strategies to prevent relapse of PPD
Lifestyle	• Exercise (Home workout, Standard routine) • Engaging in leisure activities • Engaging in activities 2-4 times a week • Healthy and balanced nutritional intake
Communication and Support	• Communication strategies • Family support
Avoidance	• Excess use of alcohol and smoke

4 Role of Machine Learning Algorithms

Machine learning is a fundamental component of artificial intelligence (AI) that allows computers to learn from data and experiences without explicit programming. It has diverse applications in tasks such as data classification, prediction, and pattern recognition. There are several types of machine learning algorithms, including supervised learning, unsupervised learning, and rule-based approaches. Supervised learning uses labeled data to make predictions and classify information, while unsupervised learning discovers patterns in unlabeled data [17]. Rule-based methods involve creating explicit sets of rules for data classification or prediction. Overall, machine learning plays a crucial role in enabling computers to learn and make informed decisions based on data analysis [18].

4.1 Rule Based Approach

Rule-based methods are a common class of machine learning and data mining techniques. They involve generating predetermined outputs based on a set of rules defined by humans. These rules are typically expressed as if-else statements and serve as a form of artificial intelligence that directs a computer to reach conclusions or make recommendations based on specific logical rules. A rule-based system consists of facts about a situation and a set of rules to handle those facts [19]. Rules can be developed through a general-to-specific or specific-to-general approach, either by adding conditions to a rule until improvements are observed or by removing conditions to cover specific circumstances. Evaluation metrics such as accuracy, knowledge gain, or likelihood ratio can guide the development and refinement of these rules [20].

In the development of the meal plan recommendation system, a meticulous rule-based approach was employed. The primary objective was to cater the meal plans to individual users based on their specific weight, ensuring that the maximum required protein intake for their body was accurately determined.

To achieve this, a set of well-defined rules was carefully crafted, following the guidance and acceptance of medical health professionals. These rules formed the backbone of the system, allowing it to make precise calculations for each user's protein requirements. The involvement of medical experts played a crucial role in ensuring the system's accuracy and reliability. By incorporating this rule-based methodology, the meal plan recommendation system could offer tailored and effective dietary suggestions, tailored to the unique needs of each user.

5 Materials and Methods

The study in the discussion was mixed-method research to design and develop a mobile application that would assist postpartum mother's in identifying postpartum and also supporting them in regaining their health. The research was performed following 4 major phases, which are discussed below.

5.1 Phase 01: Requirement Assembling

The process of gathering requirements involved two primary methods: questionnaires and interviews. To ensure a meaningful and conclusive response, a sample size calculation was performed for the questionnaire. The intention was to determine the number of participants that would provide a realistic representation of the population. To select respondents, a snowball sampling technique was employed, ensuring diversity across various factors. In order to gain deeper insights into postpartum complications, interviews were conducted with a range of medical professionals, including gynecologists, nutritionists, and medical consultants. These interviews were invaluable and helped to gain a better understanding of the challenges faced by the target audience from the perspective of healthcare providers. By engaging with these experts, It was easy to gather valuable information regarding the pain points and concerns related to postpartum complications.

Overall, the combination of questionnaires and interviews allowed for a comprehensive approach in gathering requirements. The questionnaires provided a broader view by including a diverse sample size, while the interviews provided specific insights from medical professionals, enriching the understanding of the subject matter.

5.2 Phase 02: Analysis

A meticulous analysis was conducted of the data gathered from various primary research methods, leaving no stone unturned. This thorough examination served as the foundation for deriving an extensive and detailed list of requirements. Recognizing the importance of secondary research, they also delved into competitor analysis to identify potential gaps that required attention based on the primary research findings. This comprehensive approach ensured that the proposed solution would be well-informed and capable of meeting all the necessary criteria.

To facilitate the process of defining the project's scope and capturing it accurately, a design thinking techniques, such as mind mapping, brainstorming, and creating a fishbone diagram were employed. These methods stimulated creativity and fostered a deep understanding of the problem space, leading to more innovative and effective solutions.

During this phase, the solution's scope was meticulously delineated, and a finalized list of both functional and non-functional requirements was established. To enhance the clarity of the project and provide a visual representation of its various aspects, UML diagrams were utilized, including use case diagrams, activity diagrams, and class diagrams. These diagrams offered a concise and coherent overview of the proposed solution's structure, functionality, and interactions. The proposed solution itself encompassed a wide range of features, aiming to address the diverse needs identified through the research process and offer a holistic approach to postpartum complications. These features included:

- Meal plan recommendation: Utilizing the user's required protein intake for the day, the solution generated personalized meal plans to support their postpartum nutrition goals under 4 different diet types.
- Exercise recommendation: Tailored to the user's specific postpartum stage, the solution provided targeted exercise recommendations to aid in their recovery and promote overall well-being.
- Screening for postpartum depression: An integral component of the solution involved implementing a screening process to identify potential cases of postpartum depression, enabling early intervention and support.
- Practitioner booking and verification: The solution facilitated the booking of healthcare practitioners and ensured the verification of recommended plans, reinforcing the reliability and credibility of the offered services.

- Information hub: An inclusive and accessible information hub was incorporated into the solution, providing users with relevant resources, educational materials, and support networks.

By integrating these comprehensive features into the proposed solution, it was aimed to fulfill the identified needs and provide an all-encompassing approach to addressing postpartum complications. The resulting solution would cater to the specific requirements and challenges faced by individuals during this crucial phase, offering support, guidance, and resources to promote a healthy and successful postpartum experience.

5.3 Phase 03: Design

To provide a visual representation of the proposed solution's design and architecture, a high-level architecture diagram below. This diagram presents a clear overview of how the components of the solution interact and are organized, aiding in understanding the structure and functioning of the application at a glance.

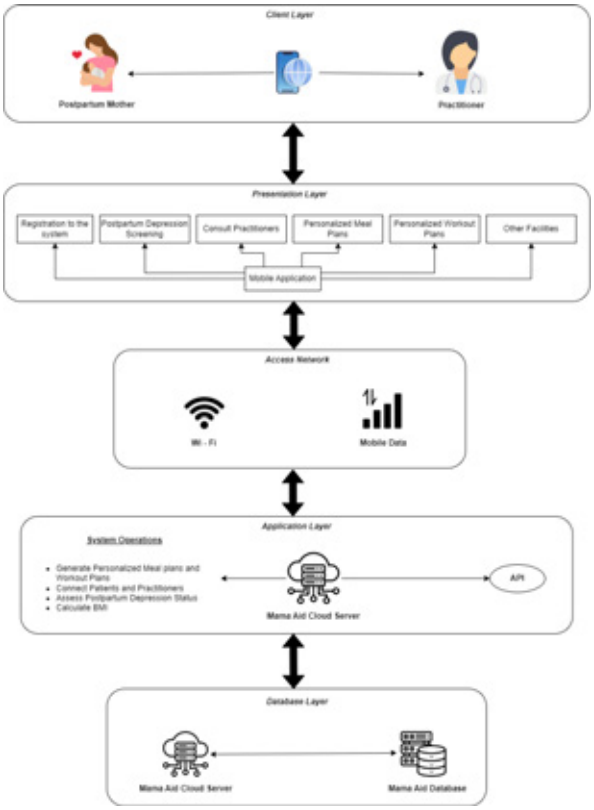


Figure 3: High Level Architecture Diagram (Author)

5.4 Phase 04: Development and Testing

The solution was built with Android Studio Code and the programming languages Java, XML, and Python. The data was saved in a Firebase database, and PhpMyAdmin was used to manage it. To propose meal plans depending on the user's necessary protein consumption, a rule-based method was built.

As for testing, Functional and user testing were carried out to confirm the correctness and expected performance of the generated solution. The core capabilities of the system were functionally tested using the black-box approach. The testing method yielded positive results, with all test cases producing the expected response.

5.5 Phase 05: Evaluation

A prototype was assessed by Target audience, domain experts, technical experts, and industry end-users to validate the effectiveness of the implemented solution. The assessment procedure was divided into two parts:

concept-based evaluation and solution-based evaluation. A specified set of questions was given to testers in order to acquire relevant feedback and ideas. Below given is a summary of evaluation results.

5.5.1 Concept Based Evaluation Summary

The feedback from the evaluators was very encouraging. Both favorable application outcomes, negative inputs and comments about how the application could be improved on specific application elements will be considered for future enhancements to the application

5.5.2 Solution Based Evaluation Summary

Technical experts praised the application, mentioning that it captured the key functionalities, but they also thought the scope could have been more clearly defined when it came to user type practitioner. Regarding the functionalities, features, and technology aspects, particularly the rule-based approach, the following feedbacks were received

Overall, given the time limits and abilities available, the technologies and methodologies chosen for the prototype were deemed to be ideal selections for constructing the prototype. All of this feedback was considered for future improvements. The expert and non-expert evaluators enjoyed the prototype's features, and technical evaluations indicated a good use of technologies and the rule-based approach, as well as the overall UI/UX of the prototype.

6 Conclusion

In conclusion, the project successfully developed an application, "MamaAid," to address postpartum complications and assist postpartum mothers in re-establishing their health. The application's uniqueness lies in its postpartum depression screening, personalized meal plan recommendations, practitioner booking and verifications, and workout recommendations based on postpartum stages. The strengths of the application include its capability to obtain assistance from Sri Lankan experts, remote and easy access, and providing a comprehensive platform for postpartum mothers' needs. However, there are some weaknesses such as the lack of Sri Lankan food recipes in meal plans, not considering allergies, and the absence of a notification feature.

Throughout the project, the gained new knowledge and skills in the postpartum domain, research, communication, modeling, technology implementation, data analysis, and documentation. Recommendations for the application include personalizing workout recommendations, enhancing PPD screening resources, considering additional nutritional aspects, incorporating email verifications and multi-language support, and improving the rule-based approach for meal plans. Future implementations could involve optimizing meal planning and training programs using machine learning, integrating with Electronic Health Records, and incorporating a social support feature for postpartum mothers to connect and share experiences. Overall, the project achieved its goals while providing valuable insights for further enhancements and the potential to positively impact postpartum mothers' well-being.

7 References

[1] Røysted-Solås, T. et al. (2022). Mothers at risk of postpartum depression in Sri Lanka: A population-based study using a validated screening tool. PLOS ONE, 17 (5), e0268748. Available from https://doi.org/10.1371/journal.pone.0268748.

[2] American Psychiatric Association and American Psychiatric Association (eds.). (2013).

[3] Diagnostic and statistical manual of mental disorders: DSM-5, 5th ed. Washington, D.C:

[4] Badr, H.A. and Zauszniewski, J.A. (2017). Meta-analysis of the predictive factors of postpartum fatigue. Applied Nursing Research, 36, 122–127. Available from https://doi.org/10.1016/j.apnr.2017.06.010

[5] Is Obesity Associated with Major Depression? Results from the Third National Health and Nutrition Examination Survey | American Journal of Epidemiology | Oxford Academic. (no date). Available from https://academic.oup.com/aje/article/158/12/1139/90785

[6] McGarry, J. et al. (2009). Postpartum Depression and Help-Seeking Behavior. Journal of Midwifery & Women's Health, 54 (1), 50–56. Available from https://doi.org/10.1016/j.jmwh.2008.07.003.

[7] ENDRES, L.K. et al. (2015). Postpartum Weight Retention Risk Factors and Relationship to Obesity at One Year. Obstetrics and gynecology, 125 (1), 144–152. Available from

https://doi.org/10.1097/AOG.0000000000000565.

[8] Bruce, D.F. and PhD. (no date). An Overview of Postpartum Depression. WebMD. Available from https://www.webmd.com/depression/postpartum-depression/postpartum-depression

[9] Gender Issues in Depression: Annals of Clinical Psychiatry: Vol 19, No 4. (no date). Available from https://www.tandfonline.com/doi/abs/10.3109/10401230701653294

[10] Christenson, A. et al. (2016). Women's Perceived Reasons for Their Excessive Postpartum Weight Retention: A Qualitative Interview Study. PLOS ONE, 11 (12), e0167731. Available from https://doi.org/10.1371/journal.pone.0167731.

[11] Houle, J. et al. (2013). Depression self-management support: a systematic review. Patient Education and Counseling, 91 (3), 271–279. Available from https://doi.org/10.1016/j.pec.2013.01.012.

[12] Farpour-Lambert, N.J. et al. (2018). Obesity and Weight Gain in Pregnancy and Postpartum: An Evidence Review of Lifestyle Interventions to Inform Maternal and Child Health

Policies. Frontiers in Endocrinology, 9, 546. Available from

https://doi.org/10.3389/fendo.2018.00546.

[13] Shloim, N. et al. (2014). Adjusting to motherhood. The importance of BMI in predicting maternal well-being, eating behaviour and feeding practice within a cross cultural setting.

Appetite, 81, 261–268. Available fromhttps://doi.org/10.1016/j.appet.2014.06.011.

[14] Mowbray, C.T. et al. (1995). Motherhood for women with serious mental illness: pregnancy, childbirth, and the postpartum period. The American Journal of Orthopsychiatry, 65 (1),

21–38. Available from https://doi.org/10.1037/h0079588.

[15] Nikcevic, A.V., Kuczmierczyk, A.R. and Nicolaides, K.H. (1998). Personal coping resources, responsibility, anxiety and depression after early pregnancy loss. Journal of Psychosomatic Obstetrics & Gynecology, 19 (3), 145–154. Available from

https://doi.org/10.3109/01674829809025692.

[16] Barlow, J. et al. (2002). Self-management approaches for people with chronic conditions: a review. Patient Education and Counseling, 48 (2), 177–187. Available from https://doi.org/10.1016/s0738-3991(02)00032-0.

[17] Halbreich, U. and Karkun, S. (2006). Cross-cultural and social diversity of prevalence of postpartum depression and depressive symptoms. Journal of Affective Disorders, 91 (2–3), 97–111. Available from https://doi.org/10.1016/j.jad.2005.12.051.

[18] Andrews-Fike, C. (1999). A Review of Postpartum Depression. The Primary Care Companion to The Journal of Clinical Psychiatry, 01 (01), 9–14. Available from

https://doi.org/10.4088/PCC.v01n0103.

[19] Yu, Y. et al. (2022). Mental Health, Behavior Change Skills, and Eating Behaviors in Postpartum Women. Western Journal of Nursing Research, 44 (10), 932–945. Available from

https://doi.org/10.1177/01939459211021625.

[20] Ramphos, E. and Barrera, A.Z. (2016). Are Pregnant and Postpartum Women Interested in Health-Related Apps? Implications for the Prevention of Perinatal Depression | Cyberpsychology, Behavior, and Social Networking. Available from https://www.liebertpub.com/doi/abs/10.1089/cyber.2015.0549

Influence of Internal Employer Branding and Organizational Trust on Employee Engagement In The IT Industry in Sri Lanka

Sashini Vasanthan

IIT Business School, Informatics Institute of Technology, Colombo, Sri Lanka, sashini.20191122@iit.ac.lk

Gayashini Ratnayake

IIT Business School, Informatics Institute of Technology, Colombo, Sri Lanka, gayashini.r@iit.ac.lk

ABSTRACT

This study was conducted to identify the influence of internal employer branding and organizational trust on employee engagement in the IT industry in Sri Lanka. The research problem in this context is the limited amount of research on the impact of internal employer branding on employee engagement especially in IT industry in Sri Lanka. While there is significant evidence linking employee engagement to factors such as job satisfaction (Susan, cited in Alam et al., 2022), intrinsic rewards (Ghose and Mohanty; Ram and Prabhakar, cited in Alam et al., 2022), extrinsic rewards (Ghose and Mohanty; Ram and Prabhakar, cited in Alam et al., 2022), and work-life balance (Wood et al., 2020), the specific influence of internal employer branding on employee engagement has not been extensively studied (Piyachat, Chanongkorn and Panisa, 2014). Therefore, this study was undertaken to examine the mediating effect of the organizational trust on the relationship between internal employer branding and employee engagement. A deductive approach was adopted, and data was collected using a convenience sampling method. Questionnaire was distributed to 300 IT employees who work at executive and non-executive positions in Sri Lanka and 205 valid responses were collected with a response rate of 68%. Based on the responses quantitative analysis was done using SPSS version 23 and reliability, validity and hypothesis testing were carried out. Additionally, Sobel test was done for mediator.

The findings of the study revealed a positive relationship between internal employer branding, organizational trust, and employee engagement but the relationship between organizational trust and employee engagement was rejected. It was also identified that organizational trust mediates the relationship between employer branding and employee engagement. Accordingly, this study concluded that Internal Employer Branding has a positive relationship with Employee Engagement and Organizational Trust. Additionally it was found that there is no relationship between Organizational Trust and Employee Engagement and that organizational trust does not mediate the relationship between employer branding and employee engagement.

Based on this study main recommendations are that employers should identify the characteristics that determine value for the employees and employers should also examine employees' main expectations. Internal organizational research should also be conducted frequently to obtain the useful information such as employee satisfaction, motivation level and effectiveness of tools used by organizations.

KEYWORDS

Internal employer branding, Organizational trust, Employee Engagement IT industry, Sri Lanka

1 Introduction

1.1 Background

In this modern era attracting and retaining human resources have become one of the main elements for organizational success (Osteraker, cited in Staniec and Kalińska-Kula, 2021). One of the main concepts of employer branding is that organization commits to employees and satisfies their needs which in turn improves the employee's ability to meet their customer needs. (Staniec and Kalińska-Kula, 2021). Employer branding has been used as a strategic tool to attract and retain future and current employees in this competitive era (Jahan and Porkodi, 2022).

Edwards & Edwards (2013) stated that employer branding spreads awareness of employees and might leverage organizational branding success for personal gain (Staniec and Kalińska-Kula, 2021). It is a process of informing the employees that the organization is highly attractive and have various benefits (Moroko & Uncles cited in Staniec and Kalińska-Kula, 2021). External and internal branding focuses on potential and existing employees respectively. (Hochegger, 2014; Sangupta, Bamel and Singh, 2015, cited in Morya and Yadav, 2017).

Internal employer branding is the internal promotion where employees are considered as target audience (Punjaisri, Evanschitzky & Wilson, cited in Morya and Yadav, 2017). Internal employer branding including; well communicating the brand values, providing training programs and development, coaching, conducting team building activities, internal

recruitment and effective individual performance appraisal system (Staniec and Kalińska-Kula, 2021; Morya and Yadav, 2017).

Employee engagement means passion, involvement and enthusiasm towards their work (Eldor and Vigoda-Godot, cited in Morya and Yadav,2017). Engaged employee understands the company's goals. Company with effective engagement strategy and highly engaged force more likely to retain and attract talents (sundar, 2020).

Main motive of employer branding activities is to increase employee engagement (Verčič, 2021). As past researchers stated, when employee engagement is high, employers are performed well which leads to high revenues, profits and overall returns on investment resulting in a sustainable competitive company. With effective branding startegy, level of engagement can be increased (Piyachat, Chanongkorn and Pansia, 2014; Porkodi and Jahan, 2022).

IT industry in Sri Lanka has been growing and a main sector driving economic development. Due to competitiveness, in recent years, IT industry has faced shortages of talent which has led to high demand for IT professionals and increased competition among employers (Anuradha; Lo, cited in Dabirian, Berthon, and Kietzmann,2019). This has increased the pressure to attract and retain talent. So, firms have to show the potential employees that the organization is a great place to work. Good employer branding communicates this and creates a competitive advantage (Dabirian, Berthon, and Kietzmann,2019). Gallup report has stated that only 38% of employees in Sri Lanka are engaged (Sanduni Tharika, Anuradha Iddagoda and Helena Bulinska-Stangrecka, 2022.)

When employer branding activities are done to increase employee engagement, organizational trust should be considered. Trust develops a strong and sustainable relationship between employer and employee (Burt and Kenz, cited in Yadav, Kumar and Mishra, 2022).

1.2 Research Gap

There is significant evidence relating to the antecedents of employee engagement such as job satisfaction (Susan, cited in Alam et al., 2022), intrinsic rewards (Ghose and Mohanty; Ram and Prabhakar, cited in Alam et al., 2022), extrinsic rewards(Ghose and Mohanty; Ram and Prabhakar, cited in Alam et al., 2022), work-life balance (Wood et al., 2020). However, limited research exists on the impact of internal employer branding on employee engagement (Piyachat, Chanongkorn and Panisa, 2014).

Organizational trust plays a key role on the degree of employee engagement (Acosta, Salanova and Liorens, 2012). Internal employer branding activities improve organizational trust (Yadav, Kumar and Mishra, 2020). Therefore, mediating effect of organizational trust on internal employer branding and employee engagement was identified and analysed in this study.

1.3 Research Question

RQ1: How does internal employer branding activities affect employee engagement?

RQ2: What is the influence of employer branding activities on organizational trust?

RQ3: What is the influence of organizational trust on employee engagement?

RQ4: What are the factors that influence employee engagement?

1.4 Research Objectives

The research objectives were derived from the research questions above.

RO1: To analyse the influence of internal employer branding activities on employee engagement

RO2: To analyse the influence of employer branding activities on organizational trust

RO3: To assess the influence of organizational trust on employee engagement

RO4: To analyse the factors that influence employee engagement

1.5 Scope of the study

This study is designed to understand the impact of internal employer branding on employee engagement. It focused on IT companies in Colombo. Those who work in IT industry at executive and non-executive level were investigated in this study.

2 Literature Review

2.1 Main Theories and Models

2.1.1 Virtuous Circle of employer branding

Virtuous circle covers the key aspects of internal and external employer branding and their effects (Moroko & Uncles cited in Staniec and Kalińska-Kula, 2021). According to this, employer branding activities lead to attraction and retention of valuable employees who represents the company proudly meet the customer expectations. These employees are likely to be engaged and committed to their work. So, customers become loyal to the brand and improves company's reputation and profitability. The connection between employee satisfaction, customer satisfaction, and company profitability is a critical aspect of the virtuous circle of employer branding. This virtuous circle closes when a well reputed profitable company attract and retain talented employee who wants to contribute to the company's success and be connected to it (Staniec and Kalińska-Kula, 2021).

2.1.2 Social Exchange Theory

Social exchange theory states that obligations are generated through a series of interactions between parties who are in the state of reciprocal interdependence (Kibara and Kiiru, 2021; Kular et al., 2008). In other words, it states that exchange is the best way to show appropriate behaviours valuable to their organization (Alshaabani, Hamza and Rudnák, 2021). According to social exchange theory, relationships between two parties are built and developed through time and are based on mutual commitment, loyalty and trust which is bound by a reciprocity rule known as exchange rule (Cropanzano and Mitchell cited in Piyachat, Chanongkorn and Panisa 2015). Such rule involve reciprocity. So, the actions of one party lead to a response or actions by the other party. For example, when employees receive economic and social-emotional resources from their workplace, they feel obligated to repay organizations (Kular et al., 2008).

Sakes (2006) states that the level of engagement is one way to repay. Employees will engage themselves based on the management practices and resources they receive from their organization. Therefore, employees are more likely to exchange their engagement for positive practices and resources and benefits provided by their organization (Alshaabani, Hamza and Rudnák, 2021; Kular et al., 2008).

Therefore, social exchange theory provides a theoretical foundation why employees choose to become more or less engaged in their work and organizations. When these are not provided by organizations employees are more likely to withdraw or disengage themselves (Kular et al., 2008).

This theory also shows the relationship between organizational trust and employee engagement. Trust in an organization will be reciprocated by engagement which will reduce the idea of quitting the organization (Yadav, Kumar and Mishra, 2020).

2.1.3 Expectancy Theory

According to expectancy theory, in the given opportunity a person will choose to work only when the value of the outcomes expected to achieve from work is more positive than not working (Vroom, cited in Piyachat, Chanongkorn and Panisa 2015). This theory explains the relationship between employees' efforts and their performance which impacts the performance of the organization (Kibara and Kiiru, 2021). Vroom and Deci (1992) stated that employee engage in behaviours as they expect these behaviours lead their goals (Piyachat, 2015).

This theory explains how individuals make decisions on various behavioural alternatives. According to the theory, employees will select the options with the greatest motivation factors associated with three conditions which are expectancy, instrumentality and valence and employees felt motivated when all three conditions are present (Kibara and Kiiru, 2021; Piyachat, 2015).

Expectancy theory uses both extrinsic motivating factors such as salary, compensations, bouns, commissions and benefits and intrinsic motivating factors such as sense of pride, dignity and values obtained from job to describe the causes for behaviour in the workplace (Victor and Hoole cited in Kibara and Kiiru, 2021; Piyachat, 2015).

2.1.4 Concentric circle model

This model explains how the employer brand, which is referred to as a set of particular employment experience attributes make the organization distinctive and attractive to its current and future potential employees. Accordingly, special importance is given on the internal perspectives by answering the fundamental questions of employees (Bendaravičienė, 2017). Therefore, this model explains how employer branding attracts current and future employees.

2.1.5 HERO model

Organizational trust is considered as one of the key elements in the HERO model. It is included under the category of healthy employees (Acosta, Salanova and Llorens, 2012). Healthy employees are employees with positive psychological resources such as organizational trust, self-efficacy, mental and emotional competence which impacts the work engagement positively (Gil-Beltrán, Salanova and Llorens, 2020; Acosta, Salanova and Llorens, 2012).

2.1.6 Job Demands-Resources model (JD-R model)

This model explains how to improve employee engagement. According to Moura and Ramos (2014), engagement can be created through working conditions such as job demands and job resources (cited in Alam et al., 2022; Wood et al., 2020). Accordingly managers have a significant impact on employee engagement in terms of coaching and training which are indicators of internal branding activities. This model also predicts work engagement and burnout (Alam et al., 2022).

3 Conceptual Model

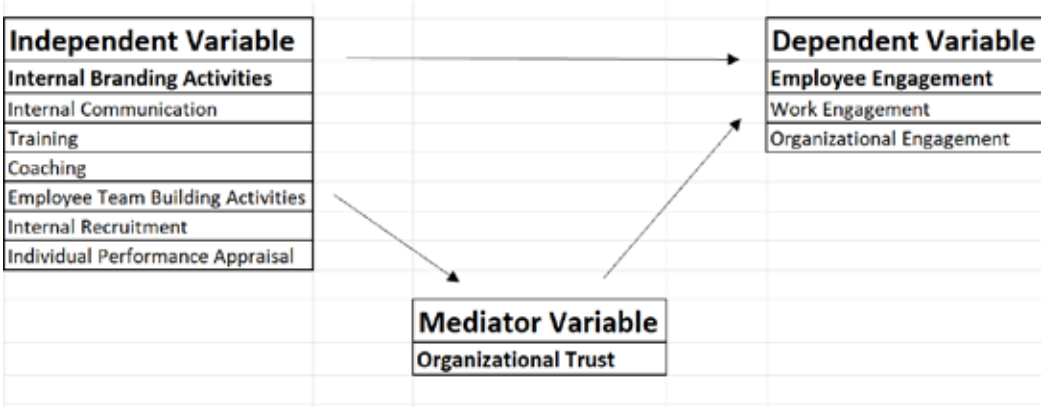


Figure 1: Conceptual Framework

3.1 Key Terms

3.1.1 Employee Engagement

The dependent variable is employee engagement. It has two dimensions; work engagement and organizational engagement (Kunerth and Mosley, 2011; Morya and Yadav, 2017; Kular et al., 2008). Employee engagement is defined as, 'the harnessing of organization members' selves to their work roles' (Alam et al., 2022; Kular et al., 2008). It is the extent of eagerness and connection employees have with their organization and can be defined by job satisfaction, involvement, organizational commitment and feeling empowered (Porkodi and Jahan, 2022). It involves physiological and physical presence in performing the work (Kahn, cited in Verčič, 2021).

Engagement was separated into work and organizational engagement firstly by Saks (2006) and it was explained through social exchange theory (Piyachat, Chanongkorn and Panisa 2015; Kular et al., 2008). He defined engagement as a distinct and unique construct that includes cognitive, emotional and behavioral components associated with individual role performance (Piyachat, Chanongkorn and Panisa 2015, p60; Morya and Yadav, 2017; Alam et al., 2022).

Cognitive dimensions were associated with commitment and satisfaction (Piyachat, Chanongkorn and Panisa 2015). The cognitive aspect concerns employees' belief in the organization, leaders and working conditions (Lockwood cited in Alam et al., 2022). Emotional dimensions were associated with the feeling about relationships with other colleagues and being able to express oneself without worrying about the negative consequences (Piyachat, Chanongkorn and Panisa 2015). The emotional aspect concerns how employees feel and their attitudes toward the organization, co-workers and their leaders (Alam et al., 2022). Behavioral dimensions such as behaviors supporting organizational effectiveness,

encouraging change and innovation and retention were associated with adaptive behavior (Piyachat, Chanongkorn and Panisa 2015; Kular et al., 2008).

Employee engagement leads to higher level of customer satisfaction, loyalty, productivity and profit (Harter, Schmidt and Hayes, cited in Verčič, 2021). According to Kahn (1992) engagement leads to superior work quality and positivity and enthusiasm in work and also leads to higher and sustained productivity and enhanced financials (Yadav, Kumar and Mishra, 2020). Harter, Schmidt and Hayes (2002) also stated that employers have impact on the level of employee engagement and that all employees can be engaged (Verčič, 2021). Therefore, Employee engagement will be achieved only when employees realize the expectation of the employers and when opportunities for their skill utilization are provided. (Lee and Ok cited in Morya and Yadav, 2017).

Employees engage themselves with different degrees in an organization. Employees are actively engaged when they feel passionate towards their work and enjoy and feel very connected (Alam et al., 2022; Morya and Yadav, 2017). Employees are not engaged when they are not enthusiastic regarding work or innovative approach and do only the allocated works and employees are disengaged when they become completely negative and shows unhappiness (Gallup cited in Morya and Yadav, 2017). In order to explain these different engagement levels, Saks (2006) proposed social exchange theory and Vroom and Deci (1992) proposed expectancy theory (Piyachat, Chanongkorn and Panisa 2015; Sundar, 2020). Empirical Studies of Piyachat, Chanongkorn and Panisa, 2015; Porkodi and Jahan, 2020; Morya and Yadav, 2017; Piyachat, 2015; Yadav, Kumar and Mishra, 2020; Mayuran and Kailasapathy, 2022 show that internal employer branding and employee engagement has a positive relationship. Additionally, virtuous circle of employer branding, social exchange theory and Job-Demands resources model support the relationship of internal employer branding and employee engagement.

H1: There is a significant positive relationship of internal branding activities on employee engagement

3.1.2 Employer Branding

The independent variable is internal branding activities. It has six dimensions; internal communication, training, coaching, employee team building activities, internal recruitment and individual performance appraisal (Staniec and Kalińska- Kula, 2021)

Employer branding is defined as the package of functional, economic and psychological benefits provided by employment and identified with the employing company [Sokro, 2012, p165; Porkodi and Jahan, 2022; Verčič, 2021]. Ambler and Barrow (1996) applied the concept of brand to HRM initially by considering the employer as the brand and employees as customers (Staniec and Kalińska-Kula, 2021; Sokro, 2012; Morya and Yadav, 2017). This provides economic and psychological benefits to employees.

Employer branding includes firm's values, systems, policies and behaviour which aim to attract, retain and motivate employees (Edwards, 2009). Organizations with strong employer branding standouts from rivals and are more likely to attract and retain employees which leads to the achievement of a long-term competitive advantage through their employees. Therefore, employer branding also increases productivity and competitiveness (Porkodi and Jahan, 2022).

Management of employer image is considered as external employer branding and management of identity is considered as internal employer branding (Lievens & Slaughter cited in Staniec and Kalińska-Kula, 2021). Accordingly, Ouchi (1981) defines internal employer branding as an activity that creates an environment where current employees will support brand and achieve the organization goals (Staniec and Kalińska-Kula, 2021). Internal branding shapes the organizational identity and culture and brings better employee outcomes (Backhaus, cited in Kashyap and Chaudhary, 2019). Empirical studies of Kashyap and Chaudhary, 2019; Ekhsan and Saroh, 2021 show that internal branding activities and organizational trust has a positive relationship and social exchange theory studies about this relationship.

H2: There is a significant positive relationship between internal branding activities and Organisational Trust

3.1.3 Organizational Trust

The mediator variable is organizational trust (Yadav, Kumar and Mishra, 2020). Organizational trust is defined as a feeling of confidence and support in an employer and it is believed that when an organizational trust exists, employees are straightforward and will follow through the commitments (Gilbert and Tang, 1998). In other words, organizational trust means the willingness of the employees to be exposed to the decisions made by their companies, whose behaviour and actions cannot be controlled (Acosta, Salanova and Liorens, 2012). Trust is considered as the basic concept that holds the balance in an organization and fulfilment among employees (Cook and Wall cited in Yadav, Kumar and

Mishra, 2020). It is considered as an important thing in work life and organizational effectiveness (Acosta, Salanova and Liorens, 2012). According to Burt and Kenz (1996) trust is seen as a cure for adamancy in organizations and it is also considered as a moderator between employer and employee to develop the robust and sustainable affiliations (Yadav, Kumar and Mishra, 2020).

Butler and Cantrell (1984) stated integrity, competence, consistency, loyalty and openness as fundamental traits of trust (Yadav, Kumar and Mishra, 2020; Ekhsan and Saroh, 2021) and Mayer et al., (1995) stated ability, benevolence and integrity as attributes of trust (Yadav, Kumar and Mishra, 2020). In addition, Cunningham and McGregor (2000) and Mishra (1996) added predictability as an important feature of trust (Yadav, Kumar and Mishra, 2020). Trust is strengthen by organizational aspects such as structure, people management policies and organizational culture (Whiteneret et al., cited in Yadav, Kumar and Mishra, 2020).

There can be disappointment among employees and they can lose trust on the organization, if the organization is not able to keep up their promises announced through the employer branding initiatives (Barrow and Mosley and Pilbeam and Corbridge, cited in Yadav, Kumar and Mishra, 2020). Trust is considered as an important mechanism which drives a positive interrelation between employer branding and employee engagement (Kashyap and Chaudhary, 2019). Therefore, employer branding should increase trust in organization. Stronger the employer branding, stronger should be trust in organization (Yadav, Kumar and Mishra, 2020). Relationship between organizational trust and employer branding can be showed through Social Exchange theory (Yadav, Kumar and Mishra, 2020). Acosta, Salanova and Liorens, 2012; Kashyap and Chaudhary, 2019; Alshaabani, Hamza and Rudnák, 2021; Mohanty and Arunprasad (2020) shows that there is a positive relationship between organizational trust and employee engagement. Social exchange theory and Hero model studies and supports this relationship.

H3: There is a significant positive relationship between organizational trust and employee engagement

Additionally studies of Yadav, Kumar and Mishra, 2020; Kashyap and Chaudhary, 2019 show that organizational trust mediates the relationship between internal employer branding activities and employee engagement

H4: Organizational trust mediates the relationship between internal employer branding activities and employee engagement

4 Research Methodology

4.1 Research Philosophy

This research study aims to answer the question on how internal employer branding activities affect the employee engagement and on how organizational trust mediate the relationship between internal employer branding activities and employee engagement. This study identifies and empirically test these relationships and hence the research philosophy is positivism (bell, 2022). Positivism assumes scientific approach such as observation and knowledge to the development of knowledge rather than intuition or speculation (Bougie and Sekaran, 2019).

4.2 Research Approach

This study aims to test the relationship between the variables which were identified based on existing theories and models. Therefore, this research uses the deductive approach (Bougie and Sekaran, 2019).

4.3 Research Method

Quantitative research method was used to test the relationships using quantitative research techniques according to positivistic paradigm. It uses statistical tools and models (Bougie and Sekaran, 2019)..

4.4 Research Instrument Development

A cross sectional survey was used to test the hypotheses and achieve the research objectives. The survey instrument was designed by adopting measures used in prior empirical research studies. The cross sectional survey was based on a 5-point Likert Scale ranging from Strongly Agree to Strongly Disagree was used. The Likert scale offered theoretically equal distances between the responses and was considered suitable for the research study. Internal branding activities was operationalized using 6 indicators (Communication, Coaching, Employee Team Building, Training, Internal Recruitment, Individual Performance Appraisal Systems) on an ordinal scale ranging from 1 to 5

adopted by Staniec and Kalińska-Kula, (2021). Organisational trust was operationalized using measures adopted from Yadav, Kumar and Mishra, (2020) using an ordinal scale ranging from 1 to 5. Employee engagement was operationalized using two indicators (Work Engagement, Organizational Engagement) using measures adopted by Morya and Yadav (2017) using an ordinal scale ranging from 1 to 5.

4.5 Sampling Strategy

4.5.1 Sampling Technique

Convenience sampling method was used as the sampling technique, where responses are collected from members of the population who are conveniently available (Bougie and Sekaran, 2019). It was chosen because respondents were easy to access and was able to collect the responses in a short period of time. Similar sampling technique was used by Staniec and Kalińska-Kula (2021), Sokro (2012), Acosta, Salanova and Liorens (2012) for the studies related to internal employer branding, organizational trust and employee engagement.

4.5.2 Sample Size

This research was aimed to collect data from employees at executive and non-executive level in any department who work for a IT company in Sri Lanka. Prior to the survey, a pilot study was carried out among 30 respondents. Later questionnaire was distributed to 242 participants and 205 valid responses were collected. The acceptable response rate was 84.7%.

5 Data analysis and Representation

5.1 Data Preparation

Data was checked for normality using Skewness and Kurtosis test, Kolmogrov and Smirnov and Shapiro Wilk test and it was identified that the data distribution is normal. In order to check the missing value using univariate testing was done and missing values were deleted. Existence of multicollinearity was tested using Tolerance and Variance Inflation Factor values and concluded that multicollinearity does not exist.

5.2 Correlation Analysis

Correlation analysis identifies the significance of the relationships between the variables. According to Cohen's Pearson Correlation value, internal branding and organizational trust have a strong positive relationship (0.536), internal branding and employee engagement have a strong positive relationship (0.651) and organizational trust and employee engagement have a moderately positive relationship (0.394).

Table 1: Correlation Analysis Relationship

Relationship	Pearson Correlation R Value	Strength of the Relationship
Internal Branding Activities – Organizational Trust	0.536	Large
Internal Branding Activities – Employee Engagement	0.651	Large
Organizational Trust – Employee Engagement	0.394	Medium

Cohen (1988)

5.3 Regression Analysis

Regression analysis is used to explain the significance of the key variables and R Square value is used to analyze this. Accordingly, below table shows that internal branding activities and organizational trust predict employee engagement 42.7% of the time.

Table 2: Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.653	.427	.421	.40068

Predictors: (Constant), Organizational Trust, Internal Branding Activities

As the ANOVA table and coefficient table shows that the significance value is within the threshold value p<0.05, it can be concluded that internal branding activities predict organizational trust significantly. Therefore hypothesis 1 is achieved.

Table 3: ANOVA

Model		Sum of Squares	Df	Mean Square	F	Sig
1	Regression	24.165	2	12.082	75.259	0.000
	Residual	32.430	202	.161		
	Total	56.595	204			

Dependent Variable: Employee Engagement

Predictors: (Constant), Organizational Trust, Internal Branding Activities

Table 4: Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig
		B	Std. Error			
1	(Constant)	1.278	.257		4.974	.000
	Internal Branding Activities	.642	.066	.618	9.791	.000
	Organizational Trust	.086	.087	.062	.989	.324

Dependent Variable: Employee Engagement

Based on the beta value, internal branding activities 0.618 makes the largest contribution to the dependent variable, employee engagement.

Internal branding activities has a positive relationship on organizational trust (R=0.536) (Table 7), and the significance value is 0.000 (Table 11)and the there is a positive and significant relationship between internal branding activities and organizational trust and hence hypothesis 2 is achieved.

Organizational trust has a positive relationship on employee engagement (R=0.394) (Table 7), but since significant value is greater 0.001 (Table 10), the relationship is not significant and hence hypothesis 3 is not achieved.

5.4 Mediator Analysis

Table 5: Mediator Coefficients 1

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error			
(Constant)	1.638	0.173		9.469	0.000
Internal Branding Activities	0.405	0.045	0.536	9.054	0.000

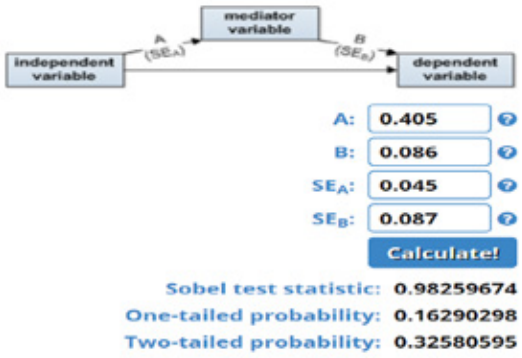
Dependent Variable: Organizational Trust

Table 6: Mediator Coefficients 2

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error			
(Constant)	1.278	0.257		4.974	0.000
Internal Branding Activities	0.642	0.066	0.618	9.791	0.000
Organizational Trust	0.086	0.087	0.062	0.989	0.324

Dependent Variable: Employee Engagement

Table 7: Sobel Test



The p value is not significant (since P value is greater than 0.05) and organizational trust does not play a significant mediator effect. Therefore, Hypothesis 4 is rejected.

5.5 Justification of the Research Instrument

5.5.1 Reliability Analysis

Using Cronbach Alpha method reliability of the research instrument was assessed. If the Cronbach Alpha value exceeds 0.7, research instrument is considered to be reliable (Mishra et al., 2019; Tavakol and Dennick, 2011. As Cronbach's Alpha is 0.953, the research instrument is reliable.

5.5.2. Validity

5.5.2.1. Sample Adequacy

The adequacy of the sample was measured using Kaiser-Meyer-Olkin and as it exceeds 0.5 (Budaev, 2010; Goni et al., 2020), the sample is adequate. As the p value is less than 0.05 Bartletts test of sphericity is highly significant and hence the sample size is adequate.

5.5.2.2. Convergent Validity

In order to satisfy the convergent validity, inter-item correlations should exceed 0.3 (Campbell and Fiske, 1959). All main variables satisfy this threshold.

5.5.2.3. Construct Validity

A pilot test was carried out among 30 respondents and the Cronbach Alpha Value exceeded 0.7. Hence, the construct validity was achieved.

5.5.2.4. Discriminant Validity

Discriminate Validity assesses whether the measure is distinct from other measures of different constructs. As significant values of Wilks Lambda test are less than 0.05, independent variables have a significant relationship with the dependent variable. Eigen values are less than 1 hence, the construct is discriminant. Smaller the eigen value, higher the discriminant validity.

.

5.6 Demographic Analysis

Majority of the responses are from male employees (53.7%) in the IT sector and majority of the respondents are age between 21-29 years (61%). Majority respondent's income level is between below 30,000 LKR (22.9%) and based on the data analysis it was also identified that majority of the respondents who earn more than 150,000 LKR are male. Based on the data analysis significant difference in employee engagement between male and female employees were identified in the IT industry by performing t-test. This supports the empirical findings of Porkodi and Jahan, 2020; Kular et al., 2008. Kular et al., (2008) states that compared to men women engage more in the organization.

6 Conclusion and Recommendations

6.1 Discussion of Research Objectives and Findings

This study revealed that internal employer branding activities have a significant and positive influence on employee engagement. This supports the empirical findings of Piyachat, Chanongkorn and Panisa, (2015); Porkodi and Jahan, (2020); Morya and Yadav, (2017); Piyachat, (2015); Yadav, Kumar and Mishra, (2020); Mayuran and Kailasapathy, (2022). Based on the previous studies (Staniec and Kalińska-Kula, 2021) and frameworks it was identified that communication, training, coaching, employee team building activities, internal recruitment and individual performance appraisal system are indicators of internal employer branding activities that influence the employee engagement and the data findings support these empirical findings.

The study revealed that internal employer branding activities have a significant and positive impact on organizational trust. Kashyap and Chaudhary (2019); Ekhsan and Saroh, (2021) empirical research support these findings. Regression analysis also revealed a strong and positive relationship between internal branding activities and organizational trust.

Organizational trust did not have a positive and significant influence on employee engagement. This supported the previous empirical finding of Herminingsih (2017). But this contradicted the empirical findings of Acosta, Salanova and Liorens, (2012; Kashyap and Chaudhary, (2019); Alshaabani, Hamza and Rudnák, (2021); Mohanty and Arunprasad (2020) where they found that organizational trust has a significant influence on employee engagement.

Organizational trust does not mediate the relationship between internal employer branding activities and employee engagement and this contradicted the previous empirical findings by Yadav, Kumar and Mishra, (2020); Kashyap and Chaudhary, (2019).

Regression analysis identified that internal employer branding predicts employee engagement in an organization by 65.3%. Some of the other variables that predict the employee engagement are; employee expectation, empowerment, commitment, autonomy, job rotation, business culture and corporate networking. This supports the empirical findings of Yadav, Kumar and Mishra, 2020 where they have identified that employer branding and organizational trust predicts employee engagement 72% of the time

Data distribution was identified as normal based on the normality test and was identified that multicollinearity does not exist. This ensured that the results are reliable and accurate and hence, data set was considered appropriate for further analysis. Further, in order to identify the reliability of the variable, Cronbach's Alpha method was used and the result was more than 0.7. KMO value came more than the threshold value 0.5 and the sample size was considered adequate.

Based on the analysis, following regression equation is proposed.

Employee Engagement = 1.278 + 0.086 * (mean of Organizational Trust) + 0.642 * (mean of Internal Branding Activities)

Companies in the IT sector relies on limited skilled labour (Dabirian, Berthon and Kietzmann, 2019). Employer branding shows current and future employees that the firm is suitable and beneficial to work. So it is important for companies in IT industries to implement a suitable branding strategy and attract and retain employees. As identified in the study, employer branding and organizational trust have a positive relationship with employee engagement which leads to increase in organizational performance. So, companies should focus more on improving the employee engagement.

6.2 Limitations of the study

• Inadequate research articles as most research were conducted in different industries and not in IT industry in Sri Lanka. Therefore, research specifically in the IT industry identifying the influence of internal employer branding and organizational trust on employee engagement was low.
• Time was constraint to find out the underlying factors that impacts the employee engagement. therefore, only two variables were considered as influencing employee engagement. There can be other variables that influences the employee engagement such as individual factors, country factors and industry factors.
• Samples selected were too small and selection was based on convenience sampling which may have the selection bias on choosing the respondents.
• As data was collected through 5-point Likert scale, respondents had limited response options which may not accurately reflect respondents' options/ attitudes. Additionally, social bias may be present where respondents maybe forced to respond according to socially desirable rather than their true opinions/ attitudes.

6.3 Recommendations

6.3.1 Recommendations from the findings

- Characteristics that determine value for the employees should be identified and when identifying it is necessary to inspect employees' main expectations (Staniec and Kalińska- Kula, 2021).
- It is necessary to conduct internal organizational research frequently. This helps to obtain valuable information in terms of level of employee motivation and satisfaction, effectiveness of the tools used by the organization such as internal communication, training (Staniec and Kalińska- Kula, 2021).
- As internal branding activities and employee engagement have a strong positive relationship, organizations should Focus more on building the employer branding to improve employee involvement and satisfaction (Verčič, 2021).
- Organization can enhance the HR practices such as flexible work arrangements, training programmes, and incentive rewards which positively relates to employee engagement (Kular et al., 2008)

6.3.2 Recommendations for future research

- In order to generalize the results more, future study should consider other main cities with IT firms such as Kandy, Galle and Jaffna with more representatives.
- Additionally, research can be done to find out how each indicators of Internal Employer Branding impact the employee engagement and how employee engagement can be further improved.
- Qualitative method should be applied in the future research in order to gain more insight information with regard to this topic.
- The future research can identify the possible variables into models which can be moderators or mediators that can lead to a better understanding

References

[1] Amir Dabirian, Pierre Berthon and Jan Kietzmann, J. 2019. Enticing the IT crowd: employer branding in the information economy. Journal of Business and Industrial Marketing, 34, 7, 1403-1409. DOI: https://doi.org/10.1108/JBIM-11-2018-0333
[2] Ana Tkalac Verčič. 2021. The impact of employee engagement, organisational support and employer branding on internal communication satisfaction. Public Relation Review, 471, 1-7. DOI: https://doi.org/10.1016/j.pubrev.2021.102009
[3] Anik Herminingsih. 2017. The Influence of the Organizational Justice and Trust to the Leaders on Employee Engagement with Job Satisfaction as Intervening Variable. Archives of Business Research, 5,2, 56-59. DOI: 10.14738/abr.52.2602
[4] Ayman Alshaabani, Khadija Aya Hamza and Ildiko Rudnák. 2021. Impact of Diversity Management on Employees' Engagement: The Role of Organizational Trust and Job Insecurity. Sustainability, 14, 1, 420-441. DOI: **https://doi.org/10.3390/su14010420**
[5] Bernard Kunerth and Richard Mosely. 2011. Applying employer brand management to employee engagement. Strategic HR Review, 10, 3, 19-26. DOI: https://doi.org/10.1108/14754391111121874
[6] Campbell, D. and Fiske, D. 1959. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological bulletin, 56, 2, 81-105. DOI: https://psycnet.apa.org/doi/10.1037/h0046016

[7] Ester Gil-Beltrán, Marisa Salanova and Susana Llorens. 2020. Employees' Physical Exercise, Resources, Engagement, and Performance: A Cross-sectional Study from HERO Model. Journal work and Organizational Psychology, 36, 1, 39-47. DOI: https://doi.org/10.5093/jwop2020a4
[8] Iwona Staniec and Magdalena Kalińska-Kula, M. 2021. Internal employer branding as a way to improve employee engagement. Problems and perspective in management, 19, 3, 33-45. DOI: http://dx.doi.org/10.21511/ppm.19(3).2021.04
[9] Jacqueline A. Gilbert and Thomas Li-Ping Tang. 1998. An Examination of Organizational Trust Antecedents. Public Personnel Management, 27, 3, 321-338. DOI: https://doi.org/10.1177/009102609802700303
[10] Jasmine Alam, Morris Mendelson, Mustapha Ibn Boamah and Mathieu Gauthier. 2022. Exploring the antecedents of employee engagement. International journal of organizational analysis, DOI: https://doi.org/10.1108/IJOA-09-2020-2433
[11] Logendran Mayuran and Pavithra Kailasapathy. 2022. To engage or not? Antecedents of employee engagement in Sri Lanka. Asia Pacific Journal of Human Resources, 60, 3, 584-607. DOI https://doi.org/10.1111/1744-7941.12270
[12] Martin R. Edwards. 2009. An integrative review of employer branding and OB theory. Personnel review, 39, 1, 5-23. DOI: https://doi.org/10.1108/00483481011012809
[13] Mohammed Dauda Goni, Nyi Nyi Naing, Habsah Hasan, Nadiah Wan-Arfah, Zakuan Zainy Deris, Wan Nor Arifin, Tengku Hussin, Abdulwali Abdulrahman, Aisha Baaba and Muhammed Arshad. 2020. Development and validation of knowledge, attitude and practice questionnaire for prevention of respiratory tract infections among Malaysian Hajj pilgrims. BMC Public Health, 20, 1, 1-10. DOI: https://doi.org/10.1186/s12889-020-8269-9
[14] Mohsen Tavakol and Reg Dennick. 2011. Making sense of Cronbach's alpha. International journal of medical education, 2, 53-55. DOI: https://doi.org/10.5116%2Fijme.4dfb.8dfd
[15] Muhamed Ekshan and Tim Saroh. 2021. The Role of Organizational Trust on the Influence of Employer Branding on Employee Performance. International Journal of Management 44 Science and Information Technology, 1, 2, 31-35. DOI: https://doi.org/10.35870/ijmsit.v1i2.353
[16] Rita Bendaravičienė, 2017. TOWARDS SUSTAINABLE ORGANIZATION: INTEGRATIVE CONCEPTUAL MODEL FOR EMPLOYER BRANDING. JOURNAL OF SECURITY AND SUSTAINABILITY ISSUES, 6, 4, 649-662. DOI: https://doi.org/10.9770/jssi.2017.6.4(10).
[17] Samarendra Kumar Mohnaty and Arunprasad P. 2020. Identification of drivers of employee engagement in Indian power companies. International Journal of Productivity and Performance Management, 70, 6, 1263-1290. DOI: https://doi.org/10.1108/IJPPM-08-2019-0414
[18] Sanduni Tharika, Anuradha Iddagoda and Helena Bulinska-Stangrecka. 2022. Employee Engagement and its Dynamics: An Empirical Study of an Information Technology Company Sri Lanka. International Journal of Academic and Institute of Research,2, 4, 1-15. DOI: https://doi.org/10.53378/352082
[19] Sergey V. Budaev, S. 2010. Using principal components and factor analysis in animal behaviour research: caveats and guidelines. Ethology, 116, 5, 472-480. DOI: **https://doi.org/10.1111/j.1439-0310.2010.01758.x**
[20] Sundar, A. 2020. Employee Engagement in Retail Sector. International Journal of Science and Research, 9, 12, 708-715. DOI: 10.21275/SR201127123210
[21] Vaneeth Kashyap and Richa Chudhary, 2019. Linking Employer Brand Image and Work Engagement: Modelling Organizational Identification and Trust in Organization as Mediators. South Asia Journal of Human Resources Management, 6, 2, 177-201. DOI: https://journals.sagepub.com/doi/pdf/10.1177/2322093719844644
[22] Wood et al. 2020. The Relationship Between Work Engagement and Work–Life Balance in Organizations: A Review of the Empirical Research. Human Resource Development Review, 19,23, 240-262. DOI: https://doi.org/10.1177/1534484320917560
[23] Yadav, L., Kumar, A. and Mishra, S. 2020. Can Organizations Leverage Employer Branding to Gain Employee Engagement and Reduce Turnover Intention? An Empirical 47 Study with Organizational Trust as a Mediator. Vision, 24,4, 496-505. DOI: https://doi.org/10.1177/0972262920943774

Ambula - A Mobile Application Solution to Promote Sri Lankan Food Culture

Sarah Hausotter

Business School, Informatics Institute of Technology/University of Westminster, Colombo 06, Western Province, Sri Lanka, sarah.2018061@iit.ac.lk

Lakshika Paiva

WSO2 Lanka Pvt. Ltd., Colombo 04, Western Province, Sri Lanka, lakshika.paiva@gmail.com

ABSTRACT

Tourism plays a crucial role in Sri Lanka's economy, contributing significantly to the country's Gross Domestic Product. The tourism sector encompasses various sub- sectors, including Food Tourism, also known as Gastro-Tourism, which is a popular and effective method of promoting destination loyalty by combining cultural and culinary experiences. Despite Sri Lanka's diverse and flavourful cuisine, there has been little momentum to propagate resulting-showing growth and visibility on the global food stage. This thesis aims to study the underuse of Gastro-Tourism in Sri Lanka and propose a digital solution that enables tourists to have an authentic and transparent experience of Sri Lankan cuisine while also expanding the market opportunities for Small and Medium-sized Enterprises (SMEs). The research approach adopted for this study includes a mixed-methods approach, incorporating a literature review, interviews with domain experts, and a survey of both tourists and vendors. The findings underscore the significance of culinary experiences for tourists and highlight the potential of technology in promoting Sri Lankan cuisine. The application prototype developed for this project, "Ambula," leverages features such as Image Classification via Machine Learning, a recommendation module based on dietary and allergy requirements, and a portal for vendors to advertise and sell their food products. The application prototype was designed based on user feedback and evaluated using usability and functional testing. The results of the project demonstrate that the application can enhance gastronomic experiences and contribute to the development of sustainable tourism practices for tourists, vendors, and other stakeholders involved in Sri Lanka's tourism industry.

KEYWORDS

Gastro-Tourism, Tourism, Sri Lanka, Food, Food Allergies, Diet Types, Destination Loyalty, Tourists, Vendors, SMEs, Machine Learning, Recommendation module, E-Commerce

1 Introduction

Sri Lanka is a country that thrives off Tourism, with the Tourism and Travel sector typically coming in as the third largest foreign exchange earner [16]. However, there has been a notable decline in the arrivals. One method to boost the number of tourist arrivals is via Gastro-Tourism. The World Food Travel Association introduced the following as a more inclusive definition for Gastro-Tourism: “Food tourism is the act of traveling for a taste of place in order to get a sense of place” [34]. In recent years, mainstream holiday providers have understood the potential of niche tourism, with gastro-tourism identified as one [13]. Gastronomy allows tourists to access the cultural and historical heritage of destinations and immerse themselves in customs, traditions, recipes, dishes, cooking techniques, utensils, cooking stories, unique ingredients, and history. That is, it makes it possible to approach culture in a more experiential and participatory way [37] Gastro-tourism has emerged as a significant trend in the tourism industry. It has become a strong motivator for international travelers, contributing to the development of tourist destinations. Food has the potential to shape a country's brand identity and is increasingly popular due to food and drink programming in various media platforms. Millennials and Generation Z are seeking more food and drink experiences, with more importance and emphasis placed on ethical, sustainable, and local dining experiences than previous generations [37].

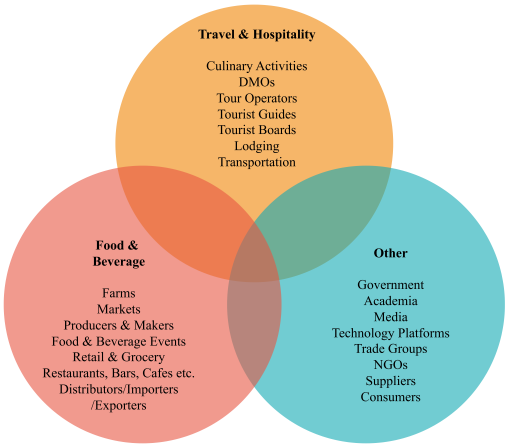


Figure 1 The Food Tourism Industry Cluster [37]

In 2019, the global culinary tourism market was valued at US$ 1,116.7 billion [5]. It is estimated at US$ 865.71 Billion in 2022 and is projected to reach US$ 4,009.77 Billion [32]. Given the size of the market and its growth potential, gastro-tourism is a niche market that deserves the attention of countries that depend greatly on tourism for their GDPs. However, Sri Lanka is not exploiting this sub-sector to its fullest extent [20].

Therefore, the key problem statement of this project is the under-utilisation of the food culture in Sri Lanka as an income generator in the Tourism Industry that will allow tourists to experience authentic Sri Lankan gastronomy with full transparency while widening the market scope for SMEs. On diving into the problem domain, the research identifies four key root causes towards the under-utilization of gastro-tourism which are lack of initiatives, lack of innovation, lack of technology and external factors. The following paragraphs discuss the problems in depth.

1.1 Lack of Initiative on Gastro-Tourism

The lack of initiative in Gastro-Tourism in Sri Lanka has resulted in various issues. Outdated dining experiences and limited menu options fail to accommodate tourists with dietary restrictions or allergies [19][22]. Inadequate standardization of rules and regulations, particularly concerning food safety, creates a perception of risk among tourists [19]. The unstable infrastructure, encompassing economic, financial, and political aspects, further hampers the exploitation of Gastro-Tourism and negatively affects Small and Medium Enterprises (SMEs) that contribute significantly to the tourism industry [30]. The absence of a stable value chain between local SMEs and global markets, caused by coordination failures and miscommunication among government institutions, impacts SMEs and women entrepreneurs who lack financing opportunities [18]. Despite the potential for increasing awareness and appeal through effective marketing and public relations programs, the political landscape in Sri Lanka lacks stability, preventing the formulation of long-term policies for tourism development [10]. The frequent change of government leads to discarded policies and a lack of meaningful collaboration between the public sector, private sector, and the tourism industry [18].

1.2 Lack of Innovation

Sri Lanka ranks 74th out of 117 economies Travel & Tourism Development Index 2021 indicating a need for improvement in the sector [35]. The tourism industry in Sri Lanka has struggled to adapt to changing trends and understand the behavior of modern travelers. The traditional tourism model of undercutting prices has led to financial challenges for many businesses [31]. This has hindered the recovery of the declining tourism sector in Sri Lanka, including the Gastro-Tourism niche. The diminishing generational knowledge of indigenous and traditional food, as well as limited presentation of local products on mainstream platforms, has created a gap in the Gastro-Tourism sector [21]. Inconsistent quality and standards, lax policies and regulations, and concerns over food adulteration deter repeat customers [7][8]. Sri Lankan restaurants often

fail to present traditional food in a modern and innovative manner, both in fine dining establishments and street food offerings.

1.3 Lack of Technology Used

The tourism industry has failed to recognize and adapt to its new audience, including Millennials and Gen Z. To connect with this demographic, building brand awareness through digitalization is crucial. Social media platforms provide an effective means to communicate the unique experiences a destination has to offer, showcasing local culture and emphasizing sustainable travel practices [17].

The government has been slow in embracing technological advancements, particularly in terms of digitalizing information. To leverage the increasing volume of data, the public sector needs to expedite the process of digitization to remain competitive on a global scale [38]. Resistance to technology adoption stems from bureaucratic processes, conformity, and a lack of tech-savviness among stakeholders in the tourism sector [36]. This resistance hinders the long-overdue shift to digital marketing, as noted by Kimarli Fernando, the former Sri Lanka tourism chief [2]. Limited knowledge, training, and resources in developing economies contribute to the delay and reluctance to embrace digitalization [1]. Furthermore, applications pertaining to Food in Sri Lanka consist of third-Party Delivery Services such as Uber Eats and PickMe Food. This proves that the development of the intended solution is imperative.

1.4 External Factors

The tourism industry in Sri Lanka faced multiple challenges in recent years. It started with the Easter Sunday Attack in 2019, which caused a significant drop in tourist arrivals from 2,333,796 in 2018 [24] to 1,913,702 in 2019 [25]. This event raised concerns about the country's security, leading to travel warnings from several nations and economic pressure on businesses. The COVID-19 pandemic in 2020 further worsened the situation, with tourist arrivals declining to 507,311 [25] and a complete halt in travel between April and November 2020. This resulted in a significant increase in unemployment within the industry. In 2022, an economic crisis hit the country, and the tourism industry struggled to recover with arrivals stagnating at 568,258 as of October 2022 [28] fuel had a severe impact on the tourism industry and its supply chain. Sri Lanka's international image was also damaged, leading to negative perceptions among tourists [29].

1.5 Research Questions

There are several important considerations outlined in the problem domain. Considering these concerns, this project explores the answers to the following research questions:

- Are there gaps in the local market for gastro-tourism in Sri Lanka?
- How can gastro-tourism help stakeholders like Small and Medium Businesses generate revenues?
- How can this research introduce a novel digital initiative to promote gastro-tourism in Sri Lanka and what are the key requirements?

With these questions in mind, the literature review delves into existing literature to explore the gaps and opportunities to take this research forward.

2 Literature Review

2.1 Introduction to Gastro-Tourism and Its Importance

Gastro-tourism has emerged as a significant trend in the tourism industry [3]. It has become a strong motivator for international travelers, contributing to the development of tourist destinations. Gastronomy, on par with other attractions, plays a vital role in attracting tourists [14] and has the potential to shape a country's brand identity [20]. It is increasingly popular due to food and drink programming in various media platforms as Millennials and Generation Z prioritise ethical, sustainable, and local dining experiences [34]. The food

tourism industry comprises interconnected sectors, emphasizing the need for collaboration and a cooperative approach to create a successful gastro-tourism strategy and build a strong brand identity [15].

2.2 Sri Lankan Food Culture

In terms of Gastronomy, Sri Lanka is teaming with potential. With various food cultures, offering diverse dishes, packed with flavour and nutritional benefits, Sri Lanka has a lot to offer. The cuisine of Sri Lanka is multi-layered, reflecting the long history of trade activities, royal marriages, and invasions. The influence of Arabic, African, European, and neighbouring Asian countries on the indigenous food culture has enriched the gastronomy of the country. The traditional diet of Sri Lankans may be summed up as "rice and curry", with all ethnicities in the country adapting the dish to reflect cultural, religious, and social norms unique to them. The dish generally consists of rice, served with a fish or meat curry, along with other curries made from vegetables, lentils, or fruits [21]. The generous use of coconut products and a plethora of spices in the cooking gives these preparations their signature Sri Lankan taste. The influence of indigenous and ayurveda (traditional medicine practice) can be seen in the deep understanding of nutrition as well as health benefits of foods, with different food consumed according to the type of person, physiology, degree of activity, and the type of meal. Food is commonly used for therapeutic and treatment purposes, with such knowledge commonplace in communities across the country [24]. A great example of the preservation of this knowledge can be seen in Mahasupavamsaya - The Great Chronicle of Sri Lankan Culinary Art by Dr. Desabandhu T. Publis Silva [23].

2.3 The dietary inclusivity of the Sri Lankan cuisine

Two of the most common dietary restrictions that can hinder travel plans are gluten- free and vegetarian/vegan diets. The Sri Lankan cuisine is special in that most dishes are naturally gluten-free and vegan. The "rice and curry" menu consists of gluten-free rice as the primary carbohydrate, along with tubers, vegetables and fruits cooked using coconut milk, making them vegan. The fish and meat dishes are often cooked and presented separate from other curries so there is little to no cross-contamination. Other commonly found foods such as Hoppers, String hoppers, Appa, Pittu, Idly, and Dosa are gluten-free when made in the traditional manner. There is a risk of gluten contamination with some restaurants or hotels adding wheat flour for ease of producing larger volumes at a lower price. Several gluten-free ancient grains such as Kurakkan (finger millet), Olu haal (water lily rice), Meneri (Proso millet), and Thana haal (Foxtail millet) are also used in traditional cooking, both for sweet and savoury dishes [23].

In the same manner, most traditional sweet and savoury Sri Lankan recipes are vegan, with the use of coconut milk instead of dairy products as commonly used in neighbouring Asian countries [21]. The most common foods that have been reported to cause allergic reactions are known as the "Big Nine": milk, eggs, nuts, fish, crustaceans, shellfish, wheat, soy, and sesame [12]. The abundance of vegetarian and vegan dishes in the cuisine means that milk, egg, fish, shellfish, and crustacean allergens can be avoided by tourists with relative ease. The Tamil food culture in the country is often vegetarian, buy tends to contain more dairy products. When dairy is found in traditional Sri Lankan diet, it is either buffalo curd, or in the form of "Pasgorasa" (five flavours) milk, curd, ghee, cream/butter, and whey. But with 73% of Sri Lankans showing lactose intolerance in varying degrees, it is limited in daily use [31].

The gluten-free (no wheat products) nature of traditional cooking allows for a great culinary experience by the tourist without compromising on flavour. Soy allergens can be found mostly in fusion cooking in Sri Lanka (Chinese fusion being the most common), but not in traditional cooking. The most commonly available nuts in Sri Lanka are cashews and peanuts, mostly eaten as snacks. However, cashews can be found in rice and curry meals, but can be easily removed from the menu. Sesame is mostly used in sweet dishes in Sri Lanka, with the most common being sesame rolls and sesame balls, consumed as tea-time snacks. Sesame oil is rarely used in traditional cooking, making a more common appearance in fusion cooking like soy products. Thus, the Sri Lankan cuisine can be identified as a very allergen free cuisine, opening the doors to promoting Sri Lankan food as inclusive of dietary restrictions without compromising of flavour, nutrition, or variety [21].

Considering the complexity of dietary needs and difficulties in identifying their constituents in prepared food, Tourists have a challenge in front of them to make conscious decisions not only about consuming but enjoying Sri Lankan cuisine to its fullest. Hence, we take these learnings forward to be considered in the design of a potential digital solution.

2.4 SMEs

One of the research questions explores the ability of SMEs to generate revenue through this research initiative. This section explores this aspect and presents the potential of SMEs and the opportunities available to them. SMEs comprise more than 75% of enterprises in Sri Lanka and, account for more than 20% of exports, and 52% of GDP [33]. Of the SMEs functioning in Sri Lanka, about 40% are involved in the food sector and employs an estimated 2 million [9]. The processed/value-added food industry in Sri Lanka has been identified to hold great potential in terms of production, consumption, export, and growth prospects [33]. If an increase in demand for local food and value-added products can be created, SMEs and the local communities will enjoy direct benefits, with an increase in local farming and manufacturing [11]. Despite the potential, there isn't a reliable link in the value chain between regional SMEs and international markets [4]. There is very little published research discussing the operation of food-based SMEs in the tourism industry in Sri Lanka [11]. Restaurants and cafes can be considered as the most common form of food-based SMEs in tourism, with cooking demonstration and food tour SMEs less so [12]. Food-based SMEs that manufacture and sell value-added traditional food, such as traditional sweets, spice blends, ready-to-eat foods, and snacks etc can be considered as the least visible form of SMEs in the tourism sector. The average profit of tourism-related SMEs is considerably higher than the average profit of non-tourism SMEs, showing the benefits of combining the food-based SMEs sector with tourism [8]. The introduction of digital platforms will encourage the entry of food-based SMEs to newer markets, such as tourism and global value chain [15]. Ms. Kimarli Fernando, at a business to business (B2B) interactive session with the Member states of the Conference on Interaction and Confidence-Building Measures in Asia (CICA), stated that "The primary focus at this point is the sector's speedy recovery and SMEs have a crucial role to play in this transition," [6].

2.5 Review on Existing Competitor Products

As outlined in the research questions, this section focuses on exploring the gaps in the gastro-tourism market. At the outset of the research, the researcher outlines a potential digital solution which will be explored and compared against this review. The table below is a gap analysis conducted on existing solutions along with "Ambula", the proposed solution.

Table 1 Gap Analysis

Features	Competitor				
	Withlocals	**Tourplus**	**Drinking with Beers**	**Traveling Spoon**	**Ambula**
Set Diet Type	✓	✓	✗	✓	✓
Set Allergy triggers	✓	✓	✗	✓	✓
Browse for meal information	✗	✗	✗	✗	✓
Food recommendation based on chosen diet type	✓	✓	✗	✓	✓
Food recommendation based on location	✓	✓	✓	✓	✓

Product recommendation based on Diet type and Allergy type	✗	✗	✗	✗	✓
In-App Purchases	✓	✓	✓	✓	✓
Browse for Food products	✗	✗	✗	✗	✓
Save favourite meals, Vendors, Food Souvenirs	✓	✗	✓	✗	✗
Create business listing and advertise	✓	✓	✗	✓	✓
Host Food Tours	✓	✗	✓	✓	✗
Food recognition using Image processing	✗	✗	✗	✗	✓
Available in Sri Lanka	✓	✗	✗	✓	✓

Withlocals is a platform that focuses on personalized and sustainable experiences, offering tours by local guides and customized group activities. They also partner with passionate local hosts to provide authentic food experiences, although this feature is not available in Sri Lanka and does not include food information or the ability to browse/purchase food products Tourplus is a similar platform that offers booking services for tours, transportation, and hotels. It allows merchants to advertise their restaurants and food products, but it is not available in Sri Lanka and does not provide detailed information about individual dishes such as ingredients or allergy triggers.

Drinking with Beers is a service and mobile application centered around beer tours, owned by Amy Beers. It operates in Indianapolis and includes visits to breweries, restaurants, and other beer-related locations while exploring the historical and cultural aspects of beer. The application offers extensive knowledge on beer but is focused solely on that topic. Traveling Spoon is a platform that connects tourists with local hosts in various locations worldwide, offering authentic food experiences such as cooking classes and dining with the host. It provides the opportunity to learn about food history. It is available in Sri Lanka and has generally received positive ratings, although there is limited information regarding the food components used in creating dishes. Overall, these applications have little to no significance in Sri Lanka and do not cater to the local economy or Tourists visiting Sri Lanka. They are limited to the capabilities they offer in the locality where they operate. Hence, there is a need for a localised digital solution which addresses the needs of Tourists and local vendors in Sri Lanka.

With the literature review findings, it was important that there is a need to explore the birth of a novel solution that caters to the local market and Tourists visiting Sri Lanka, to uplift the gastronomical experience and promote a different aspect of the country. Hence, the next section presents the methodology taken to investigate the target audiences for this digital solution initiative.

3 Methodology

An empirical survey was carefully designed and conducted with the following key objectives in mind:

- To validate the problem statement in the local gastro-tourism sector
- To investigate individual dietary preferences and market needs related to gastronomy.
- To investigate user and system requirements for a digital solution catering to gastro-tourism

The research follows a mixed methodology approach to investigate quantitative and qualitative aspects to gather information for the three objectives outlined above. The empirical survey adopted both questionnaire and interview approaches. Two sets of questionnaires were employed to retrieve quantitative data from a

large audience of Tourists and Vendors, respectively. The Tourist survey was available in English and German and was delivered both physically via paper forms and online via Google Forms to have a better reach. The Vendor survey which was solely online and was accessible in English, Sinhala, and Tamil. Overall, 157 Tourists responded to the Tourist Survey and 33 responses came from Vendors for the Vendor Survey. The tourists were selected through probability sampling as they are easily found via online audiences and communities. However, as vendors were a specific audience, they were directly approached and selected using the non-probability sampling method.
Interviews were conducted after selecting domain experts through non-probability sampling as required for this research to obtain expert opinion. Three domain experts were consulted, who are the Former Sri Lanka Tourism Development Authority (SLTDA) Chairperson Ms. Kimarli Fernando, Founder Director of the Responsible Tourism Partnership (RTP) Ms. Charmarie Maelge and Former Personal Manager - Avani Bentota Resort and Hemas Hotel Project Consultant Mr. Tissa Zoysa. The semi-structured nature of the interviews allowe to develop a qualitative analysis perspective. The following section presents a summary of the key findings.

4 Key Findings and Discussion

4.1 Justification of the problem domain

4.1.1 Lack of initiatives in gastronomy

During the interviews, it became evident that Sri Lanka lacks a well-established brand image for its gastronomy. Despite some successful initiatives, there is a significant gap in efforts to promote Sri Lankan cuisine as a whole, both locally and internationally. The country's diverse cuisine, with its numerous health benefits and regional variations, as well as its rich historical and cultural significance, remains largely untapped and underutilized. One of the contributing factors to this issue is the prevalent practice of offering full package tours for tourists, which often restricts them from exploring and experiencing authentic local food.

4.1.2 Lack of visibility of SMEs among Tourists

Regarding Sri Lankan Vendors and SMEs, there have been projects like the Women Agriculture Extension Programme that aim to promote traditional local foods and support women entrepreneurs. However, these vendors still face challenges in connecting with a tourist customer base and accessing the global value chain. To address these challenges, it was suggested by Ms. Maelge that vendors should be made aware of regulations and standards without enforcing strict standardization, as this could be counterproductive. Additionally, the use of sustainable materials can enhance product value while also benefiting the environment.

4.1.3 Lack of marketing efforts in the food category

Ms. Fernando and Mr. Zoysa both agreed that there is a pressing need to improve the marketing of Sri Lankan food products. The presentation of these products needs to meet international standards, especially due to strict food laws that hinder export opportunities. As potential solutions, Ms. Fernando and Ms. Maelge proposed incorporating cooking classes and cookbooks into the offerings. Furthermore, all three experts concurred that a digital marketing system would be instrumental in promoting Sri Lankan food cuisines, given its ability to reach a wide audience of both tourists and vendors.

4.2 Gastro-tourism, tourist preferences, and digital solution

4.2.1 Demographics of Tourist respondents

The Tourist survey provides valuable insights into the demographics, preferences, and experiences of tourists regarding Sri Lankan cuisine. Most survey participants were in the age range of 30-39, indicating a significant interest in exploring the local food culture among this age group. Germany emerged as the country with the highest number of respondents, followed by the UK and Russia, highlighting the international appeal of Sri Lankan cuisine.

4.2.2 Tourist interest in Sri Lankan gastronomy

A notable finding is that majority of participants (80%) were regular visitors to Sri Lanka, showcasing a consistent interest in experiencing the local cuisine. While the spiciness of the local food emerged as a common issue (28%) faced by participants, many tourists found the local cuisine to exceed their expectations (91%). However, some expressed concerns about the limited variety and carb-heavy nature of the food in certain establishments.

4.2.3 Demand for food and related souvenirs among Tourists

The Tourist survey also highlights the significant market potential for food-related souvenirs, with almost half of the participants purchasing such items during their visit. Spices, tea, local alcohol, and cashews were the most popular choices among the tourists.

4.2.4 Demand for a digital solution and individual preferences.

The participants expressed a strong desire for a digital solution to assist them in navigating the local cuisine. Features like ingredient listings, spice levels, and historical information were highly desired. The majority preferred a mobile application as the preferred platform (69%), although a notable percentage of older participants preferred a website (30%).

While a significant portion of participants had not used a similar digital solution before, the survey findings suggest that the proposed solution would be of great interest to the majority (67%). This validates the problem statement as there is a demand from tourists to assimilate with the food culture and push to improve this experience if coming up short.

4.2.5 Vendor demographics and the vendor ecosystem

The survey on participating vendors revealed insights about their business categories, products sold, sales platforms, target customer base, online presence, and preferences for a digital solution. Most vendors were categorized as vendors (44%), followed by restaurant owners (18%) and street food vendors (12%). The vendors offered a diverse range of products from 27 different categories, with traditional rice and curry being the most common (38%). While dedicated retail outlets were the primary sales platform (58%), a considerable number of vendors also maintained an online presence (24%). A significant portion of vendors (74%) targeted the tourist/foreign customer base, with some facing challenges in catering to this market. Uber Eats was the most used online review and delivery platform (38%).

4.2.6 Vendor preference for a digital mobile app solution and related requirements

Vendors expressed strong preferences for a digital solution that enables business registration, analytics viewing, customer ratings, and transaction capabilities. A mobile application was the preferred mode of delivery for the digital solution (82%), with ease of use and reliability being some of the most important requirements. Furthermore, a large majority of vendors (93%) expressed interest in utilising the proposed digital solution. These findings highlight the potential benefits of a tailored digital solution to assist vendors in accessing and serving the tourist/foreign customer base, providing valuable tools for business growth and customer engagement.

Based on these key findings, a digital solution based on mobile technology was designed, developed and tested, as per the requirements brought out by the target audiences. A methodical approach was followed in this process by developing a fully-fledged requirements specification by prioritizing requirements and designing analysis diagrams using UML techniques. Then a business solution and a mobile application was designed with proper consideration given to the architecture, user interface design and choice of technology. Then the solution was tested using black box and white box testing methods to validate the requirements along with evaluations for user acceptance tests by the target audiences. The following section explains the digital solution and the evaluation results.

5 Proposed Digital Solution: ‘Ambula’

“Ambula” is a mobile application that lifts the veil on Sri Lankan food to Tourists while simultaneously allowing Vendors to promote their specialty products. It is a recommendation app that considers parameters

such as diet type, intolerances and location when recommending Sri Lankan food and food souvenirs. This criterion is based on personal information (allergies, preferences) provided by the Tourist during account creation. The application also implements image processing functionalities which allow Tourists to identify food by uploading a picture to the system. This is a key unique feature of the Ambula application compared to rivals which was validated via the competitor analysis conducted. Verified vendors are allowed to create listings for their business and directly advertise and connect with Tourists. This platform aims to contribute to the promotion of Sri Lankan food culture with clarity and accuracy. It is a completely developed product. The solution was developed with the hope that it could have a tri-fold impact on Sri Lanka's Gross Domestic product:

1. Encouraging Tourists to step out of all-inclusive packages and explore the cuisine at local eateries benefiting small business owners.
2. Promoting the concept of food souvenirs to include foods other than Tea and Spices with direct purchasing from Small and Medium Enterprises (SMEs), cutting out middleman corporations such "food cities".
3. Create a demand for Sri Lankan food and food products overseas, allowing SMEs to enter the global value chain.

A minimum viable product of "Ambula" was developed using a variety of tools and technologies ranging from Flutter, MongoDB, Dart, JavaScript, Node, Google Sheets API and more. For the Machine Learning component in relation to the Image Classification, Keras, Opencv2, Python, PyCharm, Jupyter Notebook, TensorFlow and Flask were utilised. The model in the Image Classification API was trained via Transfer learning using the mobilenetv2 model which was originally trained using an ImageNet dataset. This model was later finetuned using a custom dataset and made to detect ten different unique Sri Lankan food. The unique dataset was created specifically for "Ambula" and was done with two sets of images: one to train and one to test. The training set consisted of 55 images of a dish and the training used 25 images of the same dish. Once it receives the base64 string, it uses cv2 library to retrieve the image from the base64 string. Then it uses the image to predict the type of food in the picture and return an output with information pertaining to the recognised dish.

6 Expert and User evaluation

The proposed digital solution, "Ambula," underwent evaluation by domain experts and potential users, including both tourists and vendors to validate the solution and to ensure market acceptance. The application was shared with these individuals and a questionnaire was used to gather valuable feedback on the fulfillment of the digital solution requirements and its usability. The feedback received from the evaluators was predominantly positive, affirming that the digital solution had met the established requirements and was extremely useful. They also provided suggestions for future improvements and enhancements.

One aspect of the application that received praise was its clear-cut look and easy navigation. Users found the interface intuitive and user-friendly, allowing for seamless interaction with the platform. Additionally, the Image Classification model implemented in the solution was well-received, as it facilitated the identification and classification of various food items, aiding both tourists and vendors in their culinary experiences.

Among the suggestions for improvement, users expressed a desire for expanded functionality of the current features and the Image Classification database. The current model gave an accuracy rate of 83% which is in a favourable range, yet expansion would allow for a more comprehensive coverage of food items and further enhance the accuracy of the classification model.

Additionally, users proposed the introduction of new services such as food tours and cooking classes, providing tourists with immersive experiences and vendors with additional avenues for promoting their products. Another area for improvement highlighted in the feedback was the enhancement of vendor analytics. Users suggested incorporating more robust and comprehensive analytics tools that would enable vendors to gain valuable insights into their business performance, customer preferences, and market trends. By leveraging advanced analytics capabilities, vendors would be better equipped to make informed decisions and optimize their operations.

Overall, the evaluators acknowledged its potential to have a significant impact on the tourism and gastronomic dynamics in Sri Lanka. The positive reviews, coupled with the valuable feedback received, provide a strong foundation for further development and refinement of the “Ambula” digital solution, ensuring its continued success in supporting both tourists and vendors in navigating the local cuisine and fostering a rich culinary experience.

7 Conclusion

The research posed three questions to which answers have been derived through the research components, including literature review, empirical survey, analysis of findings, solution design and development and evaluation. The literature review and gap analysis of existing applications have highlighted the absence of dedicated applications in Sri Lanka that focus on promoting local cuisine and assisting vendors in expanding their reach to an international audience. Existing platforms either limit their services to third-party delivery or lack transparency. However, "Ambula" stands out as a unique solution in this landscape. Its mobile-based platform makes it easily accessible and practical for both international tourists and Sri Lankan vendors. By comparing it to other applications in the gap analysis, it becomes evident that "Ambula" offers a diverse range of features that are not available elsewhere. From e-commerce capabilities to serving as an information portal and even providing image classification, "Ambula" offers a comprehensive digital solution that sets it apart. The application ensures transparency when it comes to Sri Lankan cuisine, providing a comprehensive list of ingredients, information on possible allergens, and regional favorites.

While the internet can provide similar information, "Ambula" consolidates everything into one convenient location, allowing for personalized food and product recommendations based on preferred dietary needs and allergy triggers. However, it should be noted that the current availability of the application only in English creates a barrier for Sri Lankan vendors who are interested in using the platform but face difficulties due to language fluency.

In summary, "Ambula" fills a significant gap in the market by providing a unique and comprehensive digital solution for promoting Sri Lankan cuisine and aiding vendors in reaching an international audience. With its user-friendly mobile platform, transparency in culinary information, and personalized recommendations, "Ambula" has the potential to revolutionize the way Sri Lankan cuisine is showcased and experienced globally.

8 Recommendations and Future work

Recommendations and future work arise from two aspects, the research, and the digital solution. From a research perspective, the following areas have been identified:

1. There is potential for extensive research on indigenous and Ayurvedic food communities focusing on health benefits, ingredient acquisition and preparation, along with how it can be promoted to an international audience.

2. There is an unexplored area of research which focuses on Sri Lankan food-based SMEs to identify the requirements needed to enter the global value chain along with understanding the importance of brand image/marketing and packaging.

From the digital solution perspective, the following functionalities are recommended as enhancements to the “Ambula” application:

1. Recommending Restaurants based on Food items, location, and dietary preferences/restrictions.
2. Push notifications for recommendations based on Location.
3. Integrating with Uber, PickMe, and couriers for delivery services.
4. Integrating with a secure payment gateway service provider.
5. Implementing a Complaint and Chat feature so that Tourists can directly communicate Vendors instead of just leaving ratings and reviews.

6. Improving Vendor analytics to provide more complex insights such as individuals product data behaviour.
7. Offering the Vendor portal in Sinhala and Tamil.
8. Offering a subscription-based service model for Vendors that is structured according to their business size and requirements such as number of listings, complexity of analytics, etc.
9. Displaying information such as Income Tax, Terms and Conditions, and return policies for Vendors.
10. Introducing services such as food tours and cooking classes for Tourists to purchase.
11. Bank Account connection for Vendors.

ACKNOWLEDGMENTS

First and foremost, I would like to express my sincerest gratitude to my project supervisor, Ms. Lakshika Paiva, for her invaluable patience and constructive feedback. I thank her immensely for taking the time from her busy schedule and imparting knowledge and guidance which encouraged me to successfully bring this project to a close. My sincere thanks also go to Ms. Kimarli Fernando, Ms. Charmarie Maelge, and Mr. Tissa Zoysa for providing insights regarding the project idea as industry experts. Finally, I thank my parents and sisters for their unwavering support throughout the project.

REFERENCES

[1] Mudalige Uthpala Indeelinie Alahakoon and Shahzadah Nayyar Jehan. 2020. Efficiency of Public Service Delivery—A Post-ICT Deployment Analysis. Economies 8, 4 (December 2020), 97. DOI:https://doi.org/10.3390/economies8040097

[2] Peden Doma Bhutia. 2022. Sri Lanka Tourism Chief Resigns Under Pressure, Then Warns of Industry Cartel. Skift. Retrieved November 17, 2022 from https://skift.com/2022/05/26/sri-lanka-tourism-chief-resigns-under-pressure-then-warns-of-industry-cartel/

[3] Stephen Chaney and Chris Ryan. 2012. Analyzing the evolution of Singapore's World Gourmet Summit: An example of gastronomic tourism. International Journal of Hospitality Management - INT J HOSP MANAG 31, (June 2012). DOI:https://doi.org/10.1016/j.ijhm.2011.04.002

[4] Chen Chen and Sabyasachi Mitra. 2022. Five Ways to Connect Sri Lanka's Small Businesses to Global Value Chains. Retrieved November 17, 2022 from https://blogs.adb.org/blog/five-ways-connect-sri-lanka-s-small-businesses-global-value-chains

[5] Cision US. 2020. Global Culinary Tourism Market (2020 to 2027) - by Activity Type, Age Group and Mode of Booking. Retrieved November 1, 2022 from https://www.prnewswire.com/news-releases/global-culinary-tourism-market-2020-to-2027---by-activity-type-age-group-and-mode-of-booking-301106550.html

[6] Daily News. 2021. Recognizing vital role of SMEs in Sri Lanka's Tourism Sector. Daily News. Retrieved May 8, 2023 from https://www.dailynews.lk/2021/09/02/business/258202/recognizing-vital-role-smes-sri-lanka%E2%80%99s-tourism-sector

[7] N. P. Ravindra Deyshappriya and B. C. H. Maduwanthi. 2020. Impact of Global Value Chains on Performance of Small and Medium-Sized Enterprises in Sri Lanka: Evidence from Sri Lanka. Asian Development Bank. Retrieved November 6, 2022 from https://www.adb.org/publications/impact-global-value-chains-performance-smse-sri-lanka

[8] Ravindra Deyshappriya and A Nawarathna. 2020. TOURISM AND SME DEVELOPMENT: PERFORMANCE OF TOURISM SMES IN COASTAL TOURIST DESTINATIONS IN SOUTHERN SRI LANKA Asian Development Bank Institute. (July 2020). Retrieved from https://www.researchgate.net/publication/355844995_Contribution_of_Traditional_Cooking_Demonstrations_for_Gastronomy_Tourism_Development_in_Sri_Lanka_The_Case_of_Galle_District_Sri_Lanka

[9] Export Development Board Sri Lanka. 2021. The Role of SMEs in Sri Lanka Exports - EDB Sri Lanka. Retrieved November 6, 2022 from https://www.srilankabusiness.com/blog/role-of-smes-sri-lanka.html

[10] Amalini Fernando. 2019. Can Gastrodiplomacy Revive Tourism in Sri Lanka? The Lakshman Kadirgamar Institute. Retrieved May 6, 2023 from https://lki.lk/publication/can-gastrodiplomacy-revive-tourism-in-sri-lanka/

[11] Rishini Fernando, Hamzayini Pirashath, and Suvimali Rathnayake. 2021. Contribution of Traditional Cooking Demonstrations for Gastronomy Tourism Development in Sri Lanka: The Case of Galle District, Sri Lanka. Retrieved from https://www.researchgate.net/publication/355844995_Contribution_of_Traditional_Cooking_Demonstrations_for_Gastronomy_Tourism_Development_in_Sri_Lanka_The_Case_of_Galle_District_Sri_Lanka

[12] FoodAllergy. 2023. Facts and Statistics - FoodAllergy.org. Retrieved May 4, 2023 from https://www.foodallergy.org/resources/facts-and-statistics

[13] GlobalData. 2022. Gastronomy Tourism Market Trend and Analysis of Traveller Types, Key Destinations, Challenges and Opportunities, 2022 Update. Market Research Reports & Consulting | GlobalData UK Ltd. Retrieved May 4, 2023 from https://www.globaldata.com/store/report/gastronomy-tourism-trend-analysis/

[14] Jeou-Shyan Hong and Chen-Tsang (Simon) Tsai. 2010. Government websites for promoting East Asian culinary tourism: A cross-national analysis. Tourism Management 31, 1 (February 2010), 74–85. DOI:https://doi.org/10.1016/j.tourman.2009.01.009

[15] International Labour Organization. 2020. Platformization of the tourism sector in Sri Lanka: Impacts on labour market and decent work opportunities. Retrieved May 8, 2023 from http://www.ilo.org/colombo/whatwedo/publications/WCMS_755007/lang--en/index.htm

[16] U.S. Department of Commerce International Trade Administration. 2021. Sri Lanka - Travel and Tourism. Retrieved October 21, 2022 from https://www.trade.gov/country-commercial-guides/sri-lanka-travel-and-tourism

[17] Dulekha Kasturiratne and Carleen Martin. 2022. Attracting Millennial and Gen Z customers to Sri Lanka. University of Plymouth. Retrieved November 17, 2022 from https://www.plymouth.ac.uk/schools/plymouth-business-school/attracting-millennial-and-gen-z-customers-to-sri-lanka

[18] Sydnella E. Kpundeh. 2017. Sri Lanka tourism - Strategic plan 2017-2020. Retrieved November 5, 2022 from https://documents.worldbank.org/en/publication/documents-reports/documentdetail/333581539112950320/sri-lanka-tourism-strategic-plan-2017-2020

[19] Rohan Laksiri. 2019. Sri Lankan Food in Tourism Experience: Neophobia, Food Quality Perception and Food Experience Linkages. Global Review of Research in Tourism, Hospitality and Leisure Management (January 2019). Retrieved May 7, 2023 from https://www.academia.edu/39639621/Sri_Lankan_Food_in_Tourism_Experience_Neophobia_Food_Quality_Perception_and_Food_Experience_Linkages

[20] Tasha Marikkar. 2020. Why don't Sri Lankans invest in Sri Lankan food? | Daily FT. Retrieved May 3, 2023 from https://www.ft.lk/columns/Why-don-t-Sri-Lankans-invest-in-Sri-Lankan-food/4-703593

[21] Sachithra Mihiranie, Jagath K. Jayasinghe, Chamila V. L. Jayasinghe, and Janitha P. D. Wanasundara. 2020. Indigenous and traditional foods of Sri Lanka | Journal of Ethnic Foods | Full Text. Retrieved September 14, 2022 from https://journalofethnicfoods.biomedcentral.com/articles/10.1186/s42779-020-00075-z

[22] Jagath Munasinghe, Achini De Silva, G. Weerasinghe, Anil Gunaratne, and Harold Corke. 2015. Food safety in Sri Lanka: Problems and solutions. Quality Assurance and Safety of Crops & Foods 7, (January 2015), 37–44. DOI:https://doi.org/10.3920/QAS2014.x007

[23] Dr. Publis Silva. 2018. Mahasupavamsaya - The Great Chronicle of Sri Lankan Culinary Art.

[24] SLTDA. 2018. Annual Statistical Report 2018. Sri Lanka Tourism Development Authority. Retrieved September 14, 2022 from https://www.sltda.gov.lk/storage/common_media/Tourist%20Board%20Annual%20Statistical%20Report%202018_Web784216427.pdf

[25] SLTDA. 2019. Annual Statistical Report 2019. Sri Lanka Tourism Development Authority. Retrieved September 14, 2022 from https://srilanka.travel/SLTDA_documents/ASR%202019.pdf

[26] SLTDA. 2020. Annual Statistical Report 2020. Retrieved from https://www.sltda.gov.lk/storage/common_media/Annual%20Stastistical%20Report%202021%20-Final%2025.4.20223624932970.pdf

[27] SLTDA. 2020. Monthly Tourist Arrivals Reports 2020. Retrieved October 15, 2022 from https://www.sltda.gov.lk/en/monthly-tourist-arrivals-reports-2020

[28] SLTDA. 2022. Monthly Tourist Arrivals Reports 2022. Retrieved November 6, 2022 from https://www.sltda.gov.lk/en/monthly-tourist-arrivals-reports-2022

[29] SLTDA. 2022. The Impact of the Current Sri Lankan Economic Crisis on Tourist Arrivals. Retrieved from https://www.sltda.gov.lk/storage/common_media/Economic%20crisis%20and%20arrivals%202022%20(1)%20final765641986.pdf

[30] Small Enterprises Development Division. 2015. National Policy Framework for SME Development.

[31] Christian Løvold Storhaug, Svein Kjetil Fosse, and Lars T. Fadnes. 2017. Country, regional, and global estimates for lactose malabsorption in adults: a systematic review and meta-analysis. Lancet Gastroenterol Hepatol 2, 10 (October 2017), 738–746. DOI:https://doi.org/10.1016/S2468-1253(17)30154-1

[32] Transparency Market Research. 2022. Travel and Tourism Market. Retrieved May 4, 2023 from https://www.transparencymarketresearch.com/travel-tourism-market.html

[33] United Nations. 2020. The Impact of COVID-19 to MSME sector in Sri Lanka Sustainable Development Knowledge Platform. Retrieved May 8, 2023 from https://sustainabledevelopment.un.org/index.php?page=view&type=13&nr=3300&menu=1634

[34] UNWTO. 2012. Affiliate Members Global Report, Volume 4 - Global Report on Food Tourism. Default Book Series. Retrieved May 4, 2023 from https://www.e-unwto.org/doi/book/10.18111/9789284414819

[35] Lauren Uppink and Maksim Soshkin. 2022. Travel & Tourism Development Index 2021: Rebuilding for a Sustainable and Resilient Future. World Economic Forum. Retrieved November 17, 2022 from https://www.weforum.org/reports/travel-and-tourism-development-index-2021/

[36] Sivarajah Uthayasankar and Vishanth Weerakkody. 2018. Barriers to digital diplomacy: why are governments slow in adopting technology? British Politics and Policy at LSE. Retrieved November 6, 2022 from https://blogs.lse.ac.uk/politicsandpolicy/barriers-to-digital-diplomacy/

[37] Erik Wolf, Matthew Stone, Roberta Garibaldi, Steven Migacz, and Nate Stein. 2020. 2020 State of the Food Travel Industry Report. Retrieved from https://worldfoodtravel.org/news-the-economic-impact-of-food-tourism/

[38] World Economic Forum. 2022. Public sector risks losing trust as digital transformation lags. World Economic Forum. Retrieved November 6, 2022 from https://www.weforum.org/agenda/2022/05/the-public-sector-must-accelerate-digital-transformation-or-risk-losing-sovereignty-and-trust/

Influence of Supervisor Behavior & Social Support on Employee Wellbeing in the Apparel Industry of Sri Lanka *

Osandi Udalamatta†

Business Management, Informatics Institute of Technology, Sri Lanka, Osandi.20191166@iit.ac.lk

Gayashini Ratnayake

Informatics Institute of Technology, Sri Lanka, gayashini.r@iit.ac.lk†

ABSTRACT

Employees in the apparel industry receive less support from coworkers and managers (Vallina and Alegre, 2018 and Liaw and Chuang, 2020). Due to the poor incentive for work, rising employee absenteeism rates, rising lead times, and low productivity of the task, these issues arise (Tselebis et al., 2021). Employee well-being suffers as a result of these circumstances since they work in an unpleasant atmosphere. On the other hand, supervisor behavior, social support, and employee welfare all help to lower employee turnover (Orth, 2021). A better work environment with employee wellness, social support, and supervisor behavior is required to lessen these implications because it can identify an employee's job level and their happiness as a cycle (Aerts, 2019). Numerous studies have examined the impact of leadership style or social support on employee happiness in addition to the practical research challenge. This paper is to identify the influence of supervisor behavior & social support on employee wellbeing in the apparel industry of Sri Lanka. For employee wellness in the apparel industry, numerous studies have pointed to the importance of supervisor behavior and social support. But most of them haven't explored this subject in the context of Sri Lanka. Deductive research approach using quantitative methods. Data was collected using a validated questionnaire and analyzed using SPSS version 23. Respondents are 207 operational level workers from four leading apparel firms. The findings revealed that there is a positive and significant relationship between supervisor behavior and employee wellbeing supporting previous empirical findings by Zoha et al.,(2021); Fathima and Azam, (2020); Carlo et al.,(2018); Ajantha and Priyadharshani, (2019); Karunanayake et al.(2019). Social support mediates the relationship between supervisor behavior and employee wellbeing.

Keywords: Supervisor Behavior, Social Support, Employee Wellbeing, Apparel sector

1.Introduction

Supervisor relationships and social support contribute to making employees happy at the workplace (Mansour and Tremblay, 2019). It also includes the behavior including rewards and punishments of the supervisors, interaction with the supervisors and the peers, and group/team bonds are some factors that help employees' well-being (Vallina and Alegre, 2018 and Liaw and Chuang, 2020). Employee's good interpersonal relationship with leaders/ supervisors and social is helps to contribute to employees' happiness in the workplace (Tselebis et al., 2021) and finally create employees' well-being. Also, it helps to make positive outcomes for the employees including satisfaction, and efficiency (Sender and Fleck, 2018, Schaufeli, 2021). In the apparel sector, operational workers are get supervised by a senior employee. Employees are spending one-third of their time in the workplace. Therefore, supervisor behavior and social support is helps to determine how much employees get happy with the work environment and how it contributes to employee wellbeing (McNeely et al, 2019, Terk et al. 2021). Supervisor behavior can explain as activities which have been taken by the supervisor to influence activities on workers to get done the organizational objectives (Aerts, 2019). According to that supervisor, behavior helps to determine the level of employee happiness and employee happiness helps to make an employee's well-being (Hettiarachchi, 2022, Terk et al. 2021).

Employees cannot work independently. To get work success and change necessary for organizational requirements, employees need social support (Herscovitch and Meyar, 2018). Social support can explain as actions from others which help to conduct work. Social support can be the flow of emotional concern or psychical support of surroundings (Chou, 2017,Jaros 2019).

1.1 Research Gap

The apparel industry required continues to change and this helps to make the organization sustain and make the well-being of the organization and individuals by increasing the capability of workers' time management, enhancing performance and uplifting energy (Jaros, 2019). Apparel sector employees suffered from stress for a long period (Jaskyte, 2020). This arises with the increase in job demand and the need for the high requirement for work (Kotter and Cohen, 2016). This becomes the main issue for increasing employee health problems. Employees look forward to getting supervisory support and social support to carry out the work smoothly and efficiently. (Kim, & Wang, 2021; Liu, Wang & Sun, 2021). The apparel sector employees receive less social support and supervisor support (Liu and Ipe, 2017). Thiscreates problems with the low incentive for work, increase labour turnover, increase employee rate of absenteeism, increase lead time and low productivity of the work (Orth, 2021). Those situations make less employee well-being because employees work in an unhappy work environment (Mathushan and Kengatharan, 2022, Orth 2021). Also on the other hand employee wellbeing, social support and supervisor behavior reduce employee turnover. It can identify employees' work level and their happiness becomes a cycle and to reduce those implications its need to have better work environment with employee wellbeing, social support and supervisor behavior (Mathushan and Kengatharan, 2022, Orth 2021). Apart from the practical research problem, in literature aspects, many studies (Mathushan and Kengatharan, 2022, Munim et al. 2022, Law et al. 2020, Mounika et al. 2022) have analyzed the effect of leadership style or social support on employee happiness. Many studies have identified need of supervisor behavior and social support for make wellbeing in apparel industry. But most of them are un-tapped this topic on Sri Lankan context.

1.2 Research Questions

The following research questions were identified.
RQ1: What is the influence of supervisor behavior on social support in the Apparel sector?
RQ2: What is the impact of supervisor behavior on employee well-being in the Apparel Sector?
RQ3: What is the influence of social support on employee well-being in the Apparel Sector?
RQ4: What is the mediating effect of social support on the relationship between supervisor behavior and employee wellbeing in the Apparel Sector?

1.3 Research Objectives

Based on the Research questions, the following research objectives were identified.
RO1: To identify the influence of supervisor behavior on social support in the Apparel sector.
RO2: To identify the impact of supervisor behavior on employee well-being in the Apparel Sector
RO3: To identify influence of social support on employee well-being in the Apparel Sector
RO4: To analyze the mediating effect of social support on the relationship between supervisor behavior and employee wellbeing in the Apparel Sector

1.4 Scope of the Study

The scope of this study is limited to supervisor behavior investigated through Apparel industry of Sri Lanka. The export of garments is one of the most significant economic areas in Sri Lanka, and it has experienced rapid expansion over the past four decades and also, the sector's value increased to $5.42 billion in 2021, with a 22.93 percent year-over-year growth in exports (Perera & Fernando, 2022). This study is focused on the context of the apparel industry in Sri Lanka, providing insights specific to this sector and the impact of the factors on employees working in this industry.

2. Literature Review

To carry out the basic factors of the research, this chapter of the literature review will cover significant concepts employing definitions, theories, and models*!*

2.1 Theoretical Background and Models

2.1.1 Social exchange theory

The social exchange theory, which holds that people interact with others in hopes of gaining something in return, forms the theoretical foundation of this literature review (Macleod, 2019). Social interaction takes place in the workplace between staff members and their managers, coworkers, and the company (Emil, 2021). As compensation for their contributions to the organization, employees anticipate receiving assistance, resources, acknowledgement, and other advantages from these sources. The theory of social exchange argues that the consequence of an exchange process is observed to be social behavior. The purpose of the trade is to maximize benefits while simultaneously minimizing costs. According to this theory, people try to strike a balance between the positive and negative effects of the relationships they maintain. If the potential risks outweigh the potential benefits, they will either sever or leave the link (Cherry, 2022).

2.1.2 Job Demands-Resources (JD-R) model.

Arnold Bakker and Evangelia Demerouti created the 2006 Job Demands-Resources (JD-R) Model. The model states that job demands and resources/positives increase stress and burnout (Jane, 2020). However, a work's many positives can offset its many obligations (White, 2022). In Sri Lanka's apparel company, supervisor behavior and social support are important sources of assistance that influence employee wellbeing (Ş Koroglu, O Ozmen, 2022). The Job Demands-Resources (JD-R) model states that while job demands like workload and stress can have a detrimental influence on employees' wellbeing, job resources like social support and supervisory behavior can have a beneficial effect (Bakker & Demerouti, 2014). To preserve employee wellbeing and performance, the JD-R model emphasizes the significance of striking a balance between job demands and resources (Arnold, 2018).

2.1.3 Conservation of Resources (COR) Theory

The Conservation of Resources (COR) Theory is a stress theory that clarifies why humans are driven to both protect and expand their existing resource networks. In 1989, Dr. Stevan E. Hobfoll contributed to the literature on stress by proposing this idea. According to COR theory, people work hard to preserve and expand their resources in order to maintain their wellbeing. Social support, supervisor behavior, job control, and other employment resources are examples of resources at work. According to the COR hypothesis, those who suffer from resource depletion or loss, such as a lack of social support or unfavorable supervisor conduct, may experience work stress, burnout, and diminished wellbeing (Demerouti, 2020).

2.1.4 The Job Characteristics Model (JCM)

According to the Job features Model (JCM) (Hackman & Oldham, 1976), job features like skill variety, task identity, autonomy, and feedback can have an impact on employee motivation, satisfaction, and wellbeing. The JCM places a strong emphasis on the value of job design and job enrichment to improve employee performance and welfare (Towler, 2021). The JD-R model, the COR theory, the JCM, and social exchange theory serve as the foundation for this literature review's theoretical background. These theories offer a framework for comprehending how social support and supervisor behavior in Sri Lanka's garment industry might improve employee wellness. According to the literature, fostering a supportive work environment that strikes a balance between resource requirements and job expectations as well as offering opportunities for workplace enrichment can improve employee performance and wellbeing.

3. Methodology

3.1 Conceptual Framework

Supervisor behavior is the independent variables, while social support is the mediator variable. Employee wellbeing is the dependent variable.

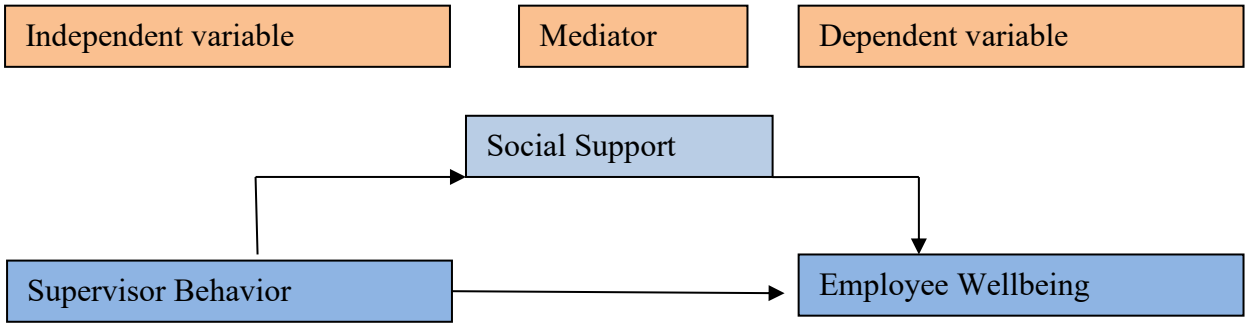


Figure 1: Conceptual Framework (Source: Developed by author, 2023)

3.1.1 Supervisor Behavior

Supervisor behavior can be defined as dynamic and creative roles act by the supervisors/ leaders in the organization for improve organization/ employee performance (Michael, 2021). The behaviors and interactions of supervisors with their subordinates at work are referred to as supervisor behavior. A supervisor's actions can take many different forms, such as giving instructions and feedback, praising workers for their accomplishments, encouraging teamwork and collaboration, and fostering a happy work atmosphere (Michael, 2021). Effective supervisor conduct is essential for employee well-being and performance since it can affect workers' psychological health, motivation, and job satisfaction (Guohao, 2020).

Different behaviors are displayed by the supervisor. In certain circumstances, managers give bonuses to workers who produce quality work and discipline those who violate corporate policies (Melwin, 2016). The supervisor oversees flow, style analysis, line arrangement, machine selection, control, operator training, target production, and product quality (Jayarathne, 2020). Most workers desire a kind, supportive boss, which improves morale (Robertson, 2019). Empirical studies have shown that supervisor behavior can indeed have a positive impact on employee well-being. Social Exchange Theory suggests that employees perceive their relationship with supervisors as an exchange of resources and support. When supervisors provide positive support and resources, employees reciprocate with increased commitment and well-being (Cherry, 2022).

H1: There is a significant positive impact of supervisor behavior on employee well being

3.1.2 Social Support

The resources, knowledge, and emotional support that people receive from their social networks, including coworkers, friends, and family members, are referred to as social support (Cooper and Marshell, 2019). Social support can provide people with a sense of community, connection, and approval, which can improve their well-being and lower stress (Swarnalatha, 2020). Coworker, supervisor, and organizational support in the workplace, such as availability of employee assistance programs or wellness programs, are examples of social support (Fukada, 2019). Supporting others at work involves teaching, encouraging, supporting, and exchanging experiences for other employees' new challenges, which helps to identify solutions. Additionally, it can be supervisor assistance, which includes providing better direction, exhibiting knowledge, abilities, and attitude, and helping staff advance their skills through training programs. In addition, it might be organizational assistance (Mathushan and Navaneethakrishman, 2022). Empirical studies have indeed demonstrated that supervisor behavior can have a positive impact on social support in the workplace. Social Exchange Theory, employees perceive their relationship with their supervisors as an exchange of social support and resources. When supervisors exhibit supportive behaviors, employees are more likely to reciprocate with positive attitudes, including increased social support for their colleagues (Cherry, 2022). Organizational Support Theory emphasizes the importance of perceived organizational support (POS) in influencing employees' attitudes and behaviors. Supportive supervisor behavior contributes to higher POS, which, in turn, encourages employees to provide social support to their colleagues and the organization.

H2: There is a significant positive influence between supervisor behavior and social support

3.1.3 Employee Wellbeing

The status of an employee's physical, emotional, and psychological health at work is referred to as their wellbeing (Senevirathne, 2020). The term "employee wellbeing" refers to a broad variety of aspects, such as physical and mental health, job satisfaction, stress levels, burnout, and job satisfaction. The performance of employees depends on their sense of wellbeing since it affects their commitment to the organization and their willingness to work hard (George, 2019). An apparel company's employee well-being has six categories. Sustainable workload, individual control over job obligations, support from superiors and core employees, constructive communication and connection, and clear organizational change management are among them (Mathushan and Navaneethakrishman, 2022). Employee well-being has become a priority in Sri Lanka's garment industry. Employee well-being affects physical and mental health, absence, and working conditions (Johnson, 2020). Empirical studies have provided evidence of the positive impact of social support on employee well-being. Conservation of Resources (COR) Theory suggests that individuals strive to acquire and protect their resources. Social support functions as a resource that helps employees cope with job demands and stressors, leading to improved well-being and work-related outcomes (Schwarzer, 2020). Social Support Theory posits that social support, whether from supervisors, colleagues, or friends, serves as a buffer against stress and adversity. When individuals feel supported, they experience reduced psychological strain, leading to enhanced well-being.

H3: There is a significant positive influence between social support on employee well being

Empirical findings reveal that social support acts as a mediator between supervisor behavior and employee well-being. In other words, the impact of supportive supervisor behavior on employee well-being is partially explained by the level of social support perceived by employees. Empirical studies demonstrate that the relationship between supervisor behavior and employee well-being weakens when controlling for the impact of perceived social support. This weakening of the direct relationship between supervisor behavior and well-being suggests that social support plays a significant role in explaining the link between these two variables.

H4: Social support mediates the supervisor behavior to employee wellbeing relationship

4.1 Research Methodology

4.1.1 Research Philosophy & Research Approach

A positivistic research philosophy drives this research study as it seeks to empirically observe, measure and test empirical data (Saunders et al., 2016). The study seeks to test relationships between main variables and employs a deductive approach. A quantitative research method is employed using SPSS version 23 for data analysis.

4.1.2 Research Instrument

To test the hypotheses, a cross sectional survey based on a 5-point Likert Scale ranging from Strongly Agree to Strongly Disagree was used. The Likert scale offered theoretically equal distances between the responses and was considered suitable for the research study. Supervisor Behavior was operationalized using ten indicators from the ordinal scale ranging from 1 to 5 proposed by (Guohao, 2020). Social Support was operationalized using ten indicators from the ordinal scale ranging from 1 to 5 by (Mathushan and Navaneethakrishman, 2022). Employee wellbeing was operationalized using ten indicators from the ordinal scale ranging from 1 to 5 proposed by (Johnson, 2020).

4.1.3 Sampling

The population of the research is 580 apparel sector operational level workers of MAS Active Trading (Pvt) Ltd, MAS Intimates (Pvt) Ltd, Hirdramani International Exports (Pvt) Ltd and Omega Line LTD. The sample of the study was selected based on convenience sampling method because the way of gathering data was easily accessible. Convenience sampling is nonprobability sampling method (Mathushan and Navaneethakrishman, 2022). This method is useful for select sample participants with interest and connivance of the researcher scope.

4.1.4 Data Collection

A cross sectional survey was used to collect data from the target sample Among the target 300 sample 207 usable responses were collected. The response rate is 69%.

Table 1: Summary of Sample selection (Source: Developed by author, 2023)

Company	Sample Number of Operational Level Workers	Number responded	Response Rate
MAS Active Trading (Pvt) Ltd	80	51	63.75%
MAS Intimates (Pvt) Ltd	120	64	53.33%
Hirdramani International Exports (Pvt) Ltd	70	63	90%
Omega Line LTD	30	29	96.66%
Total	300	207	69%

4 Data Analysis and Representation

4.1 Normality

The test of Kolmogorov- Smirnov test and Shapiro Wilk test is used to determine the normality of the data. The rule of thumb to identify normality of the data is the p value greater than 0.05 (Mishra, 2019).

4.2 Kolmogorov- Smirnov and Shapiro-Wilk test

The test of Kolmogorov- Smirnov test and Shapiro Wilk test is used to determine the normality of the data. The rule of thumb to identify normality of the data is the p value greater than 0.05 (Mishra, 2019).

Table 2: Normality test (Source: Developed by author, 2023)

Tests of Normality						
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
SB	.145	207	.080	.879	207	.078
SS	.142	207	.074	.900	207	.064
EW	.157	207	.054	.872	207	.052
a. Lilliefors Significance Correction						

4.3 Reliability and Validity for Main Instrument

Validity and Reliability use for determine the data accuracy and the consistency. According to that the current survey is conducted use of 300 respondents and check for the measurements to determine whether researcher

can use the questions for further advance analysis. The chapter analysis and present the findings use of SPSS statistical Software. The reliability of the research instrument is assessed analyzing the Cronbach Alpha value. According to the SPSS results in a reliability test, Cronbach Alpha value is 0.949 and exceeds the threshold value of 0.7 (Sekaran, and Bougie, 2016). Pilot study is carried out using 50 respondents and Cronbach alpha value is 0.949.

4.3.1 Reliability- Main Analysis

Reliability of the research instrument is assessed analyzing the Cronbach Alpha value. According to the SPSS results in a reliability test, Cronbach Alpha value is 0.949 and exceeds the threshold value of 0.7 (Sekaran, and Bougie, 2016). All the main variables exceeded the Cronbach Alpha value of 0.7 and can be considered reliable

Table 3: Supervisor Behavior – Reliability Statistics (Source: Developed by author, 2023)

Variable Name	Cronbach Alpha Value	Number of items
Supervisor Behavior	0.972	10
Social Support	0.9076	8
Employee Wellbeing	0.9536	12
Whole Instrument	0.949	30

4.3.2 Validity- Sample Adequacy

Table 4: Validity 1 (Source: Developed by author, 2023)

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.743
Bartlett's Test of Sphericity	Approx. Chi-Square	456.253
	Df	3
	Sig.	.000

Kaiser-Meyer-Olkin measure is used to determine whether the sample is adequate, and it was 0.743 and exceeding the threshold value of 0.5 (Budaev, 2010; Goni et al., 2020) hence the sample is adequate.

4.3.3 Discriminant Validity

Wilks Lamda test used to determine the Discriminant validity. The Eigen values are less than 1 and hence discriminant validity exists. The p value is less than 0.05 and hence there is a significant relationship between independent and dependent variable. This means that there is discriminant validity between these variables.

Test of Equality of Group mean					
	Wilks Lambda	F	df1	df2	sig
SB	.711	14.251	4	207	.000
SS	.689	35.452	4	207	.000

Eigenvalues				
Function	Eigenvalues	Varianve	Cumilative	Canonical Correlation
1	.585	100.0	100.00	.602

a. First canonical discriminant functions were used in the analysis

4.4 Regression Analysis

Table 5: Model Summary (Source: Developed by author, 2023)

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.868[a]	.753	.750	.25122
a. Predictors: (Constant), SB, SS				

Regression analysis explained the impact of independent variable on dependent variable. The R Square value shows that 75.3% variance in the dependent variable Employee Wellbeing can be explained by Supervisor Behavior and Social Support.

Table 6: ANOVA (Source: Developed by author, 2023)

ANOVA[b]						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	39.182	2	19.591	310.409	.000[a]
	Residual	12.875	204	.063		
	Total	52.057	206			
a. Predictors: (Constant), SB, SS						
b. Dependent Variable: EW						

In the ANOVA table, the p value is less than 0.05 and hence there is a positive and significant relationship between dependent and independent variables.

Table 7: Coefficient (Source: Developed by author, 2023)

Coefficients[a]						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.813	.176		4.609	.000
	SS	.288	.059	.258	4.872	.000
	SB	.537	.043	.657	12.404	.000
a. Dependent Variable: EW						

Supervisor behavior and social support contribute to determine employee wellbeing with beta values 0.258 and 0.657. Of these, Supervisor Behavior makes the largest contribution to employee wellbeing (Beta value of 0.657).

4.5 Hypothesis Testing

Based on the deductive approach, the Hypothesis was developed to evaluate the link between independent, mediator and dependent variables.

H1: There is a significant positive impact of supervisor behavior on employee well being

The correlation coefficient between Supervisor behavior and Employee wellbeing is 0.851, indicating a strong positive relationship. The regression analysis resulted in a p-value of 0.000, which means the relationship is highly significant. The hypothesis is accepted because of the high correlation and statistical significance, suggesting that Supervisor behavior has a significant positive impact on Employee wellbeing. The hypothesis is accepted.

H2: There is a significant positive influence between supervisor behavior and social support

The correlation coefficient between Supervisor behavior and social support is 0.851, indicating a strong positive relationship. The regression analysis resulted in a p-value of 0.000, which means the relationship is highly significant. The hypothesis is accepted because of the high correlation and statistical significance, suggesting that Supervisor behavior is significantly related to the level of social support employees receive. The hypothesis is accepted.

H3: There is a significant positive influence between social support on employee well being

The correlation coefficient between social support and Employee wellbeing is 0.752, indicating a strong positive relationship. The regression analysis resulted in a p-value of 0.000, which means the relationship is highly significant. The hypothesis is accepted because of the high correlation and statistical significance, indicating that social support has a significant positive impact on Employee wellbeing. The hypothesis is accepted.

H4: Social support mediates the supervisor behavior to employee wellbeing relationship

. The p-value for this mediation analysis is 0.000, which indicates that the mediation effect is highly significant. The hypothesis of mediation is accepted, suggesting that the positive impact of Supervisor behavior on Employee wellbeing is partly explained by the level of social support employees receive.

In summary, it demonstrates that Supervisor behavior and social support have strong positive relationships with Employee wellbeing. Additionally, it highlights the mediating role of social support, indicating that the positive influence of Supervisor behavior on Employee wellbeing is partially explained by the level of social support provided to employees. This hypothesis is accepted.

4.6 Mediator Analysis

Table 8: Model 1 (Source: Developed by author, 2023)

Coefficients[a]						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.420	.132		10.762	.000
	SB	.696	.030	.851	23.183	.000
a. Dependent Variable: EW						

Table 9: Model 2 (Source: Developed by author, 2023)

Coefficients[a]						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.813	.176		4.609	.000
	SS	.288	.059	.258	4.872	.000
	SB	.537	.043	.657	12.404	.000
a. Dependent Variable: EW						

Table 10: Sobel test (Source: Developed by author, 2023)

Input:			Test statistic:	Std. Error:	*p*-value:
a	0.696	Sobel test:	4.77676825	0.0419631	0.00000178
b	0.288	Aroian test:	4.77252463	0.04200041	0.00000182
s_a	.030	Goodman test:	4.78102321	0.04192575	0.00000174
s_b	.059	Reset all	Calculate		

The above results indicate that p values are significant and social support play significant mediator effect. The point effect can be calculated by multiplying the unstandardized beta coefficient values (0.696*0.288). therefore 0.20 is the estimate of the indirect effect between supervisor behavior and employee wellbeing through social support at the p value of 0.000 as shown by the SOBELS test.

5 Conclusions

5.1 Main Conclusions

Supervisor behavior has a positive impact on employee wellbeing. This indicates the mean value is higher than the established mean hence a positive relationship was shown. These supported empirical findings by Zoha et al.,2021; Fathima and Azam, 2020; Carlo et al.,2018; Ajantha and Priyadharshani, 2019; Karunanayake et al.2019. Regression analysis revealed that supervisor behavior and social support change employee wellbeing by 75%. Also, from the regression analysis it has founded that supervisor behavior significantly impact on employee wellbeing with beta value 0.696. The findings supported by the empirical findings Jaiswal, (2017), Michael, (2021), Jayarathne, (2020). Supervisor behavior has a positive impact on

social support. Mean value is greater than the established mean and hence a positive relationship was demonstrated. These supports empirical findings by Priyanga and Thenuwara, 2019; Ahlstrom et al., 2022; Yasir et al., 2020; Manu and Sinosh, 2020. Social support has a strong positive impact on employee wellbeing. These empirical findings by Jaiswal, (2017), Michael, (2021), Jayarathne, (2020) support these findings. Also, it has been found that social support is very much important to determine employee wellbeing with mean value 4.51. Social support mediates the relationship between supervisor behavior and employee wellbeing. This supports empirical findings by Senevirathne (2020), Fukada, (2019), Mathushan and Navaneethakrishman, (2022), Vahtera, (2019).

Supervisor Behavior makes the largest contribution to the dependent variable, employee wellbeing based on the beta value of 0.657 (Table 21). The correlation is moderate, and this supports the regression findings as well (Refer table 17). From correlation analysis of supervisor behavior and social support, it has identified that there is positive significant relationship among those two variables with Pearson value 0.753 and p value .000. The study findings are supported by Kularathne and Senevirathne (2020), Khan et al (2020).

Based on the multiple regression analysis conducted, the regression equation is below.

Employee Wellbeing = 0.813 + 0.288 (Mean of SS) + 0.537 (Mean of SB)

REFERENCES

Aerts, C. (2019). The relationship between Work-Family Conflict and Burnout and the moderating role of Age and Gender: A secundaire data analysis.

Arnold, B. (2018) The Job Demands‐Resources model: state of the art. Journal of Managerial Psychology, 12(4), pp. 334 - 339.

Brown, H. (2021) Behavioral Management Theory. Journal of management studies , 12(4), pp. 445 - 449.

Buchanan, B., (2017) Building Organizational Commitment, the Socialization of Mangers in Work Organizations, Administrative Science Quarterly, 19(1): 1974, 533-546.

Cherry, K. (2022) verywellmind. Available at: https://www.verywellmind.com/what-is-social-exchange-theory-2795882 [Accessed 22 04 2023].

Cliffs (2022) 'Occupational Stressors and Resilience in Critical Occupations: The Role of Personality', in Perrewe, P.L. and Ganster, D.C(eds.) Research in Occupational Stress and Well Being. Bingley: Emerald

Demerouti, E. (2020) Conservation of Resources (COR) Theory. Journal of Managerial Psychology, 21(2), pp. 124 - 132.

Drew, C. (2023) helpfulprofessor. Available at: https://helpfulprofessor.com/social-exchange-theory/ [Accessed 24 04 2023].

Dissertation, Colorado State University, Dissertation Abstracts International, 6(3), 26-35.

Emil, H. (2021) Social exchange theory. Journal of management studies , 5(2), pp. 335 - 339.

Fukado (2019). Employee engagement: a driver of organizational effectiveness. European Journal of Business and Management, 3(8), 53-59.

George, J. (2019) The wider context, costs, and benefits of work engagement.. European, 20(1), p. Journal of Work and Organizational Psychology.

Jane, H. (2020) JDR Model. Journal of management studies , 2(1), pp. 123 - 128.

Johnson, H. (2020) Psychological conditions of personal engagement and disengagement at work. Academy of management journal, 2(1), pp. 223 - 234.

Jaiswal (2017). The association of meaningfulness, well‐being, and engagement with absenteeism: a moderated mediation model. Human resource management, 52(3), 441-456.

Jaiswel (2017) Impact of Employee wellbeing on Employee Job performance –Operational employees clothing sector, Research Gate

Jaros, S. (2019). Commitment to Organizational Change: A Critical Review. Journal of Change Management, 10(1), 79-108.

Jaskyte, K. (2020). Assessing Changes in Employees' perceptions of Leadership Behavior, job Design, and organizational Arrangements and their job satisfaction and commitment. Administration in Social Work, 27(4),25-39.

Khan (2020) 'Reciprocal Effects of Work Stressors and Counterproductive Work Behavior: A Five-Wave Longitudinal Study', Journal Of Applied Psychology, 98(3), pp. 529-539

Kotter, J., & Cohen, D. (2016). The Heart of Change. Boston,MA: Harvard Business School Press

Kim, J., & Wang, M. (2021). Supervisor support and employee career adaptability: A moderated mediation model. Journal of Vocational Behavior, 128, 103495. https://doi.org/10.1016/j.jvb.2021.103495

Kularathna and Senaratha (2020) factors Causing stress among the employees in apparel factories in Sri Lanka, Research Gate

Kumar (2011) Step by step for research beginners, scientific research publish

Lucas, S. (2021) aihr. Available at: https://www.aihr.com/job-characteristics-model/ [Accessed 23 04 2023].

Lamm and Gordon(2016). Personality traits and resilience as predicators of job stress and burnout among call centre employees. Unpublished master's thesis, University of the Free State, Bloemfontein

Lamm, E., & Gordon, J. R. (2016). Empowerment, Predisposition to Resist Change, and Support for Organizational Change. Journal of Leadership & Organizational Studies, 17(4), 426-437

Mathushan, P. (2022) The Moderating Effect of Employee Well-Being between Employee Engagement and Firm Performance: Evidence from the Apparel Industry in Sri Lanka. Kelaniya Journal of Management, 11(1), pp. 129 - 136.

Melwin, M. (2016) EMPLOYEE ENGAGEMENT - AN EMPIRICAL STUDY ON IMPLICATIONS FOR PSYCHOLOGICAL WELL. International Journal of Management , 7(5), p. 183–187.

Mathushin and Nevaneethakrishnan (2022) Moderating effect of employee wellbeing between employee engagement and firm performance, research gate, https://www.researchgate.net/publication/361339172

Meyer, J. P., & Herscovitch, L. (2018). Commitment in the workplace Toward a general model. Human Resource Management Review, 11,299-326.

Vallina and Alegre (2018). Employee commitment and support for an organizational change: test of the three-component model in tow culture. Journal of Occupational and Organizational Psychology, 80(2), 185-211

iCIIT 2023

THE 4TH INTERNATIONAL CONFERENCE ON
INNOVATIONS IN INFO-BUSINESS & TECHNOLOGY

EXTENDED ABSTRACTS

DATA SCIENCE & AI



Voice-bot for book an appointment

Nawanjana Madhushankha Pathirage
SLT digital lab, Sri Lanka Telecom PLC
nawanjanamadhushankha@gmail.com

Dinoj Waruna Wickramasinghe Gunasekara
SLT digital lab, Sri Lanka Telecom PLC
dinojwaruna2@gmail.com

Anil Pradeep Kumara
SLT digital lab, Sri Lanka Telecom PLC
anil@slt.com.lk

Navindu Madanayake
SLT digital lab, Sri Lanka Telecom PLC
navindudananga123@gmail.com

Chathura Bandara
SLT digital lab, Sri Lanka Telecom PLC
chathuramahasen@gmail.com

ABSTRACT

This study examines the creation of a voice bot system that uses RASA NLU to schedule appointments. There is a rising need for effective and user-friendly appointment scheduling systems in the current digital era. Traditional methods can come with difficulties, such as lengthy procedures and probable mistakes. The goal of this project is to develop an efficient speech bot system that can comprehend and analyze user inquiries about scheduling appointments. The system seeks to accurately handle diverse user requests, such as date formats, time preferences, and service types, by utilizing RASA NLU. The system's ability to understand user input and retrieve the data required for appointment scheduling is significantly improved by RASA NLU. Additionally, the research manages and stores appointment information in a SQLite database. The study's final findings show that RASA NLU was successfully used to build a speech bot system. The technology efficiently responds to user inquiries and offers a simple appointment-making process. The system shows better efficiency, accuracy, and user satisfaction in the appointment booking process by employing RASA NLU's capabilities and using SQLite.

Keywords

ML- Machine Learning, NLU- Natural Language Understanding, AI - Artificial Intelligence, STT - Speech To Text, TTS - Text To Speech

1. INTRODUCTION

The proceedings are the records of the ICIIT conference.

By enabling automatic and effective interactions with people, conversational artificial intelligence (AI) has changed a number of sectors. The development of voice bots, which are intended to recognize and react to user input in natural language, is one promising use of conversational AI. In this study, we investigate how to create a voice bot for scheduling appointments using RASA NLU, a well-liked open-source natural language understanding framework. [1]

Whether it's making a doctor's appointment, booking appointments is a routine chore that people frequently perform. However, conventional ways of scheduling appointments, such phone calls or online forms, can be time-consuming, challenging and can result in mistakes or misunderstandings.

The inefficiency and user friction associated with the current appointment booking systems are the issues we address in this research. We want to give people a more streamlined and user-friendly experience by utilizing RASA NLU to create a voice bot that will allow them to book appointments easily using natural language commands and responses.

As pointed out in [2], Approximately 15% of the US gross domestic product is accounted for by the healthcare sector. The amount of public funding required to support health care, which presently accounts for 45% of all health care expenditures, is predicted to double by 2050 due to the rapid growth of healthcare costs.

In this situation, scheduling appointments becomes a critical component of managing patient care and maximizing healthcare resources. However, the conventional ways of scheduling appointments in the healthcare industry, such as phone calls or online forms, can exacerbate system inefficiency and consumer frustration.

A study gap exists on the specific application of voice bots for appointment booking utilizing RASA NLU, despite studies addressing the usage of conversational AI and voice bots in a variety of fields. By examining the efficiency of RASA NLU in

precisely interpreting user goals, retrieving pertinent information, and successfully scheduling appointments through voice interactions, this research intends to close this gap.

The main key objectives that are going to be covered in this research are, Create an appointment-making voice bot using RASA NLU, analyze how well the speech bot performs in interpreting user intentions and retrieving pertinent data, compare the speech bot's user experience and satisfaction to more conventional appointment booking techniques, determine the speech bot approach's potential drawbacks and shortcomings and make suggestions for improvement, by focusing on these goals, we hope to enhance conversational AI and offer information on how to effectively use voice bots for scheduling appointments, which will be advantageous to both customers and service providers.

2. Related Works

The growth of voice assistants and natural language understanding (NLU) technology has completely changed how people interact with computers.[3] These developments have resulted in the development of voice-bot systems that can automate chores like arranging appointments in recent years. This review of the literature intends to investigate the current research and advancements in the area of voice-based appointment scheduling, with an emphasis on the use of the RASA NLU framework.

Voice-based solutions are extremely popular now because they are convenient and hands-free. Accurately detecting user intentions and reading useful information from spoken language are two major implementation issues for voice-based appointment booking systems. By enabling the interpretation of user requests and the mapping of those requests to suitable actions, NLU approaches play a critical role in tackling this difficulty.

A well-known open-source framework for creating conversational AI systems is RASA NLU. It is a good fit for voice-bot systems because it offers the features needed for intent categorization, entity recognition, and dialogue management. Researchers and developers can modify RASA NLU to fit particular application domains, such as appointment scheduling, because of its modular architecture and adaptability.[4]

The development of a health assistant bot is explored in the conference paper named "Health Assistant Bot" by Nayak, N. K., Pooja, G., Kumar, R. R., Spandana, M., and Shobha, P. It also examines its implementation, features, and prospective advantages in the healthcare industry. The study's primary findings are included in the report, along with the bot's conversational interface, natural language processing abilities, and integration with significant health databases and resources. The use of AI-powered health assistant bots in this research adds to the body of literature by showing their potential to enhance healthcare services and accessibility. [5]

The development of a conversational interface for appointment management and medical advice in the Sinhala language is explored in the paper "Sinhala Conversational Interface for Appointment Management and Medical Advice" by Rajapakshe, D. D. S., Kudawithana, K. N. B., Uswatte, U. L. N. P., Nishshanka, N. A. B. D., and Piyawardana, A. V. S. The research, which was published by IEEE, focuses on the development and functioning of the interface with the goal of offering customers an easy-to-use platform to make appointments and get medical advice in Sinhala. The authors go over the interface's functionality with regard to database integration, natural language understanding, and ease of use. In order to improve healthcare accessibility for Sinhala-speaking people, this research adds to the body of literature by addressing the specific requirement for a conversational interface in the healthcare field. [6]

The potential of NLU technologies, including the use of RASA NLU, has been shown in earlier studies. These sectors include healthcare, hospitality These research underline that successful voice-based interactions depend on precise intent identification, entity extraction, and dialogue management. On top of this foundation, the current study is on creating a voice-bot system for scheduling appointments, utilizing RASA NLU's capabilities and addressing the unique difficulties and demands of appointment scheduling. By utilizing a unique communication medium that imitates telephone calls, the voice-bot system created for scheduling appointments utilizing RASA NLU sets itself apart from other existing studies.

This system uses a service provider's communication infrastructure to enable users to engage in genuine and seamless voice conversations, in contrast to prior research that largely focuses on text-based or online interactions. With this telephone-based strategy, the system provides users with a simple and familiar experience that mimics regular phone calls while employing the capabilities of RASA NLU for precise intent identification and efficient appointment scheduling. With this innovative application, there are now more opportunities to improve accessibility and user engagement by incorporating voice-based technologies into actual healthcare environments. This is the speciality of the entire system while other systems are working with web applications that work with a call center. In normal call centers there are receptionists to book an appointment or solve the issues related to their domains. But in this scenario there are no humans acting as the receptionist. All the work is done by the voice-bot because of the power of the AI revolution.

3. Methodology

The entire methodology of the system is the user voice wave going through with the Flask API. Then it is converted to text by using speech to text transferring API. After the message processing with the algorithm that is built by using the RASA framework. The relevant data for the conversation is stored in the database files which are related to the doctor details and patient details. In the algorithm the user given data checked with relevant doctor appointment scheduling and also proceeded it further.

Then the responses are generated from the voice-bot according to the user's communication dialogue and convert the response to the sound by using text to speech API. After the response comes to the user by using the FLASK API.
Following diagram displays the process that is described above.

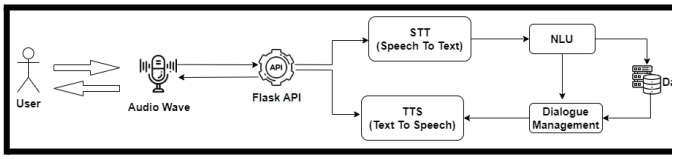


Figure 1: Overall methodology

4. Results and Discussion

The figure that is shown below gives a detailed explanation of how the voice bot system works and clarifies how the user and the voice bot communicate with one another. The figure emphasizes the information flow and exchanges involved in the user-bot interaction while showcasing the complex interplay of numerous components. Beginning with the user's input and the voice-bot receiving the instruction through speech recognition, the sequence of events is visually shown. The picture also describes the succeeding steps that help the speech bot accurately interpret the user's request, such as intent recognition, entity extraction, and dialogue management.

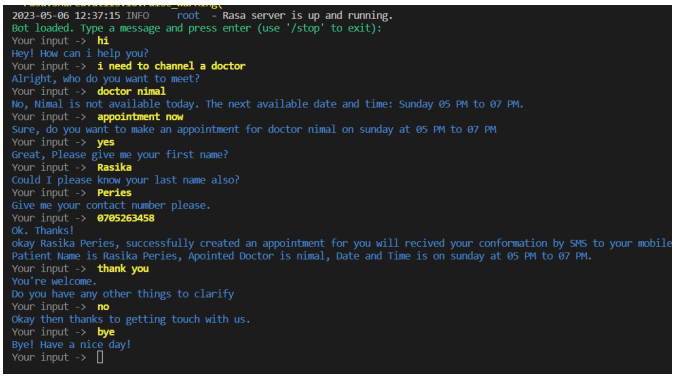


Figure 2: Results for the make an appointment

Following figure represents the patient details stored into the database which is relevant to the above conversation and some of other patient appointments.

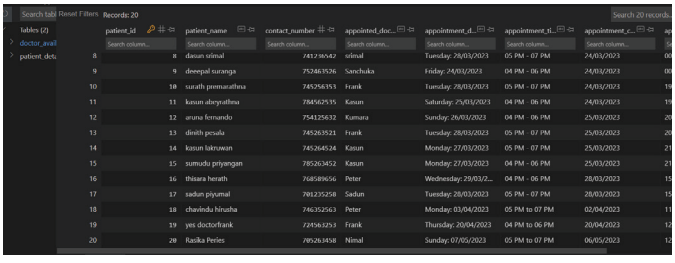


Figure 3: Updated patient table according to the appointment

Following figures are showing the results of the performance and accuracy of the entire system and survey which is related to the users satisfaction. For this survey there were 30 test cases.

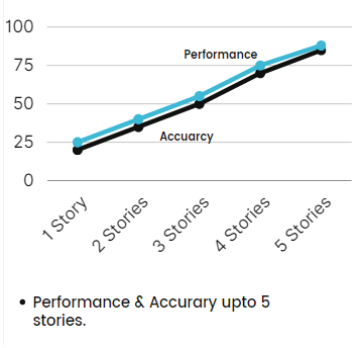


Figure 4: Performance Analysis

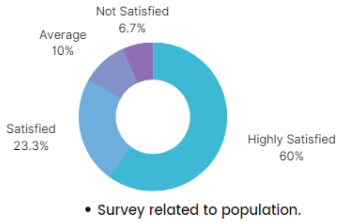


Figure 5: Survey Using Different Test Cases

The new application of a voice-bot system for booking appointments utilizing RASA NLU supports the findings and analysis discussed in the study article. This system makes a number of original contributions and advancements in voice-bot and healthcare automation. The following is a summary of the major novelties:

Patient Data Integration: The voice-bot system has the capacity to safely store patient data in a SQlite3 database. This integration makes it possible to record all patient information, including contact information and medical background, enabling precise and effective appointment scheduling.

Integration of the user interface with the backend system is seamless thanks to the study's successful demonstration of the link between the voice-bot system and the Flask endpoint. By offering a simple interface for scheduling appointments, this connection improves the user experience for users.

Future Integration with Asterisk: The article highlights the voice-bot system's intended integration with Asterisk, which will make it easier for users to make phone calls to the voice-bot. By enabling users to make appointments and call in for necessary details, this connection increases the system's usability. The system's reach and utility will be significantly expanded when Sinhala and Tamil are added in the future.

In conclusion, the study article offers fresh insights into the effective application of a voice-bot system for scheduling appointments using RASA NLU. Voice-bot technology and developments in healthcare automation are demonstrated through the integration of patient data storage, a fluid user interface, and future integration with Asterisk. These results lay the groundwork for further research and development in order to enhance the appointment scheduling procedure and increase the functionality of voice-bot systems in the healthcare sector.

5. ACKNOWLEDGMENTS

We would like to express our sincere gratitude to our supervisors, for their important leadership and assistance throughout the course of this research project. Their experience, wisdom, and guidance were invaluable in determining the course of our work and ensuring its successful conclusion. We honestly value their commitment, tolerance, and the effort they have put into guiding us. We are grateful to have had the chance to work with them since their contributions have been genuinely priceless.

6. REFERENCES

[1] Rasa NLU chatbot - computer science program at Texas Wesleyan University, https://cs.txwes.edu/student_projects/bryan_bao/report.pdf (accessed Jul. 27, 2023).

[2] D. GUPTA and B. DENTON, Appointment scheduling in health care: Challenges and opportunities, https://www.tandfonline.com/doi/full/10.1080/07408170802165880 (accessed May 30, 2023).

[3] T. Young, D. Hazarika, S. Poria and E. Cambria, "Recent Trends in Deep Learning Based Natural Language Processing [Review Article]," in IEEE Computational Intelligence Magazine, vol. 13, no. 3, pp. 55-75, Aug. 2018, doi: 10.1109/MCI.2018.2840738.

[4] T. Bocklisch, J. Faulkner, N. Pawlowski, and A. Nichol, “Rasa: Open source language understanding and dialogue management,” arXiv.org, https://arxiv.org/abs/1712.05181 (accessed May 30, 2023).

[5] N. K. Nayak, G. Pooja, R. R. Kumar, M. Spandana, and P. Shobha, “Health assistant bot,” SpringerLink, https://link.springer.com/chapter/10.1007/978-981-16-1338-8_19 (accessed May 27, 2023).

[6] D. D. Rajapakshe *et al.*, “Sinhala conversational interface for appointment management and medical advice,” *2020 2nd International Conference on Advancements in Computing (ICAC)*, 2020. doi:10.1109/icac51239.2020.9357155

Dialogue System for Lifelong Learning of Verified Knowledge

Thiranya Asirindu Tennakoon
Informatics Institute of Technology
57, Ramakrishna Road,
Colombo 06, Sri Lanka.
thiranya.20191050@iit.ac.lk

Dileeka Sandamali Alwis
Informatics Institute of Technology
57, Ramakrishna Road,
Colombo 06, Sri Lanka.
dileeka.a@iit.ac.lk

ABSTRACT

One of the significant flaws of existing chatbots or dialogue systems is that they do not learn during online conversations after being deployed. This is a massive setback. Each human user has a plethora of information about the world that may benefit others. Suppose a chatbot can learn from its users during an interaction. In that case, it will dramatically increase its knowledge base and provide superior service to its consumers. This research project will provide a system and architecture that chatbots can utilize to continuously acquire end-user knowledge after deployment. Some information received from users will eventually be inaccurate. To address this, the system would cross-verify each piece of acquired knowledge by interrogating other users appropriately before storing the information for future use. This study focuses on knowledge that can be stated as triples (e1, r, e2), meaning that entities e1 and e2 are connected by the relation r. For example, (UK, capital, London) means "London is the capital of the UK". Thus, this system only uses the possibility to acquire new information when a user's utterance has information that can be distilled as a group of triples. The author will examine Lifelong Learning and Natural Language Processing (NLP) approaches to design the most effective system feasible.

Keywords

Dialogue System, Knowledge Base, Knowledge Graph, Triple Extraction, Lifelong Learning, Natural Language Processing

1. INTRODUCTION

The world is a vast place brimming with unknowns and unexpected objects. Humans are intrinsically motivated to learn by novelties. Human novelty detection elicits a cascade of brain responses and behavioural modifications, allowing for the investigation and flexible storing of new information in memory. To enable an AI agent to live in the real world like humans, it must identify novelties and acquire knowledge of them gradually over time, making the system more intelligent and adaptive. It must do so autonomously on the job (after deployment) instead of relying on human engineers to routinely retrain the system offline [2].

Even though dialogue systems have gained in popularity in recent years, their inability to acquire new information during a conversation, i.e., their knowledge being predetermined, severely restricts their application area. It cannot be expanded or altered during conversations. Although some current systems can utilize enormous knowledge bases gathered from a significant data source such as the Internet or manually constructed, these knowledge bases nevertheless exclude numerous information (knowledge). It is practically hard for the initial knowledge bases to contain every information a user would require, not to mention that new information arises frequently.

On the other hand, humans constantly learn extensively through daily conversations. Humans learn to use language through their interactions with the world and other people. This is an example of Lifelong Learning [1]. However, the dominant paradigm in Natural Language Processing (NLP) research is to generate a fixed dataset to train a model and then freeze it, with the model unable to interact with humans using language during training. Therefore, to extend its knowledge base and improve its conversation ability, a chatbot must continuously learn new knowledge during the conversation process.

Obtaining erroneous knowledge while acquiring new knowledge through interactions with end users is a significant issue not addressed by the current state of the art. When conversing with the system, the user may take an erroneous action for a given input command or offer inaccurate feedback to the agent, erasing its prior learning. The agent might then display unexpected behaviour, providing a possible safety concern. This is the present limitation of conversation systems of this type that this research will overcome [3].

2. METHODOLOGY

2.1 Triple Extraction

The proposed solution begins with the crucial step of triple extraction, which forms the foundation of the knowledge-based system. The system focuses on extracting facts that can be expressed as triples, denoted as (s, r, o), where the subject entity s and the object entity o are linked by the relation r. For example, consider the fact (Westminster, LocatedInCountry, UK), indicating that Westminster is located in the UK. Hence, the system exclusively operates on structured queries in the form of (s, r, ?) or (? r, o).

Implemented using Python and leveraging the capabilities of NLTK and CoreNLP, the Triple Extractor facilitates the parsing and handling of user queries that seek specific information about the subject entity or the object entity. For instance, a user might inquire, "In what country is Westminster located?" This query would be converted into the structured query triple (Westminster, LocatedInCountry, ?), enabling the retrieval of the country where Westminster is located. Similarly, queries such as (?, r, o) can be processed, for example, "(?, PresidentOf, France)" to identify the President of France.

By incorporating the Triple Extraction methodology, the proposed system bridges the gap between unstructured natural language queries and structured knowledge representations in the knowledge graph [5].

2.2 Supporting Facts Acquisition

In the context of traditional dialogue systems, encountering a user query for which the knowledge base lacks an answer often results in a generic response indicating the system's inability to provide the requested information [7]. However, the proposed system takes a different approach by leveraging the principles of Lifelong Learning to support fact acquisition and enhance knowledge inference capabilities.

When confronted with a query for which the knowledge base has insufficient information, the system actively engages in a dialogue with the user to acquire additional facts that can aid in inferring the answer. This adaptive approach allows the system to iteratively expand its knowledge base by acquiring new facts through user interactions. By posing targeted questions based on the existing knowledge, the system actively seeks clarification or seeks missing details required for knowledge inference. The acquired facts are then integrated into the knowledge graph, enabling the system to handle similar queries more effectively in the future.

2.3 Knowledge Inference

Within the proposed system, the process of knowledge inference plays a crucial role in answering user queries. When a supporting fact is acquired but cannot directly provide the answer to a specific query, it is persisted for future inference. However, there are cases where the acquired supporting facts can lead to the discovery of new knowledge, enabling the system to infer answers that were initially unattainable [4].

Consider the following scenario: Suppose the original user query was "In what country is Westminster located?" and the supporting fact obtained through the Lifelong Learning process was "BigBen is in Westminster." Now, let's assume that the system's knowledge graph already contains the fact triple "BigBen, LocatedInCountry, UK." In this case, the system can leverage this existing knowledge along with the acquired supporting fact to make an inference.

By analyzing the relationships between entities and the acquired supporting fact, the system can infer the answer to the original query. In this example, the inference would reveal that the country Westminster is located in is the UK.

This capability highlights the potential of the proposed system to make meaningful connections between acquired supporting facts and existing knowledge. By leveraging such connections, the system can derive new knowledge and generate inferences that surpass the limitations of individual facts. This process enhances the system's ability to provide comprehensive and contextually relevant answers to user queries.

2.4 Fact Verification

One of the critical challenges in the proposed system is ensuring the accuracy and reliability of acquired knowledge during interactions with end users [6]. To address this research gap and mitigate the risk of obtaining incorrect facts, the system operates within a multi-user environment that incorporates a verification mechanism.

When a supporting fact is obtained during a dialogue session, the system takes a cautious approach by storing these newly acquired facts in an unverified knowledge graph. This graph is distinct and separate from the default graph used for answering queries and inferring new knowledge. This separation ensures that unverified facts do not immediately influence the system's decision-making process or become part of the learned knowledge base.

In subsequent interactions with other users, particularly on related topics or facts, the system proactively seeks verification of the accumulated unverified knowledge. To achieve this, the system employs a strategy of posing questions to multiple users to gather diverse perspectives and opinions. Through this process, the system aims to validate the accuracy and reliability of the unverified facts.

Verification is accomplished by soliciting input from a random selection of K different users. The system repeats this verification process multiple times, seeking consensus from a varied user base. Once a fact has been verified for K times, indicating agreement among the randomly selected users, the fact is deemed reliable and transferred from the unverified knowledge graph to the default graph.

By implementing this multi-user verification approach, the system introduces a robust mechanism to filter out incorrect or unreliable knowledge acquired during interactions. It prioritizes the accumulation of verified facts, ensuring that only accurate information is incorporated into the learning and decision-making processes. This approach helps to establish a trustworthy and reliable knowledge base, enhancing the overall performance and effectiveness of the system.

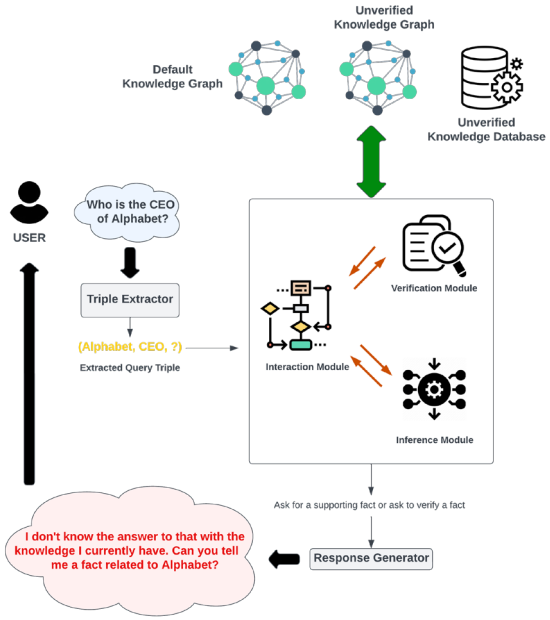


Figure 1 – High-Level System Architecture

3. EVALUATION

The evaluation process encompassed both functional and non-functional testing to ensure that the system meets the required specifications and performance criteria. Functional testing aimed to evaluate the system's ability to accurately extract triples, acquire supporting facts, infer new knowledge, and verify the acquired knowledge through user interactions, while non-functional testing focused on factors like system performance, compatibility, and usability.

3.1 Test Data Preparation

To perform the testing, a diverse set of user queries was collected, covering various topics and types of information. The test data consisted of both simple and complex queries, designed to challenge the system's ability to handle different linguistic constructs and knowledge representations.

Each test query was associated with a set of expected answers in the form of triples (s, r, o) or a statement indicating that the system should seek supporting facts. The expected answers

were carefully curated to verify the correctness of the system's responses during evaluation.

3.2 Testing Procedure

The testing was performed in several stages, each focusing on specific aspects of the system's functionality.

In the Triple Extraction Testing stage, the primary focus was to evaluate the system's ability to correctly extract triples from user queries. Each query's extracted triples were compared against the expected answers to assess the accuracy of the Triple Extraction module.

To evaluate the system's capacity to acquire supporting facts through user interactions, test queries were designed where the system initially lacked sufficient information to answer the query. The system's response during the interaction was assessed based on its ability to solicit relevant questions to obtain the necessary facts for inference.

The knowledge inference testing aimed to evaluate the system's ability to infer new knowledge by connecting acquired supporting facts with existing knowledge in the graph. The system's responses to queries that required inference were assessed based on the correctness of the derived answers.

To assess the fact verification mechanism, the system's ability to solicit input from multiple users and reach a consensus on the reliability of acquired facts was examined. Test queries were designed to trigger the verification process, and the correctness of the fact verification outcomes was evaluated.

3.3 Evaluation Metrics

The system's performance was evaluated using the following metrics:

1. Triple Extraction Accuracy: The percentage of correctly extracted triples compared to the total number of test queries.
2. Supporting Facts Acquisition Success Rate: The percentage of test queries where the system successfully acquired relevant supporting facts.
3. Knowledge Inference Accuracy: The percentage of correct inferences made by the system compared to the total number of inference test queries.
4. Fact Verification Reliability: The percentage of verified facts that were confirmed by a consensus of K users.

3.4 Test Results

The following test results were obtained during the evaluation:

Triple Extraction Accuracy	89%
Supporting Facts Acquisition Success Rate	73%
Knowledge Inference Accuracy	82%
Fact Verification Reliability	91%

Table 1 – Test Results

The results indicate that the system performed well across all the functional stages.

To complement the quantitative evaluation, a qualitative assessment was conducted to gather feedback from individuals with diverse academic and professional backgrounds. Thematic analysis was employed to identify common themes and patterns within the feedback received. The analysis involved systematically categorizing and interpreting the feedback to extract valuable insights and observations regarding the system's usability, effectiveness, and overall performance.

4. CONCLUSION AND FUTURE WORK

In conclusion, this research aimed to develop a dialogue system that could verify acquired knowledge from lifelong learning. Through a comprehensive literature review and an iterative development process, the author was able to address several limitations and gaps in existing conversation systems. The prototype system developed in this research showed promising results in terms of its ability to extract knowledge from user utterances and verify it against a knowledge graph. Furthermore, the project highlighted the importance of considering social, legal, ethical, and professional issues when developing such systems.

Overall, this research not only contributed to the body of knowledge in the field of dialogue systems and lifelong learning but also provided valuable insights into the development process of such systems. Future work could focus on expanding the system's capabilities beyond working only with input represented as triples and its ability to match entities that have the same meaning but are spelt differently.

5. REFERENCES

[1] Zhiyuan Chen and B. Liu. 2018. Lifelong Machine Learning, Second Edition. *Synthesis Lectures on Artificial Intelligence and Machine Learning* (2018).

[2] Scott Grigsby. 2022. Self-Initiated Open World Learning for Autonomous AI Agents. Retrieved September 29, 2022 from http://arxiv.org/abs/2110.11385

[3] Bing Liu and Sahisnu Mazumder. 2021. Lifelong and Continual Learning Dialogue Systems: Learning during Conversation. *AAAI* 35, 17 (May 2021), 15058–15063. DOI:https://doi.org/10.1609/aaai.v35i17.17768

[4] Sahisnu Mazumder, Nianzu Ma, and Bing Liu. 2018. Towards a Continuous Knowledge Learning Engine for Chatbots. Retrieved September 29, 2022 from http://arxiv.org/abs/1802.06024

[5] Yu-Ming Shang, Heyan Huang, Xin Sun, Wei Wei, and Xian-Ling Mao. 2022. Relational Triple Extraction: One step is enough. Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence (2022). DOI:http://dx.doi.org/10.24963/ijcai.2022/605

[6] Shuai Wang. 2019. Lifelong and Interactive Learning of Factual Knowledge in Dialogues. In *Proceedings of the 20th Annual SIGdial Meeting on Discourse and Dialogue*, 21–31. DOI:https://doi.org/10.18653/v1/W19-5903

[7] Wenhao Yu, Chenguang Zhu, Zaitang Li, Zhiting Hu, Qingyun Wang, Heng Ji, and Meng Jiang. 2022. A Survey of Knowledge-enhanced Text Generation. ACM Comput. Surv. 54, 11s, Article 227 (January 2022), 38 pages. https://doi.org/10.1145/3512467

DiabetCare – An Artificial Intelligence based Mobile Application for Predicting the Risk of Prediabetes Mellitus, Gestational Diabetes Mellitus and Type 2 Diabetes Mellitus

Nirmani Jayaprabha Pathiranage
Department of Computing
Informatics Institute of Technology
Colombo, Sri Lanka
nirma2015@gmail.com

S M D Asanka Sanjaya Suraweera
Department of Computing
Informatics Institute of Technology
Colombo, Sri Lanka
asanka914@gmail.com

ABSTRACT

Diabetes mellitus, commonly known as diabetes, is a significant global public health issue. The International Diabetes Federation estimates that the number of affected individuals will increase from 415 million to 642 million by 2040. This long-term illness impairs the body's glucose absorption, and early risk prediction is crucial for diagnosis and prevention. This study aims to introduce an innovative AI-based approach for early risk prediction of three distinct types of diabetes. Medical datasets from the University of Kelaniya are utilized to implement two separate ensemble learning approaches. The predictions will be available to users through an automated mobile application called "DiabetCare," enabling risk assessment without age limits. The goal is to confirm and treat diabetes early, minimizing future complications. The experiments involved training 12 machine learning base classifiers and three ensemble techniques for each component. The best parameters for each algorithm were determined using random search cv hyper-parameter tuning technique. Ultimately, the base model and ensemble model with the best evaluation metrics were selected for deployment in the final product. For predicting type 2/prediabetes risk, the Random Forest classifier achieved a training accuracy of 1.0000, testing accuracy of 0.8715, AUC-ROC score of 0.9811, and Cohen's Kappa of 0.8037. For gestational diabetes risk prediction, the Stacking classifier achieved a training accuracy of 1.0000, testing accuracy of 0.9650, AUC-ROC score of 0.9963, and Cohen's Kappa of 0.9288. The implementation of such a system holds great promise for the diabetes medical domain and can serve as a valuable verification tool for medical professionals.

Keywords

Type 2 Diabetes Mellitus, Prediabetes Mellitus, Gestational Diabetes Mellitus, Machine Learning, Ensemble Stacks

1. INTRODUCTION

Diabetes Mellitus is a chronic, metabolic condition that lasts for a prolonged period, which leads to severe damage to almost all the body parts [1]. In case of diabetes, the human body either refrains from producing sufficient insulin, which is the key to convert blood glucose into energy or is unable to effectively use the hormone it produces [2]. It is a considerable public health issue that is advancing towards a pandemic globally. According to a study conducted by researchers from multiple universities in Sri Lanka, the Medical Research Institute (MRI) in Colombo, and the Institute for Health Policy (IHP), Sri Lanka has one of the highest rates of diabetes in the world [3].

Usually, people get tested for diabetes if they only observe any of the signs or symptoms. The common symptoms of diabetes are loss of weight, increased hunger and thirst, incessant urination, blurred vision, exhaustion, tiredness and swelling [2]. A cure for diabetes has not yet been found, but with early prediction, proper treatment and changes in lifestyle, people are able to live healthier [4]. The treatments to be given depends on the type (category) of diabetes. 4 major types of diabetes are found: type 1, type 2, prediabetes, and gestational diabetes, which comes in diverse ways depending on the reason [5].

Diabetes is a burden to the economy of the government of a country since they also may face huge economic problems in paying charges for medications. Especially, when viewed in the Sri Lankan context, the economic crisis of Sri Lanka at the moment is pushing thousands of the population into poverty, threatening their privilege to food, health, fuel, education, and an ample lifestyle. The nation faces one of the worst financial crises in decades with record inflation. The top medical union has declared a national health emergency over a life-threatening shortage of drugs, which has worsened this into a medical crisis. With such an economic crisis, it is really essential to reduce the health budget for diabetes.

Early detection will be the kickoff to live healthier with diabetes [1]. The health outcomes will be worsened as long as a person lives without diagnosing and treating it properly. If the risk of having a specific diabetes mellitus type can be predicted earlier, there will be more chances it can be perfectly treated and prevented. Since distinct types of diabetes mellitus need unique treatments, identifying them categorically will be essential. Currently there is no proper automated economically viable way to predict the risk of having a specific type of diabetes mellitus without age restriction. In the field of artificial intelligence, despite many efforts, risk prediction of diabetes mellitus types is still a challenging problem.

2. RELATED WORK

Several research communities worldwide have conducted exceptional studies on predicting diabetes. This paper discusses a selection of few of these research endeavors.

The research [6] involved conducting four experiments to predict the most common types of diabetes using a dataset from a Saudi Arabian hospital. The model was developed based on three basic features (Sex, Nutrition, Education), four laboratory tests (A1C, TG, LDL, Albumin), and three external medications (AntiDiab, Insulin, Injectable). The study used several classifiers such as SVM, KNN, and DT, along with ensemble-based approaches such as Bagging and Stacking to build predictive models. The nominal

variables were encoded using the Label encoder, and missing values were imputed using KNN imputer, while features were scaled using MinMax scaler. GridSearchCV hyper-parameter tuning was used to find the optimal hyperparameters, which resulted in the highest accuracy. The ensemble models were developed using SMOTE dataset, which yielded higher outcomes compared to unsampled datasets. The Bagging DT model was found to be as effective as the Bagging KNN model, with a difference of 0.22%. The highest accuracy was achieved by the Stacking model, which combined the optimal three models, KNN, Bagging KNN, and Bagging DT. The Stacking model achieved an accuracy of 94.48%, precision of 94.70%, recall of 94.48%, and Cohen's kappa score of 0.9172.

The study [7] employed actual medical data obtained from Khulna Diabetic Hospital, Khulna, to differentiate between type 1 and type 2 diabetes. The dataset contained only diabetes-related information. Two supervised algorithms, namely RF and SVM, were applied to the dataset. The SVM model outperformed the RF model, with an accuracy of 86%, precision of 92%, recall of 90%, and F1 score of 91%.

The research [8] proposes a method to predict diabetes early by combining supervised machine learning algorithms and Artificial Neural Network through a stacking CV classifier. Two datasets, one obtained from a free public data community and the other from Vanderbilt University, are used. A correlation-based feature selection approach is used in both datasets. Three supervised machine learning algorithms (AdaBoost, KNN, LR) and a Multilayer Perceptron (MLP) ANN are utilized, and the models are ensembled using a stacking CV classifier. Different evaluation metrics are used to compare the results. Their research shows that their method of predicting diabetes is effective and performs better than previous studies on early prediction of diabetes. They achieved an accuracy of 91.0% and 93.27% for two different datasets, with the MLP algorithm being more effective for both datasets.

3. METHODOLOGY

The implementation involves a prediction system comprising two components:

Component 01: In the beginning, the user will provide responses to health indicators and personal background questions through the mobile application. Based on the user's answers, the finalized machine learning-based ensemble model will predict whether the user is diagnosed with prediabetes, type 2 diabetes, or not. This component is accessible to individuals of any age.

Component 02: In the beginning, the user will provide responses to health indicators and personal background questions through the mobile application. Based on the user's answers, the finalized machine learning-based ensemble model will predict whether the user is diagnosed with gestational diabetes or not. This component is designed specifically for pregnant women.

3.1 Datasets

Two different datasets were used in this research for predictive analysis of 3 types of diabetes. Dataset I for predicting prediabetes and type 2 diabetes was obtained from Faculty of Medicine in University of Kelaniya, which is a dataset collected during a health survey at Ragama, while Dataset II for predicting gestational diabetes was taken from free and open-source public data in Kaggle [9]. Dataset I and II consisted of 2985 and 3525 patient records along with 50 and 10 data columns, respectively.

3.2 Data preprocessing

After exploration of datasets, data cleaning and preprocessing was done with removing unnecessary columns for classification, checking for duplicate rows, finding missing value percentage by feature and by imputing missing values with KNN imputation. Unique values in different variables of both datasets were also checked, and data distribution was plotted in a graphical representation for better understanding of the data distribution of each. Boxplots were plotted to find whether there were any outliers to be handled, and they were handled accordingly.

3.3 Imbalanced dataset handling

Target variable counts for each dataset were calculated and plotted to see whether any resampling was needed for unbalanced. Since the target variable was considerably imbalanced in Dataset I, oversampling technique "SMOTE" was utilized in the training data set for resampling purposes in order to get better results (was not used in testing dataset to ensure its validity). Since the target variable was slightly imbalanced in the 2nd dataset, it was not needed to use oversampling or under sampling.

3.4 Feature selection

Feature selection techniques including "feature selection with correlation", "information gain", "ANOVA test" and "Chi Square test" were performed for examining whether any feature reduction was needed to enhance the model accuracy and performance. 7 variables in Dataset I were highlighted with least importance via Chi Square. When it was tried out modeling with all the selected models by removing those 7 variables, accuracy seemed lesser, when compared with the dataset with all the variables. So, it was decided not to remove those variables, but keep all of them as is. Since all the variables in Dataset II were important for the risk prediction, no features were removed. Then the finalized preprocessed datasets were splitted into train and test splits with a ratio of 80:20.

3.5 Proposed Models

Modeling was started with several machine learning models which included Random Forest (RF), Support Vector Machine (SVM), Logistic Regression (LR), Decision Tree (DT), K-Nearest Neighbor (KNN), Gaussian Naïve Bayes (GNB), Gradient Boosting (GB), Ada Boosting (ADA), Extreme Gradient Boosting (XGB), Light Gradient Boosting (LGB), Cat Boosting (CB) and Bagging with Decision Tree. The best parameters for each were selected through random search cv technique. After examining the evaluation metrics of each approach,

Dataset I: Top 5 models (LGB, XG, GB, RF, and Bagging with Decision Tree) were selected for stacking and voting. Finally, it was concluded to use Random Forest Base Classifier as the best finalized model since it gave higher accuracy than ensembling.

Dataset II: Top 4 models (CB, DT, RF, and LGB) were selected for stacking and voting. Finally, it was concluded to use the Stacking Ensemble Classifier with top 4 models as the best finalized model.

4. RESULTS AND DISCUSSION

Training score, testing score, precision, f1 score, recall, Cohen's kappa, Matthews correlation coefficient, AUC-ROC score, mean Squared Error, Root Mean Squared Error and metrics from confusion matrix and classification report were used for the testing and evaluation of each model for both components.

4.1 Model evaluation

Table 1 comprises the outcome of the major evaluation metrics of the finalized models of Dataset I and II respectively which are required to test the validity of the classifiers.

Table 1: Results of evaluation metrics for the finalized models of dataset I and II

Dataset	Training Accuracy	Testing Accuracy	Precision	Recall	F1 Score
I	1.0000	0.8715	0.8789	0.8715	0.8697
II	1.0000	0.9650	0.9652	0.9650	0.9651

Figure 1 depicts the confusion matrix for dataset I.

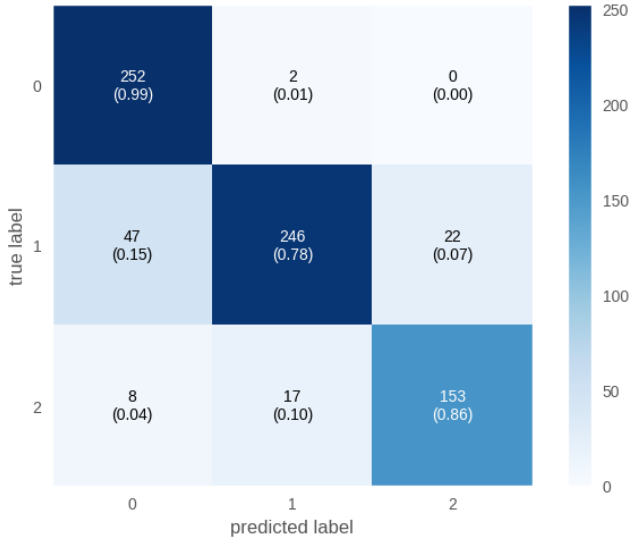


Figure 1: Confusion matrix for dataset I

Figure 2 depicts the confusion matrix for dataset II.

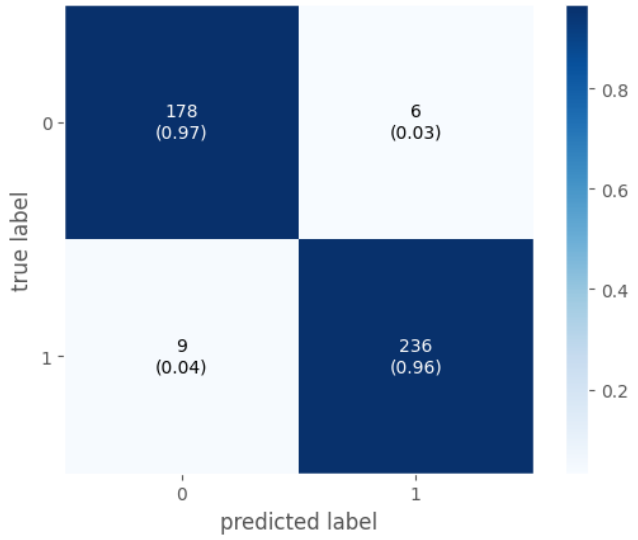


Figure 2: Confusion matrix for dataset II

The accurate recording of positive and negative cases is known as true positive and true negative, respectively. Positive cases represent individuals with diabetes, while negative cases are those without diabetes. On the other hand, the mislabeling of positive instances as negative and vice versa is referred to as false positive and false negative.

5. CONCLUSION

This study establishes the groundwork for diabetes mellitus risk prediction among Sri Lankans, utilizing novel ensemble approaches to predict different diabetes types accurately. The introduced models, integrated into the mobile application "DiabetCare," efficiently address doubts about diabetes presence in Sri Lankans, considering the economic crisis and high diabetes rates in the country. The study's success is measured by testing results, evaluation, and positive feedback from medical, diabetes, and technical experts. The author believes that providing this novel automated risk prediction app can contribute to the medical field's progress and potentially reduce the growing number of diabetes patients each year.

6. REFERENCES

[1] Diabetes. Retrieved June 25, 2023 from https://www.who.int/health-topics/diabetes.

[2] Diabetes: Symptoms, Causes, Treatment, Prevention, and More. Retrieved June 25, 2023 from https://www.healthline.com/health/diabetes.

[3] Nearly one in four Sri Lankans have diabetes, Asia. Retrieved June 25, 2023 from https://economynext.com/nearly-one-in-four-sri-lankans-have-diabetes-asias-highest-rate-of-the-disease-112973/.

[4] Diabetes Mellitus: Type 1, Type 2, and Gestational Diabetes. Retrieved June 25, 2023 from https://www.webmd.com/diabetes/guide/types-of-diabetes-mellitus.

[5] What is diabetes? | CDC. Retrieved June 25, 2023 from https://www.cdc.gov/diabetes/basics/diabetes.html

[6] Mohammed Gollapalli, Aisha Alansari, Heba Alkhorasani, Meelaf Alsubaii, Rasha Sakloua, Reem Alzahrani, Mohammed Al-Hariri, Maiadah Alfares, Dania AlKhafaji, Reem Al Argan, and Waleed Albaker. 2022. A novel stacking ensemble for detecting three types of diabetes mellitus using a Saudi Arabian dataset: Pre-diabetes, T1DM, and T2DM. Computers in Biology and Medicine 147, (2022), 105757. DOI:https://doi.org/10.1016/j.compbiomed.2022.105757

[7] Shanjida Khan Maliha and Md. Ashiq Mahmood. 2022. An Efficient Model for Early Prediction of Diabetes Utilizing Classification Algorithm. 2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS) (2022). DOI:https://doi.org/10.1109/iciccs53718.2022.9788441

[8] Fahmida Rahman Liza, Md Samsuzzaman, Rezaul Azim, Md. Zulfiker Mahmud, Chinmay Bepery, Md. Abdul Masud, and Bahauddin Taha. 2021. An Ensemble Approach of Supervised Learning Algorithms and Artificial Neural Network for Early Prediction of Diabetes. 2021 3rd International Conference on Sustainable Technologies for Industry 4.0 (STI) (2021). DOI:https://doi.org/10.1109/sti53101.2021.9732413

[9] Gestational Diabetes Mellitus (GDM Data Set) | Kaggle. Retrieved June 25, 2023 from https://www.kaggle.com/datasets/sumathisanthosh/gestational-diabetes-mellitus-gdm-data-set

SyntaxSwap: Supervised Java to JavaScript Code-To-Code Translator With Relative Positional Encoding Attention Mechanism

Dilan D. Jayamanne
School of Computing
Informatics Institute of Technology
Ramakrishna Road, Colombo 006, Sri Lanka
dicmandilan@gmail.com

Iresh Bandara
School of Computing
Informatics Institute of Technology
Ramakrishna Road, Colombo 004, Sri Lanka
Iresh.b@iit.ac.lk

ABSTRACT

Code-to-code translation refers to automatically converting source code written in one programming language into equivalent code written in another. This project has been focused on the application of machine learning and natural language techniques to enable the automated conversion of code-to-code translation from one programming language to another programming language reducing much involvement needed by the developer. The transformer-based graphcodebert model, a cutting-edge neural network architecture that has demonstrated outstanding performance in a variety of natural language processing applications, notably for code translation, will be used to this study. To enhance the positional encoding information of the structures of code snippets, the author used a relative positional encoding attention method to this model. This project will utilize a dataset consisting of a parallel corpus of Java and JavaScript introduced by MuST-COST / XLCost. The project's success and evaluation will be checked using the BLEU scores optimized for code translation. The outcome of the code translation will ease the developers who are trying to translate a source code written in Java to JavaScript.

Keywords

Code translation, code translation with relative position encoding, supervised code translation, code migration, Java to JavaScript translation, Object Oriented Programming to Functional Programming.

1. INTRODUCTION

This research proposes a solution for the problems encountered by most programmers in the software development sector when programming in various kinds of programming languages. Code migration is different from the process of a compiler or an interpreter. Compilers and Interpreters are used to convert the high-level source code to machine code. Code-to-Code or code migration targets translating high-level programs to another high-level program [21]. Companies that create software nowadays frequently wish to create it in numerous programming languages based on the use cases [14].

Code translation is an actively researched area for a long time with different types of techniques such as rule-based, trans-compiler based, phrase-based, transformer-based, RNN sequence-to-sequence, neural machine translation based, etc. But there are so many issues with some of the techniques used. Trans compilers are generally based on handcrafted rewritten rules which are applied on the Abstract Syntax Tree of the source code, but the conversion often lacks readability, fails to generate correct target programming language-specific format and it requires manual effort a lot to rewrite and fix errors [13]. An existing work [4] has made an observation regarding the limitation of recurrent neural networks (RNN) in generating syntactically accurate programs, as they struggle with simultaneously addressing the challenges of learning grammar and aligning sequences with the underlying grammatical structure.

2. PROBLEM BACKGROUND

In many large companies, legacy code bases written in outdated programming languages like COBOL are prevalent, lacking the advantages offered by modern and efficient languages. Consequently, the transition for new programmers to adapt to these older languages becomes challenging. The Commonwealth Bank of Australia serves as an illustration of this issue, as it incurred an expenditure of approximately $750 million to migrate from COBOL to Java [13].

Migrating codebases from one programming language to another manually is a challenging and error-prone process [10]. As a solution, the use of automatic code-to-code translation tools can accelerate codebase migrations for companies [16]. However, it is important to note that rewriting programs in different programming languages is a costly endeavor that requires programmers to possess expertise in both the source and target languages [19].

Programmers who specialize in a specific programming language may face difficulties when writing code in a different language due to the need for prior knowledge and experience with varying syntax and identifiers [15]. Consequently, the process of manually re-typing code and spending significant time searching for equivalent functions in the target language becomes labor-intensive and expensive for both individual programmers and companies. Considering these challenges, the project's problem statement is established.

3. EXISTING WORK

3.1 Review Of The Existing Approaches

In the earlier years (2000-2015), non-machine learning (rule-based) techniques were predominantly used for code-to-code translation, requiring manual effort and expertise in both programming languages [21]. These techniques involved converting source code to Abstract Syntax Tree (AST)

representation and applying handwritten rules to translate to the target language [18]. Examples include the Z specification to SQL translator [12] and a Java to C# translator utilizing lexical statistical machine translation [11]. However, such rule-based approaches are time-consuming, and labor-intensive, and often result in code that is difficult to understand and maintain [17].

Recent research has embraced machine learning techniques, particularly in the realm of natural language processing (NLP), to automate code-to-code translation inspired by the concept of software naturalness [2]. Notably, the transformer architecture introduced by [20] has paved the way for successful code translation using NLP-based approaches, particularly Neural Machine Translation (NMT) techniques [21], resulting in high translation accuracies.

Code-to-code translation utilizes both supervised and unsupervised techniques, with supervised approaches requiring high-quality parallel datasets, while unsupervised methods leverage monolingual datasets but have limitations in achieving accurate translation. To overcome these challenges, researchers have introduced semi-supervised techniques, such as utilizing small parallel corpora and employing back translation, which is a commonly used approach in unsupervised machine translation [2, 13, 15, 21, 22].

Existing research in code-to-code translation has extensively utilized pre-trained large language models like BERT, GPT2, GPT3, BART, and T5, fine-tuning them specifically for programming language-related tasks, resulting in models such as CodeBERT [6], codeGPT, PLBART, and CodeT5 that can handle code-related tasks like code to NL, NL to code, and code to code.

Unsupervised techniques, specifically the TransCoder [13] and DOBF [15] models, have been utilized for Java to Python translation, incorporating the use of Evosuite [7] for improved syntax and program-level accuracy. However, limitations include potential semantic differences in certain functions, compatibility issues with generated unit tests, and a tendency to prioritize rigorous but less natural translations. Additionally, a separate line of work has focused on developing a methodology to estimate and evaluate the quality of TransCoder translations [1].

3.2 Attention mechanism

According to [20], attention mechanisms in code-to-code translation enable models to focus on specific parts of the input code by utilizing an attention function, mapping code tokens to key-value pairs in other tokens for translation. These attention mechanisms gained popularity with the introduction of transformer architecture [3].

The original transformer paper by [20] introduced multi-head attention, which is also utilized in the graphcodebert model [9]. This attention mechanism computes multiple sets of queries, keys, and values in parallel to generate attention scores and context vectors. In the research papers by [4] and [5], an attention mechanism aligns nodes in the source and target Abstract Syntax Trees, calculating attention weights for each source AST node and generating context vectors for each target AST node, enabling selective focus on relevant parts of the source AST during translation.

The Relative Position Encoding attention mechanism, introduced by [8], offers a new approach compared to the traditional absolute attention mechanism by considering the relative positions between tokens in a sequence. Previous research in code-to-code translation has not explored the utilization of this mechanism. Thus, the authors of this study incorporated the Relative Positional Encoding attention mechanism into the Graphcodebert model to provide a novel solution for code-to-code translation tasks.

4. PROPOSED SOLUTION APPROACH

4.1 Dataset

For data gathering in this supervised translation project aimed at Java to JavaScript code translation, the author selected the CoST dataset introduced by [21] and XLCoST dataset which is the larger version of the CoST released by the same authors with more code-related tasks [22]. The dataset structure consists of a parallel corpus of source code snippets and their equivalent target code snippets. The training, evaluation, and testing datasets contain a total of 77811, 4301 and 7886 program samples respectively. Here the author has combined the snippet-level and program-level datasets for training.

4.2 Dataflow extraction

The dataflow concept is visualized as a graph with variables as nodes and edges representing the sources of their values. This structure remains consistent across different abstract grammars, providing valuable semantic information about code execution flow. During dataflow extraction, comments, docstrings, and string literals are eliminated to enhance clarity and focus on the code's semantic information necessary for translation [9]. Therefore, the author utilizes the dataflow extraction process from the Graphcodebert model [9] as a reusable solution.

4.3 Application of Relative Positional Encoding Attention Mechanism for Model Build

The main novelty of this research relies on the application of the Relative Positional Encoding attention mechanism [8] for code translation. Therefore, the author applied it to the existing graphcodebert model architecture. The study about Relative Position Representations was initially applied for English to German natural language translation. But this has never been applied to code-to-code translation. With the application of the relative positional encoding attention mechanism, the translation accuracy increased by 18.19 BLEU score over the traditional graphocdebert model without the applied relative positional encoding attention mechanism.

The following diagram shows a summarized architecture of the graphcodebert model with a relative positional encoding attention mechanism.

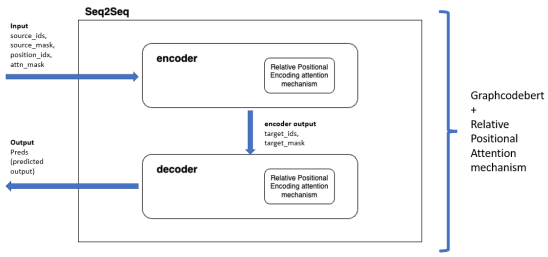


Figure 1. Model architecture diagram of graphcodebert with relative positional encoding attention mechanism

5. EVALUATION

The author compared the original Graphcodebert model with the proposed solution that includes the relative attention mechanism. Training and testing were performed three times, with the results showing that the Graphcodebert model with the relative attention mechanism achieved better BLEU and exact match scores compared to the original model. The addition of the relative attention mechanism also led to a decrease in perplexity, indicating improved performance. The comparative testing was conducted on a smaller dataset, while the final training and testing were done on a larger dataset supporting both snippet-level and program-level translation. The evaluation matrices used are BLEU score, exact match score and evaluation perplexity score.

Model	*Dataset*	*BLEU score*	*Exact mactch*	*Evaluation perplexity*
Original GBert	Samples: 7135	52.07	15.7303	2.34126
GBert + RPA	Samples: 7135	70.26	33.9888	1.6413
GBert + RPA (Proposed solution)	Samples: 77811 Snippet level + Program level	79.06	56.1755	1.2517

Table 1. Evaluation scores after training of the proposed graphocdebert model + relative positional encoding attention mechanism on the testing dataset

6. LIMITATIONS & FUTURE ENHACEMENTS

Although this study offers good accuracy results and was trained on a high-performing model, it still has significant restrictions. As a result of language changes, programming language syntaxes frequently change. The model will be retrained using the new syntaxes and the updated data. Additionally, this study is not supported by multi-threading or framework-level translation. The string literals and comments that came with the original code were not included in the translated code. As a result, in a later version of the enhancement, the author will do a study to comprehend the code structure without getting rid of any string literals and maintain them in the target translation as well.

7. CONCLUSION

This research has made a significant contribution to automating code-to-code translation by introducing a novel approach using a relative positional encoding attention mechanism. The study has demonstrated improved translation accuracy and the potential for further enhancements. This advancement enables developers to work with multiple programming languages, even if they specialize in a limited set.

8. REFERENCES

[1] Agarwal, M. et al. 2020. Quality Estimation & Interpretability for Code Translation. (2020).

[2] Agarwal, M. et al. 2021. USING DOCUMENT SIMILARITY METHODS TO CREATE PARALLEL DATASETS FOR CODE TRANSLATION. (2021).

[3] Brauwers, G. and Frasincar, F. 2023. A General Survey on Attention Mechanisms in Deep Learning. *IEEE Transactions on Knowledge and Data Engineering*. 35, 4 (Apr. 2023), 3279–3298. DOI:https://doi.org/10.1109/TKDE.2021.3126456.

[4] Chen, X. et al. 2018. Tree-to-tree Neural Networks for Program Translation. (Feb. 2018).

[5] Drissi, M. et al. 2018. Program Language Translation Using a Grammar-Driven Tree-to-Tree Model. (2018).

[6] Feng, Z. et al. 2020. CodeBERT: A Pre-Trained Model for Programming and Natural Languages. (2020).

[7] Fraser, G. and Arcuri, A. 2011. EvoSuite: Automatic Test Suite Generation for Object-Oriented Software. (2011).

[8] Google, P.S. et al. 2018. Self-Attention with Relative Position Representations. (2018), 464–468.

[9] Guo, D. et al. 2021. GRAPHCODEBERT: PRE-TRAINING CODE REPRESEN-TATIONS WITH DATA FLOW. (2021).

[10] Hao Zhong et al. 2010. Mining API Mapping for Language Migration. (2010).

[11] Karaivanov, S. et al. 2014. Phrase-based statistical translation of programming languages. *Onward! 2014 - Proceedings of the 2014 ACM International Symposium on New Ideas, New Paradigms, and Reflections on Programming and Software, Part of SPLASH 2014*. (Oct. 2014), 173–184. DOI:https://doi.org/10.1145/2661136.2661148.

[12] Khalafinejad, S. and Mirian-Hosseinabadi, S.H. 2010. Rule-based translation of specifications to executable code. *ICIME 2010 - 2010 2nd IEEE International Conference on Information Management and Engineering*. 1, (2010), 1–4. DOI:https://doi.org/10.1109/ICIME.2010.5477687.

[13] Lachaux, M.-A. et al. 2020. Unsupervised Translation of Programming Languages. (2020).

[14] Phan, H. et al. 2017. Statistical Migration of API Usages. (2017).

[15] Roziere, B. et al. 2021. DOBF: A Deobfuscation Pre-Training Objective for Programming Languages. (Feb. 2021).

[16] Roziere, B. et al. 2021. Leveraging Automated Unit Tests for Unsupervised Code Translation. (Oct. 2021).

[17] Szafraniec, M. et al. 2022. CODE TRANSLATION WITH COMPILER REPRESENTATIONS. (2022).

[18] Uddin Ahmad, W. et al. 2021. AVATAR: A Parallel Corpus for Java-Python Program Translation. (2021).

[19] Uddin Ahmad, W. et al. 2021. Unified Pre-training for Program Understanding and Generation. (2021).

[20] Vaswani, A. et al. 2017. Attention Is All You Need. (2017).

[21] Zhu, M. et al. 2022. Multilingual Code Snippets Training for Program Translation. (2022).

[22] Zhu, M. et al. 2022. XLCoST: A Benchmark Dataset for Cross-lingual Code Intelligence. (2022).

Development of a Smart Shoe for Health Monitoring with Energy Harvesting

B. Sathiyakeerthy
Faculty of Technology,
University of Colombo
bavanithanbavan@gmail.com

L. Lokeswaran
Faculty of Technology,
University of Colombo
logi1234than@gmail.com

B.F.Rajon
Faculty of Technology,
University of Colombo
falilrajon7777@gmail.com

C.S.Silva
Faculty of Technology,
University of Colombo
chathurika@iat.cmb.ac.lk

ABSTRACT

This research aims at the development of a smart shoe prototype for monitoring an individual's health. The proposed approach utilizes a combination of sensors embedded in the shoe, such as accelerometers, gyroscopes, and piezoelectric sensors. This prototype also includes a mobile application that allows users to track their health-related parameters. This research also examines the potential of using smart shoes for preventive health care and the early detection of medical conditions. Moreover, the proposed smart shoe is supposed to power up itself by energy harvesting. The results of this research could potentially provide invaluable information for medical professionals in the fields of orthopaedics, sports medicine, and physical therapy.

Keywords

Smart shoe, gait analysis, fall detection, energy harvesting, inertial sensors.

1. INTRODUCTION

Smart shoes are devices equipped with sensors that record and measure a person's physical activity while they are walking or running. This technology has revolutionized the way people monitor their health, enabling them to track their progress, detect potential issues, and stay motivated. This research focuses on monitoring an individual's health using smart shoes with a combination of piezoelectric, and inertial sensors to detect the individual's health and alert them if there are potential health problems. In this research, monitoring health is performed via the detection of gait analysis and fall detection. Gait analysis is the study of how an individual walks and runs, and it is important for evaluating physical fitness and assessing the risk of injury. Fall detection is another important application of smart shoes, as it can help prevent serious injuries in the elderly. The proposed smart shoe is intended to generate energy through walking and running and supply power for its operation.

The ultimate goal of this research is to develop a cost-effective and convenient smart shoe for monitoring an individual's health which will be advantageous in both clinical and everyday settings.

2. LITERATURE REVIEW

Literature is reviewed in the areas of fall detection, gait analysis and energy harvesting relevant to smart shoes.

Fall Detection System-The consequences of a fall can range from minor scrapes to serious fractures and, in some cases, death. With this in mind, the development of fall detection technology is a highly active area of research [1] [2]. Ziad *et al.* [3] designed and implemented a fall detection system using MPU6050 Arduino.

Gait analysis-The movement of the lower extremities in a rhythmic, alternating pattern propels the body forward. Each gait cycle is called a stride and is divided into two phases: the stance phase and the swing phase. The stance phase, which makes up about 62% of the gait cycle, begins with a heel strike and ends with a toe-off of the same foot. The remaining 38% of the gait cycle is the swing phase, during which the foot is not weight-bearing and moves from one step to the next [4]. A smart shoe for the classification of gait patterns has been proposed by Unang *et al.* This research suggests a new state-of-the-art gait classification approach, specifically for detecting human gait abnormalities. Sayali has developed a wearable smart shoe to track health and fitness using IoT [4]. Kimin *et al.* [5] developed an artificial neural network-based abnormal gait pattern classification using smart shoes.

Energy Harvesting-Piezoelectric energy harvesting in smart shoes is a technology that utilizes the piezoelectric effect to convert mechanical energy generated by the user (during walking and/or running) into electrical energy. The electrical energy generated can then be used to power various sensors and other components inside the shoes. This technology has the potential to reduce the reliance on batteries and reduce the need for frequent charging of devices. Power harvesting using piezoelectric shoes for external power storage was designed and developed by Mazalan *et al*. This paper aims to harvest energy by designing a shoe embedded with a piezoelectric disc element [6].

By reviewing the literature it is understood that there is a lack of research on the integration of different sensors and technologies in a single wearable prototype for conveniently monitoring individual health. Therefore, this research aims to develop a smart shoe prototype that can accurately monitor health.

3. METHODOLOGY

Data was collected from a group of eight healthy subjects which includes 5 males and 3 females whose having ages between 23 years to 25 years and heights around 5 feet and 5 inches. It is important to select subjects with approximately the same age and same height since both age and height have a high impact on the stride length [7]. Subjects were instructed to wear the smart shoe and walk 50 m flat tack. A pilot study was conducted to give experience with the smart shoe and the way how to operate it before the exact experiments.

The smart shoe as shown in Figure 1 is equipped with an MPU6050 sensor module a 6-axis motion tracking device that combines a 3-axis gyroscope, 3-axis accelerometer, digital motion processor, and temperature sensor all in one small package. The Arduino Uno is a microcontroller board used as the processor module which is based on the Microchip ATmega328P microcontroller and is open-source.

A Piezoelectric sensor is a device that detects changes in pressure, acceleration, and force and translates them into an electrical signal using the piezoelectric effect. A 3.7 V 1800mAh LiPO battery is used. The primary code was developed using Arduino integrated development environment (IDE) software and MIT App Inventor to create an Android app. The HC-05 is a device that uses serial communication to establish short-range wireless connections with the mobile phone.

Ultimately, the best position for each sensor would be determined based on the desired measurement, sensor specifications, and the overall design of the shoe. MPU6050 is placed in the sole of the shoe to measure foot motion and gait. Gyroscopes can be used to track the orientation of the foot and is placed in midsole. Piezoelectric sensors are placed in the insole of the shoe.

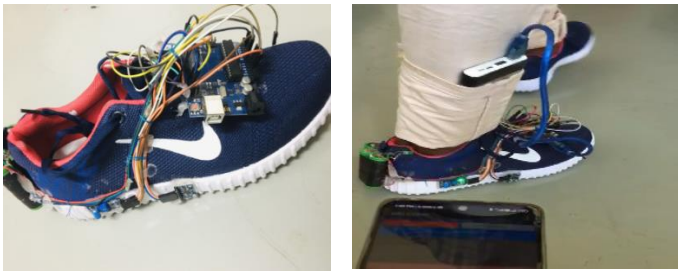


Figure 1. Prototype of the smart shoe

3.1 Fall Detection

Accelerometer and gyroscope sensors are used to get the measurements related to fall detection. The MPU6050 module contains 16-bit raw data that is represented in 2's complement form. Performing a 2's complement operation is necessary to obtain the raw data from the accelerometer and gyroscope sensors. After dividing the sensor raw data by their respective sensitivity scale factors, the acceleration (A_x, A_y, A_z) and angular velocity (W_x, W_y, W_z) can be calculated.

A fall can be detected when the acceleration and angular velocity values exceed a certain critical threshold. In this research, acceleration and angular velocity are calculated based on Eq. (1) [3] and Eq. (2) [3]. The total sum acceleration vector is calculated based on A_x, A_y, and A_z.

$$A = \sqrt{(A_X)^2 + (A_y)^2 + (A_z)^2} \quad (1)$$

Similar to acceleration, angular velocity can be calculated from sampled data.

$$W = \sqrt{(W_X)^2 + (W_y)^2 + (W_z)^2} \quad (2)$$

3.2 Gait Analysis

First, the subjects were asked to walk for 50 meters in a straight line at a normal speed and this was used as the ground truth dataset. Second, they were asked to purposely change their normal speed and walk on the same track of 50 meters. Next Measured gait parameters values were validated against the standard values of the same for a group of subjects. For that, the calculated measurements obtained from the second experiment were compared with the ground truth values obtained from the first experiment to identify the abnormalities of the gait patterns. This will allow us to identify the gait malfunction of the second experiment when compared with the first experiment.

The output signals of the accelerometer and gyroscope are sent to the app developed via the HC-05 Bluetooth module. Obtained data was stored as a text file in the smartphone app. Then this file was imported to MATLAB which is the software that used to implement the gait abnormality detection algorithm is MATLAB. The gait detection algorithm [8] is used to extract gait parameters and it has 4 steps as indicated in Figure 2.

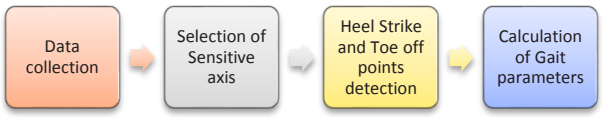


Figure 2. Process of gait analysis

An MPU6050 sensor is used to collect the angular rate about the x, y, and z-axes of the sensor frame. It is usually recommended to use the Y-axis (in the sagittal plane direction) of the angular rate signals to detect the heel strike and toe-off points when calculating gait parameters. The heel strike is characterized by a negative angle reaching its highest values at the initiation of the stance phase. Following this, the toe-off is characterized by a maximum positive angle. By detecting the heel strike and toe-off, the maximum and minimum peaks of the right and left shank signals are then identified. Once the toe-off and heel strike points of the right foot are found for each gait cycle, the gait cycle can be categorized into two phases: the stance phase and the swing phase. The stance phase is initialized at the heel strike, while the swing phase is initiated at the toe-off of the same foot. Gait parameters like time duration of stride, stance, swing, right step also stride length and step length [9] are obtained and then calculate the total number of strides within a particular time.

Out of the gait parameters, stride duration is the time between the heel strick of stride *i* and heel strike of strike *i+1* as shown in Eq. (3)

$$Stride(i) = Heel\ Strike(i+1) - Heal\ Strike(i) \quad (3)$$

3.3 Energy Harvesting

The energy harvest system is designed to generate voltage using human force, specifically footstep stress. In this research, energy harvesting is accomplished by utilizing eight piezoelectric sensors with a quartz crystal material. Piezoelectric sensors convert the gait force into electrical energy, which is then rectified into DC using a full-bridge rectifier. A KBP307 bridge rectifier is used along with the piezoelectric sensor as shown in Figure 3. The output voltage is stored in a LiPo battery. The pressure exerted on the piezoelectric accoutrements during this time generated electrical energy, which was stored in the battery. The battery's charge was successfully tested using a multimeter. The piezoelectric sensors are placed insole as shown in Figure 4.

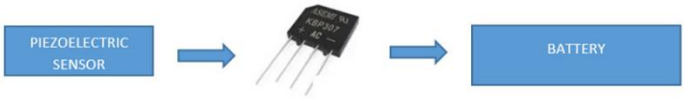


Figure 3. Energy harvesting circuit

Piezoelectric sensors are placed inside the sole as shown in Figure 4.

3.4 Smartphone Application

Smartphones are playing major roles in human health monitoring. Many apps are already invented for individual health monitoring. There are different types of app development tools such as android

studio, MIT App Inventor, etc. The current implementation of the smartphone application has been done on the MIT App Inventor [10]. This smartphone application communicates with a smart shoe through Bluetooth. This app is capable of showing the results from the sensors attached to the shoe as shown in Figure 5. The sensor values are written in a text file and stored on the smartphone's storage for future analysis.

4. RESULTS AND DISCUSSION

Gait events are calculated from the gait data of healthy volunteers and the measured gait parameters are compared to the standard values in [7]. MPU6050 sensors are attached to the shoe and used to obtain two different series of samples HS and TO. Using these points, gait parameters such as the time duration of stride, stance, swing, right step, stride length, and step length are calculated. Next, the total number of strides within a set time is also calculated. Following results are obtained as indicated in Table.1.

These values indicated in the table act as the thresholds and if the stride length is more than this it indicates an abnormal gait. If the subjects run or walk fast than their normal rate the stride length seems to increase. Increase stride length cause a risk of falls and injuries.

Table 1. Average stride of women and men

	Women	Men
Average Stride Length (inches)	51.83	53.02

Voltage and current generation corresponding to three different physical conditioning exercises are conducted with each subject over 15 seconds as shown in Table.2.

Table 2. Piezoelectric sensor readings in multimeter

Types of physical activity	Average voltage(V)	Average current (μA)
Walking	1.368	0.396
Jogging	3.624	1.272
Jumping	4.146	1.466



Figure 4. Placement of piezoelectric sensors within the sole

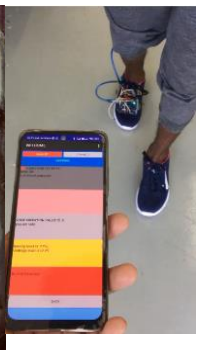


Figure 5. Smartphone application



Figure 6. Harvested energy stored in the batteries

LED lit up from the current drawn from harvested energy as shown in the Figure 6.

5. CONCLUSION

This research has proposed a novel approach to monitoring an individual's health using smart shoes. By analyzing the gait pattern of the user, the system can detect any abnormality in their walking and alert the user or their caretaker. This prototype can help in preventing falls, which is a major concern among the elderly or people with mobility issues. Moreover, the proposed approach was able to generate power through the force that emerged on shoe soles. The proposed smart shoe could be further enhanced to include more sensors to collect more data to provide a more comprehensive view of an individual's health. Further, it is expected to integrate the prototype with other wearable devices, such as smart watches, to provide a more seamless experience for the user. The ultimate goal of this research is to implement the circuit with sensors which will be able to fit into any shoe by making the components as small as possible by utilizing a wireless connection between them.

6. REFERENCES

[1] S. Ahn, J. Kim, B. Koo and Y. Kim, "Evaluation of Inertial Sensor-Based Pre-Impact Fall," 2019.

[2] S.-H. Liu and W.-C. Cheng, "Fall Detection with the Support Vector Machine during Scripted and Continuous Unscripted Activities," 2012.

[3] Z. T. Al-Dahan, N. K. Bachache and L. Nasseer Bachache, "Design and Implementation of Fall Detection System," in *International Conference on Smart Homes and Health Telematics*, 2016.

[4] U. Sunarya, Y. S. Hariyani, T. Cho, J. Roh, J. Hyeong, I. Sohn, S. Kim and C. Park, "Feature Analysis of Smart Shoe Sensors for," 2020.

[5] K. Jeong and K.-C. Lee, "Artificial Neural Network-Based Abnormal Gait Pattern Classification Using Smart Shoes with a Gyro Sensor," 2022.

[6] M. S. Mazalan, R. Mohamad, M. Kassim and S. Shahbudin, "Power Harvesting Using Piezoelectric Shoe For External Power Storage," 2018.

[7] A. Sayer, "Average Stride Length Statistics: Stride Length By Height And Sex," 17 November 2022. [Online]. Available: https://marathonhandbook.com/average-stride-length/#:~:text=Many%20fitness%20pedometers%20and%20watches,What%20is%20this%3F. [Accessed 20 December 2022].

[8] T. Gujarathi and K. Bhole, "Gait analysis using IMU sensor," in *10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, Kanpur, India, 2019.

[9] M. Boutaayamou, O. Bruls, V. Denoel, C. Schwartz, M. Demonceau, G. Garraux and J. Verly, "Segmentation of gait cycles using foot-mounted 3D accelerometers," in *International Conference on 3D Imaging (IC3D)*, 2015.

[10] "MIT App Inventor," 2022. [Online]. Available: https://appinventor.mit.edu/.

Sri Oryzaugur: A Multivariate Approach to Forecast Rice Prices in Sri Lanka

Luqman Rumaiz
Informatics Institute of Technology
Colombo, Sri Lanka
luqman.rumaiz@gmail.com

Gihan Liyanage
Ascentic Pvt Ltd
Colombo, Sri Lanka
gihan.liyanage94@gmail.com

ABSTRACT

Sri Lanka is heavily dependent on rice as a crucial part of food security, and it is an important part of the Sri Lankan economy. Accurate forecasting of rice prices is vital in order to ensure economic stability. Current forecasting models that use univariate approaches have not been sufficient to capture the complexities of rice price fluctuations. This study introduces Sri Oryzaugur, a multivariate approach based on the Temporal Fusion Transformer (TFT) model, specifically designed for rice price forecasting in Sri Lanka. It takes into account various covariates specific to Sri Lanka, to provide more accurate predictions. Results demonstrate that the use of the TFT model and multivariate approach has vastly improved the accuracy of forecasts, yielding a MAPE of 2.52, MAE of 2.30 and an RMSE of 2.77.

Keywords

Time Series, Multivariate Time Series, Price Forecasting, Data Science, Machine Learning

1. INTRODUCTION

It has been well documented that price instability for food commodities is a major concern, especially in developing countries such as Sri Lanka, where agriculture accounts for 7.4% of the countries' GDP and approximately 30% of the population's workforce is employed in this sector [1]. The rice market is prone to a lot of fluctuations due to factors such as changes in farmgate prices, government policies and so forth. Accurate forecasting of prices is therefore essential to help prevent negative economic situations in the country [2]. Furthermore, according to a study conducted in the Journal of Agricultural Sciences in Sri Lanka, farmers and consumers lack awareness on price variations [3] and need a better ability to make informed decisions.

Research has revealed that a significant proportion of farmers have access to Smartphones and the internet [4]. In fact, Mr. Manoj Thibbotuwawa, the Head of the Agriculture Economic Policy at the Institute of Policy Studies of Sri Lanka (IPS), shows that young to middle-aged males use mobile phones for various purposes [5]. This confirms the possibility of farmers using such methods in Sri Lanka.

There is research done on predicting commodity prices with several machine learning, deep learning, and time series models in various countries. However, considering other factors is redundant in most research and is regarded as future work in some papers. Moreover, this research is dynamic for Sri Lanka as it is more effective at forecasting than using a global model, "the weakness of global models is that they cannot fit each dynamic region very well" [6] therefore, a dynamic approach towards each region is needed. This is particularly true for Sri Lanka, where accurate forecasts become increasingly crucial.

2. RESEARCH GAP

Research into price forecasting of Sri Lankan rice has traditionally been based on univariate models, meaning they only rely on past price history in order to predict future prices. Unfortunately, these models have proven insufficient in considering the myriad of factors that determine price movements in a real-world commodity market [7].

Furthermore, there is a lack of research that utilizes transformer models in the broader domain of commodity price forecasting. This research will also explore the potential of transformer models for commodity price forecasting.

3. METHODOLOGY

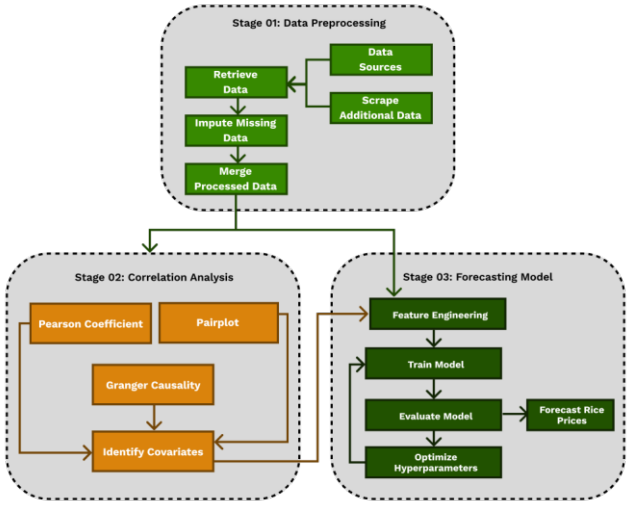


Figure 1: Proposed Architecture Diagram

3.1 Data Preprocessing

Starting off the author had prepared a list of exogenous factors that would be used to forecast the data; therefore, the author had conducted interviews with researchers and also read through literature to understand what factors affect the price of rice. From the compiled list, the author gathered data from government and non-government organizations, namely HARTI, CBSL and CEYPECTO. The historical time series data used in this research was from 1996 January to 2019 December, with a monthly

granularity. The author concluded that all the data beyond 2019 December were outliers, attributed to the unfavorable government policies and an economic crisis that have ensued since then.

Any missing data from the collected data will be imputed using Kalman's Smoothing due to its ability to handle noise and uncertainty. This method is reported to have the lowest error metrics in studies such as [8]. All the gathered data is then merged into a single dataset, ready to be passed to subsequent stages.

3.2 Correlation Analysis

The author put forth a simple framework for examining the association between multiple variables in a dataset, as it is crucial for time series forecasting with exogenous factors [9]. This includes utilizing the Pearson and Spearman correlation coefficients, plotting the data in a pair-plot, and executing Granger's Causality test. This creates a system for carefully analyzing the associations between different variables. The Pearson and Spearman correlation coefficients are commonly used measures of linear and non-linear connections respectively between two continuous variables [10][11]. A pair-plot is a strong visual tool that exhibits the correlation between several variables, while Granger's Causality test is utilized to explore potential causal relationships between time series data [12].

3.3 Temporal Fusion Transformer

Finally, the processed data from the 1st stage and the covariates identified to have a good correlation to forecast the price of rice from the 2nd stage are taken into account during feature engineering. To determine the parameters used to train the model, Optuna for hyperparameter tuning. Forecasts were then generated from the best model in the trials, Table 1 contains the details and parameters of the model.

Table 1: Time Series Dataset and Training Parameter Details

Parameter	Value
TS Dataset Details	
Encoder Length (Tested [60, 100, 276])	60 (Lowest APE)
Batch Size (Tested [32, 64, 128])	64 (Lowest APE)
Lags	[3,6,12]
Prediction Length	12
Network & Training Parameters	
Dropout Rate	60
Learning Rate	64
Gradient Clipping Value	0.8
Attention Head Size	4
Hidden Size	48
Hidden Continuous Size	35

The author chose the Temporal Fusion Transformer (TFT) to model rice prices in Sri Lanka as it offered theoretical soundness, empirical performance, and computational efficiency [13]. TFT combines the strongpoints of Recurrent Neural Networks and the attention mechanism to take into account complex temporal dependencies, seasonal patterns, and any external influences affecting prices. TFT has been observed to produce more precise estimates than traditional time series models in various time series forecasting experiments.

4. RESULTS

4.1 Benchmarking

Table 2: Extensive Model Comparison

Time Series Model	MAPE	MAE	RMSE
TFT	2.52	2.30	2.77
TFT (Univariate)	7.62	6.93	7.57
SARIMAX	2.98	2.80	3.36
ARIMAX	3.93	3.76	4.99
SARIMA	5.35	5.13	6.55
ARIMA	4.07	4.07	6.30

The TFT model has the lowest error metrics among all the tested models, it outperforms the univariate variation of the TFT model and all 04 statistical models. This demonstrates the importance of incorporating multiple factors in the forecasting process, even for the statistical models since the variants of the autoregressive models with exogenous factors outperforms the univariate statistical models.

4.2 Feature and Attention Importance

The feature importance analysis conducted on the Temporal Fusion Transformer (TFT) model has provided valuable insights into the factors influencing rice price forecasting.

4.2.1 Feature Importance

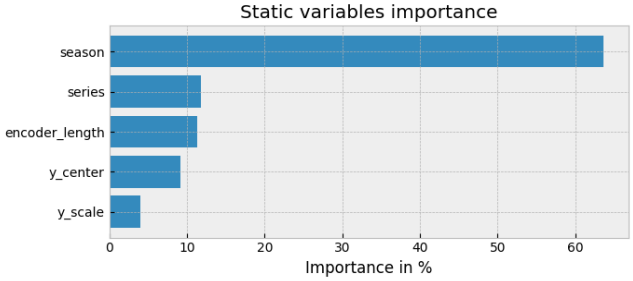


Figure 2: Static Variable Importance

The analysis of the static covariates in the Temporal Fusion Transformer (TFT) model showed that season had the highest importance, accounting for 60% of the total. This emphasizes the strong influence of seasonal patterns on rice price predictions in Sri Lanka. The importance of seasonality suggests that the market dynamics associated with the paddy harvesting seasons should be taken into account for more accurate forecasts. Such findings are invaluable for developing better rice price forecasting models and making more informed decisions in the agricultural sector.

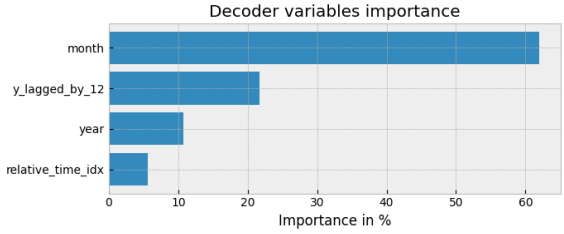


Figure 3: Decoder Variable Importance

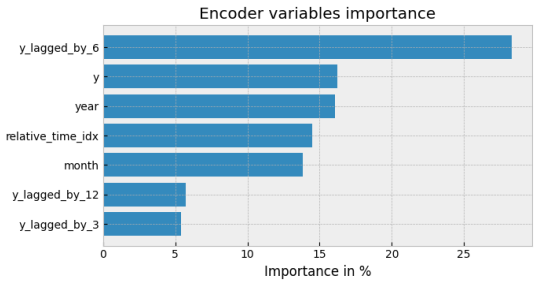


Figure 4: Encoder Variable Importance

The results of the analysis on the encoder and decoder components indicated that lagged prices, seasonal patterns, and temporal indicators play a significant role in predicting future rice prices. Variables such as lagged prices from six-time steps prior, retail price, year, and the sequential order of the timestep were found to have the highest influence in the encoder, demonstrating the importance of utilizing past price trends for accurate forecasts. For the decoder, the month was determined to have the most impact, emphasizing the significance of seasonal patterns in forecasting. Additionally, lagged prices at different time intervals were also found to be important in the model's performance.

These findings can guide stakeholders involved in the rice industry towards a better understanding of the influential factors in rice price forecasting. By taking into account the temporal dynamics and seasonal patterns, and incorporating relevant data such as lagged prices, decision-making processes such as production planning, market analysis, risk management, and policy formulation can be improved and refined.

4.2.2 Attention Importance

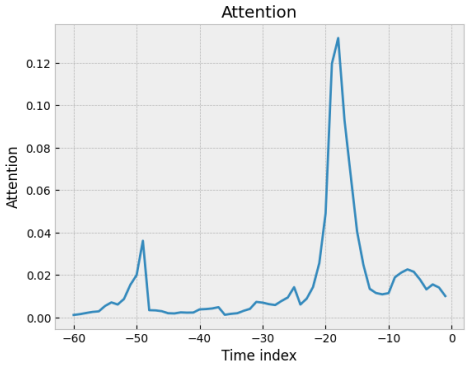


Figure 4: Attention Placed at Specific Timestamps

The TFT model has identified that the information from time steps -50 and -20 are particularly important for accurately predicting future rice prices in Sri Lanka, as evidenced by the model's attention mechanism during its training process. This finding implies that the model can prioritize specific sections of the input sequence and thereby enhance its forecasting accuracy. Subsequently, researchers can leverage this insight to identify feature importance and construct new features that capture relevant temporal information. For instance, lagged variables or derived features that account for the fluctuations observed in the rice market during these critical time steps can be included which should lead to a more accurate representation of the market dynamics and the development of more accurate forecasting models.

5. CONCLUSIONS

The purpose of this research was to develop a robust multivariate forecasting model to forecast rice prices in Sri Lanka. A learning approach in the broader domain of commodity price forecasting with deep learning was taken with the use of various exogenous factors and static covariates, which achieved the lowest error metrics compared to other statistical models. Future work should involve the development of a custom loss to handle periods of unsteadiness in data arising from issues in Sri Lanka, such as an economic crisis and so forth. Government policies and market monopolies must be considered, and analyzing deeper levels of granularity is crucial to identify short-term trends.

6. REFERENCES

[1] International Trade Administration, "Sri Lanka - Agricultural Sector," Sep. 28, 2021. https://www.trade.gov/country-commercial-guides/sri-lanka-agricultural-sector (accessed Oct. 19, 2022).

[2] World Food Programme, "Food crisis in Sri Lanka likely to worsen amid poor agricultural production, price spikes and ongoing economic crisis, FAO and WFP warn | World Food Programme," Sep. 12, 2022. https://www.wfp.org/news/food-crisis-sri-lanka-likely-worsen-amid-poor-agricultural-production-price-spikes-and-ongoing (accessed Oct. 06, 2022).

[3] B. R. P. M. Basnayake, K. D. Kaushalya, R. H. M. Wickaramarathne, M. a. K. Kushan, and N. V. C. Chandrasekara, "An Approach for Prediction of Weekly Prices of Green Chili in Sri Lanka: Application of Artificial Neural Network Techniques," *Journal of Agricultural Sciences – Sri Lanka*, vol. 17, no. 2, Art. no. 2, May 2022, doi: 10.4038/jas.v17i2.9746.

[4] A. Narmilan, G. Niroash, and N. Puvanitha, "Assessment of Current Status on Smart Farming Technologies in Batticaloa District, Sri Lanka," *Sri Lankan Journal of Technology*, p. 7, Dec. 2020.

[5] "Why the Transition to Smart Farming Is Critical in Sri Lanka," *Development Asia*, Mar. 18, 2021. https://development.asia/insight/why-transition-smart-farming-critical-sri-lanka (accessed Nov. 03, 2022).

[6] S.-C. Huang and C.-F. Wu, "Energy Commodity Price Forecasting with Deep Multiple Kernel Learning," *Energies*, vol. 11, no. 11, Art. no. 11, Nov. 2018, doi: 10.3390/en11113029.

[7] C. Bowman and A. M. Husain, "Forecasting Commodity Prices: Futures Versus Judgment," Mar. 2004.

[8] M. Sengupta, S. Bandyopadhyay, A. Habte, and A. Denhard, "A Comparison of Time Series Gap-Filling Methods to Impute Solar Radiation Data," in *Proceedings of the ISES Solar World Congress 2021*, Virtual: International Solar Energy Society, 2021, pp. 1–12. doi: 10.18086/swc.2021.38.03.

[9] L. Batista, "How to Perform Correlation Analysis in Time Series data using R? - Luba," *LOB.DATA*, Sep. 15, 2020. https://www.lobdata.com.br/2020/09/15/how-to-perform-correlation-analysis-in-time-series-data-using-r/ (accessed Apr. 28, 2023).

[10] W. Kenton, "What Is the Pearson Coefficient? Definition, Benefits, and History," *Investopedia*. https://www.investopedia.com/terms/p/pearsoncoefficient.asp (accessed Jun. 22, 2023).

[11] J. Frost, "Spearman's Correlation Explained," *Statistics By Jim*, Mar. 29, 2021. https://statisticsbyjim.com/basics/spearmans-correlation/ (accessed Jun. 22, 2023).

[12] K. Kempf-Leonard, "Encyclopedia of Social Measurement," *ScienceDirect*. http://www.sciencedirect.com:5070/referencework/9780123693983/encyclopedia-of-social-measurement (accessed Jun. 22, 2023).

[13] B. Lim, S. O. Arik, N. Loeff, and T. Pfister, "Temporal Fusion Transformers for Interpretable Multi-horizon Time Series Forecasting." arXiv, Sep. 27, 2020. doi: 10.48550/arXiv.1912.09363

Depression Detection in Sinhala-English Code-Mixed Language using Social Media Data

Fahd Fazal
Department of Computing
Informatics Institute of Technology
Colombo, Sri Lanka
fahad.2019656@iit.ac.lk

Cassim Farook
Department of Computing
Informatics Institute of Technology
Colombo, Sri Lanka
cassim.f@iit.ac.lk

ABSTRACT

The most significant and prevalent mental illness in the world today is depression. Depression is a treatable condition. In Sri Lanka, people would not recognize depression in its earliest stages, when left un-treated contributes to a high rate of suicidal behaviour. Most people tend to share depressive feelings and thoughts intentionally or unintentionally on social media. Social media is a viable source of data for researchers engaged in attempts for early detection of depression. There are numerous studies done in the field of depression detection using English language. But depression detection is less researched using code-mixed languages which is a prevalent trend on most social media users. Therefore, the objective of this research is proposing a machine learning approach to detect depression in Sinhala-English code-mixed textual contents posted in social media. Several base and ensemble machine learning algorithms were tested to detect depression using a dataset collected from social media platforms. Evaluating the performance results, it was found that ExtraTreesClassifier outperformed other machine learning algorithms with the highest accuracy of 79.13%.

Keywords

Depression Detection, Machine Learning, Natural language Processing, Ensemble Classifier, Sinhala-English code-mixed language.

1. INTRODUCTION

1.1 Depression

Depression is a mood disorder that involves a persistent feeling of sadness and loss of interest. It is a serious mental illness and it negatively affects the entire body [1]. Most of the common symptoms of depression are feelings of sadness and irritation, loss of interest in everything, hopelessness, difficulty in sleeping and suicidal thoughts [2]. According to the World Health Organization (WHO), approximately 280 million people suffer from depression. It is equivalent to 3.8% of the total global population which includes 5.0% among adults and 5.7% among adults older than 60 years [2].

In Sri Lanka, 4.1% of total population suffers from depression which includes mostly age group of 20 years and over [3]. During COVID pandemic, a higher number of people were identified with depressive symptoms because of the fear of being infected with COVID-19 [4]. Also, the economic crisis has increased the rate of depression among Sri Lankans [5]. This depicts that depression is a major problem among people all around the world and it can be cured easily when it is diagnosed in the early stage. Therefore, early depression detection has become a necessity to reduce the depression rate in society.

1.2 Sinhala-English Code-Mixed Language

Code mixing is considered as mixing two or more languages in a single sentence or speech. It is widely used by multilingual communities for better communication [6]. Code mixing can be divided into two categories as intra-sentential code-mixing and inter-sentential code-mixing [7]. Intra-sentential code-mixing is when there is no obvious shift in subject within the sentence and inter-sentential code-mixing is the process that occurs when a speaker switches between their native language and a second language to describe an experience [7]. Sri Lankans also use code mixing on social media where they mix native Sinhala with English, and it is alternatively called as Singlish [8]. There are different styles of code mixing used by the Sri Lankan social media users but the most common one is Sinhala words written in English Language (Singlish) [9]. The Table 1 shows different styles of Sinhala-English code mixing.

Code-Mixing Style	Example
Sinhala is written in English (Singlish), with some English terms in the sentence.	Mage phone eka denna
Unicode characters with mix of Singlish	Mage දුරකථනය denna
English words + Unicode characters + Singlish	මගේ phone eka දෙන්න

Table 1: Different Singlish code mixing styles

Since some of the well-known symptoms of depression can be extracted through the words used by the social media users [10], it can be considered that social media networks have developed into viable tools for researching various mental diseases including depression as well [11]. Even though depression detection using English language is a highly researched area, there are only very few studies done in depression detection using Sinhala and other code-mixed languages.

2. METHODOLOGY

As this study is focused on detecting depression in Singlish textual contents using social media data, the authors experimented different base and ensemble ML algorithms to find the best performing algorithm for the specific depression detection task. The methodology consists of main steps such as dataset collection, data pre-processing, feature extraction and classification.

2.1 Data Collection and Annotation

The authors had to manually gather data and construct an own dataset because there is no publicly available dataset for depression detection in Sinhala-English code-mixed language. So, twitter API was used to extract tweets by using a set of keywords which are mostly used by depressive people such as “dukai”, “epa wela”, “marenna hithenawa” and “kalakirila”.

Since tweets were in limited numbers, the authors had to collect Singlish texts from Facebook by using Facebook Scraper which is a python library created based on the Facebook API and by manually searching in psychology groups on Facebook. The authors were able to annotate the collected total number of Singlish texts with the help of a counsellor psychologist. All the tweets and posts collected were publicly accessible and the identity of the user who posted the text was ignored due to privacy issues. The dataset consists of 1177 data records as shown in Table 2. The label column is to label the Singlish texts as non-depressive or depressive. ‘1’ denotes the ‘depressive texts while ‘0’ denotes the non-depressive texts.

Text	Label
mata manasika gataluwak thiyenawa. katha krnna puluwan kenek innawada	1
hithana dewalma wenne natha	0

Table 2: Sample records of the dataset

2.2 Data Pre-processing

Raw data collected from social media should be preprocessed by removing the unnecessary and noisy elements in the data to extract features from it [12]. Different Natural Language Processing (NLP) techniques are used for pre-processing in text classification tasks. Tokenization, stop words removal, and removal of capitalization, punctuation, special characters, numbers, HTML tags, links and emojis were followed as pre-processing steps. Even though stemming and lemmatization are widely used for pre-processing, experts in Singlish domain suggest that it is highly difficulty to perform them on a Singlish dataset. Therefore, stemming and lemmatization were ignored in this study. Moreover, for stop words removal, custom Singlish stop words list was created using Sinhala and English stop words where Sinhala words were translated into Singlish and common variations of them were inserted into the Singlish stop words list.

2.3 Feature Extraction

Feature extraction is an important step to be performed before classification as it involves transforming text data into a numerical representation, known as features, that can be used as inputs to machine learning models. Bag of Words (BoW), word n-grams, character n-grams, Term Frequency and Inverse Document Frequency (TF-IDF) and word embeddings are some of the features used for feature extraction in the previous studies of depression detection domain.

Based on an experiment done by the authors, it was observed that ML algorithms perform well with unigram to five-gram character n-gram features than word n-gram features. These features should be extracted from the dataset and vectorized using a vectorizer. In the researches [13], [3] and [14], it was proved that TF-IDF vectorizer works better than Count vectorizer with ML algorithms. Therefore, TF-IDF vectorizer in the scikit-learn package was used to extract and vectorize character n-gram features in this study.

2.4 Classification

Due to the reason that deep learning approaches requires larger dataset for training, seven base and five ensemble supervised ML algorithms were considered in this study where Logistic Regression, SVM, Multinomial Naïve Bayes, KNN, Decision Tree, SGD Classifier and Passive Aggressive Classifier are the base ML algorithms while Random Forest, Ada Boost, Gradient Boosting, Hard voting classifier and ExtraTreesClassifier are the ensemble ML algorithms. In order to find the best parameters for each algorithm, hyper-parameter tuning was performed using grid search with five-fold cross validation. Grid search was selected over other methods as it checks for all possible combinations of hyper-parameter values. All the algorithms were implemented using Scikit Learn and the results obtained are discussed in the results section.

3. RESULTS

Performance of all the experimented supervised ML algorithms were evaluated using classification metrics such as accuracy, precision, recall and f1-score. Each algorithm was trained with unigram to five-gram character n-gram features using 80% of the created dataset and was tested using 20% of the created dataset. The performance results are shown in Table 3.

Algorithm	Accuracy	Precision	Recall	F1-score
LR	0.74	0.71	0.58	0.64
SVM	0.75	0.73	0.60	0.66
MNB	0.73	0.66	0.67	0.66
DT	0.57	0.46	0.48	0.47
Random Forest	0.73	0.68	0.59	0.63
KNN	0.66	0.57	0.67	0.61
SGD Classifier	0.66	0.60	0.49	0.54
Passive Aggressive Classifier	0.71	0.63	0.68	0.65
Ada Boost	0.68	0.59	0.63	0.61
Gradient Boosting	0.70	0.63	0.57	0.60
ExtraTrees Classifier	**0.79**	**0.77**	**0.65**	**0.71**
Hard voting classifier	0.74	0.72	0.59	0.65

Table 3: Performance of the ML algorithms

According to the results, it can be observed that the implemented ExtraTrees Classifier algorithm with character n-gram features performs better than other ML algorithms used as it has achieved higher accuracy of 79.13% and f1-score of 0.71 out of all.

As a result, the main contribution of this research is that the proposed machine learning approach can be used in other code-mixed data-based depression detection works as it performs well. Moreover, the dataset created will help other researchers to research more in this area and the custom stop words list introduced can be utilized in other research works in Singlish domain as well.

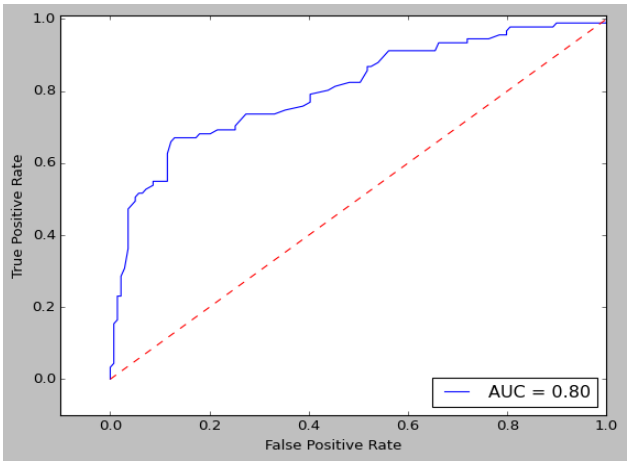


Figure 1: AUC-ROC curve of ExtraTreesClassifier

In general, a larger AUC value implies a stronger classifier, with a classifier that has an AUC close to 1.0 having good discriminating power and a classifier that has an AUC near to 0.5 having weak discrimination power [15]. According to Figure 1, since the AUC value is 0.80 which is closer to 1.0, it can be concluded that the ExtraTrees Classifier algorithm performs well in classifying Singlish text into depressive or non-depressive classes.

4. CONCLUSION

This study has proposed a machine learning approach for detecting depression in Sinhala-English code-mixed language. A dataset was created by gathering tweets and posts from Twitter and Facebook which was annotated with the help of a counsellor psychologist. In data pre-processing, a custom Singlish stop words list was created using Sinhala and English stop words for stop words removal task. The implemented ExtraTrees classifier outperformed all the other classifiers with an accuracy of 79.13% and f1-score of 0.71. Unigram to five-gram character n-gram features performed well with ML algorithms in depression detection in Sinhala-English code-mixed language.

As future work, size of the dataset could be increased to improve the performance of the ML model. Since only machine learning approaches were considered in this study, Deep Learning approaches could be explored with a larger dataset. Different elements such as emojis and time of the posts posted could be considered for better depression detection. Moreover, the proposed approach could be upgraded into a chatbot and classify the detected depression into different levels which can vary from low to severe.

5. REFERENCES

[1] Felix Torres. 2020. What Is Depression? *psychiatry.org*. Retrieved October 4, 2022 from https://psychiatry.org:443/patients-families/depression/what-is-depression

[2] 2021. Depression. *World Health Organization*. Retrieved October 4, 2022 from https://www.who.int/news-room/fact-sheets/detail/depression

[3] Lashini Rathnayake and Isuri Anuradha Nanomi Arachchige. 2021. Supervised Learning Approach for Detection of Sinhala Depressive Posts based on Twitter. In *2021 21st International Conference on Advances in ICT for Emerging Regions (ICter)*, 111–116. DOI:https://doi.org/10.1109/ICter53630.2021.9774819

[4] Bilesha Perera, Bimba Wickramarachchi, Champika Samanmalie, and Manjula Hettiarachchi. 2021. Psychological experiences of healthcare professionals in Sri Lanka during COVID-19. *BMC Psychology* 9, 1 (March 2021), 49. DOI:https://doi.org/10.1186/s40359-021-00526-5

[5] The Lancet Regional Health – Southeast Asia. 2022. Sri Lanka - Health in the middle of a crisis. *The Lancet Regional Health - Southeast Asia* 4, (September 2022), 100067. DOI:https://doi.org/10.1016/j.lansea.2022.100067

[6] D R Mabule. 2015. What is this? Is It Code Switching, Code Mixing or Language Alternating? *JESR* (January 2015). DOI:https://doi.org/10.5901/jesr.2015.v5n1p339

[7] Kasthuri Shanmugalingam and Sagara Sumathipala. 2019. Language identification at word level in Sinhala-English code-mixed social media text. In *2019 International Research Conference on Smart Computing and Systems Engineering (SCSE)*, 113–118. DOI:https://doi.org/10.23919/SCSE.2019.8842795

[8] Lahiru de Silva and Supunmali Ahangama. 2021. Singlish to Sinhala Transliteration using Rule-based Approach. In *2021 IEEE 16th International Conference on Industrial and Information Systems (ICIIS)*, 162–167. DOI:https://doi.org/10.1109/ICIIS53135.2021.9660744

[9] Ian Smith and Uthayasanker Thayasivam. 2019. Sinhala-English Code-Mixed Data Analysis: A Review on Data Collection Process. In *2019 19th International Conference on Advances in ICT for Emerging Regions (ICTer)*, 1–6. DOI:https://doi.org/10.1109/ICTer48817.2019.9023739

[10] Rodrigo Martínez-Castaño, Juan C. Pichel, and David E. Losada. 2020. A Big Data Platform for Real Time Analysis of Signs of Depression in Social Media. *IJERPH* 17, 13 (July 2020), 4752. DOI:https://doi.org/10.3390/ijerph17134752

[11] Esteban A. Ríssola, Seyed Ali Bahrainian, and Fabio Crestani. 2020. A Dataset for Research on Depression in Social Media. In *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (UMAP '20), Association for Computing Machinery, New York, NY, USA, 338–342. DOI:https://doi.org/10.1145/3340631.3394879

[12] Shaik Rasheeda Begum and Saad Yunus Sait. 2022. Effective Techniques for Depression Detection on Social Media: A Comprehensive Review. In *2022 International Conference on Computer Communication and Informatics (ICCCI)*, 1–9. DOI:https://doi.org/10.1109/ICCCI54379.2022.9740754

[13] Carmel Mary Belinda M J, Ravikumar S, Muhammad Arif, Dhilip Kumar V, Antony Kumar K, and Arulkumaran G. 2022. Linguistic Analysis of Hindi-English Mixed Tweets for Depression Detection. *Journal of Mathematics* 2022, (April 2022), 1–7. DOI:https://doi.org/10.1155/2022/3225920

[14] Oshadhi Liyanage and Krishnakripa Jayakumar. 2021. Hate Speech Detection in Sinhala-English Code-Mixed Language. In *2021 21st International Conference on Advances in ICT for Emerging Regions (ICter)*, 225–230. DOI:https://doi.org/10.1109/ICter53630.2021.9774816

[15] Jorge M. Lobo, Alberto Jiménez-Valverde, and Raimundo Real. 2008. AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography* 17, 2 (2008), 145–151. DOI:https://doi.org/10.1111/j.1466-8238.2007.00358.x

iCIIT 2023

THE 4TH INTERNATIONAL CONFERENCE ON
INNOVATIONS IN INFO-BUSINESS & TECHNOLOGY

EXTENDED ABSTRACTS

COMPUTING



Transfer Learning-based Approach to Detect Onychomycosis Fungal State & Fungal Growth on the Nail

Punsara Bandara
Department of Computing
Informatics Institute of Technology
Colombo, Sri Lanka
punsara.2018634@iit.ac.lk

Sachinthani Perera
Department of Computing
Informatics Institute of Technology
Colombo, Sri Lanka
sachinthani.p@iit.ac.lk

ABSTRACT

Onychomycosis, a fungal nail infection, poses significant dangers to diabetic patients. Patient outcomes are at risk, necessitating prioritized research for advanced diagnostic and tracking methods for Onychomycosis. In this research project, a light-weighted pre-trained model was employed, utilizing transfer learning techniques and hyper-parameter optimization. The models were trained on various real-world datasets to maximize performance. The author also employed hyper-parameter tuning to find the best parameter combination for the specific use case. From this approach, the author was able to gain 97.53% of Accuracy along with a 0.9804 F1 score, 0.9821 Sensitivity value, 0.9690 Precision value, and 0.9921 for the AUC value from the trained InceptionV3 model along with the self-taught transfer learning. Also, for the fungal segmentation process, the author gained a 0.96 F1 score, a 0.97 Recall, a 0.94 Precision for the segmented boxes, and a 0.83 for the F1 score, Recall, and Precision for the segmented Masks. This Image Processing approach can make significant improvements in dermatology.

Keywords

Machine Learning, Image Classification, Transfer Learning, Fungal Detection, Onychomycosis, Dermatology

1. INTRODUCTION

1.1 Sense Onychomycosis Through Image Processing

Toenail fungus, also referred to as Onychomycosis is a fungus that infects the nails on the toes and fingers. It may have a variety of effects on a person's life. Some common effects of Onychomycosis can be profound on a person's physical, social, and mental health. Through the use of image processing techniques on pictures of damaged nails, Onychomycosis can be discovered. Applying a pre-trained algorithm as the foundation for another task is the basic idea behind the machine learning technique known as transfer learning. Utilizing a pre-trained model that has been trained from a large dataset and applying it to a new task is the idea behind transfer learning. The new model can learn more quickly and effectively this way with its own smaller dataset [2]. For the conducting research project, the author also has to follow Inductive transfer learning since the chosen dataset consists of data classes that are not used to train the most of classification models. As follows the author will conduct the research with Sequential transfer learning also known as self-taught transfer learning.

1.2 Selection of the Dataset

In order to start the implementation, the author needed a datasets with pathologist approved Onychomycosis infected toenail image dataset and a healthy toenail image dataset. The author has contacted a researcher who has conducted research [4] on the same medical domain and requested the datasets can be used in the implementation. All the credits go to the mentioned researcher respectively. After following relevant techniques such as data cleaning, data augmentation, and data separation the author has granted himself the dataset needed for the classification process. For the segmentation process, the author needed to annotate the data set with annotated full nail area and annotated fungal spread area. Because of the unavailability of a such dataset, the author had to create a dataset on his own. The author has annotated data and sample data is shown in Figure 1.

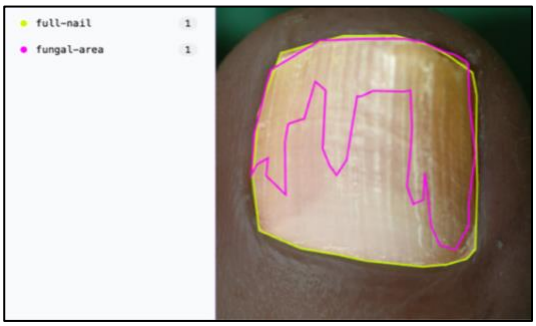


Figure 1 : Annotated Data Sample

1.3 Selection of the Pre-Trained Classification Model

The author needed to find the best-performing pre-trained model for his dataset. In order to achieve that goal, the author selected the best-performing image classification algorithms based on previous research outcomes. All the credit goes to previous researchers respectively. While selecting the best-performing pre-trained model for the dataset, the author considered the following factors on the given matrix values. *The minimum range between Accuracy and Validation accuracy, Gaining higher values for F1 score, recall, and precision, Weights of the well-performed models on the selected dataset.*

Considering the above factors, the author finds out that VGG19 and InceptionV3 are performing well than other pre-trained models. Considering the light weightiness of InceptionV3 than the VGG19, the author decides to conduct the transfer learning with InceptionV3 for conducting the model implementation along with transfer learning because it's got around 0.013 of a minimum range between accuracy and validation accuracy along with more than 0.9 values for F1 score, recall, and precision.

Matrix	Pre-trained Model			
	ResNet 152	VGG 16	Inception V3	VGG 19
Accuracy	0.7989	0.9554	0.9493	0.9468
Val Accuracy	0.7635	0.9261	0.9360	0.9261
Recall	0.97	0.97	0.91	0.94
Precision	0.80	0.96	0.93	0.98
F1 Score	0.88	0.96	0.92	0.96

Table 1 : Pre-trained Classification Model Performance on the Dataset

The way the author's dataset behaves on each selected algorithm is presented in Table 1. The author used the base settings of the model and apply the hyperparameter auto-tuning for those settings. The author set up the auto-tuning for 10 trials and record each hyperparameter combination and their output in order to find out the best combination. Considering the gap between the accuracy and validation accuracy the author has chosen the hyperparameter combination with 352 dense units and a 0.01 learning rate. Even if the accuracy value is high but the relevant validation accuracy is low, the author ignored that combination because of the tendency of overfitting. Table 2 will represent some of the best-trained outputs and their combinations.

Trail No.	Parameter Settings		Accuracy	Validation Accuracy
	Dense Layer Units	Learning Rate		
3	224	0.01	0.9550	0.9802
5	384	0.01	0.9920	0.9778
6	352	0.01	0.9697	0.9753
7	224	0.0001	0.9729	0.9576
9	128	0.001	0.9822	0.9678

Table 2 : InceptionV3 Hyperparameter Settings & Gained Results

1.4 Segmentation Process

The author trained the created annotated dataset with YoloV8 which is one of the latest object detection algorithms, in order to predict the fungal spread percentage on the uploaded nail image. In order to achieve better performance and more accurate results on segmentation model training the author had to make some changes as following. *Increasing the number of annotated train datasets, and Hyperparameter fine-tuning.*

Version	Changes	Gained results			
		mAP50 on Boxes		mAP50 on Masks	
		Full nail	Fungal area	Full nail	Fungal area
SV1	Base settings	~ 0.31		~ 0.29	
SV2	Increased the dataset	0.974	0.896	0.67	0.815
After applying auto-tuning for hyperparameters					
SV3	Learning rate = 0.000441 Epochs = 150	0.93	0.897	0.683	0.819

Table 3 : YoloV8 Segmentation Model Changes & Gained Results

1.5 Proposed Solution Architecture

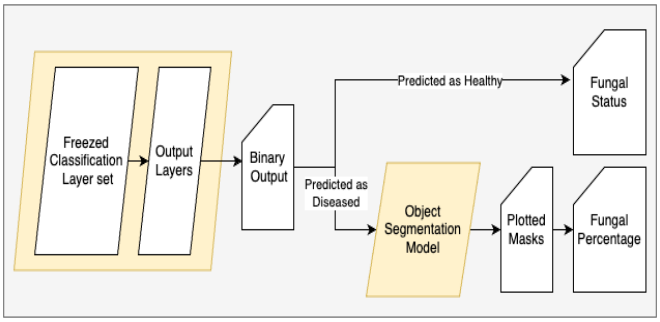


Figure 3 : Proposed Solution Architecture

Based on the output of the classification model if the nail is identified as a fungal-infected nail, the segmentation model will calculate the fungal spread by masking the fungal area. The system output will represent the onychomycosis state and the fungal spread percentage on the nail. Figure 2 is presented the proposed solution architecture.

2. MODEL RESULTS

From the conducted research the author gained the following confusion matrix values from the self-taught transfer learned classification model. Test data is a randomly divided part from the selected dataset.

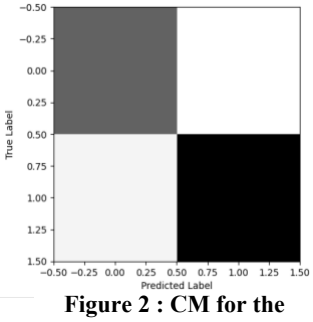


Figure 2 : CM for the Classification Model

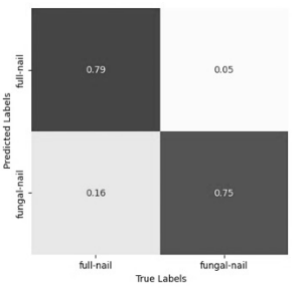


Figure 4 : CM for the Segmentation Model

	True Positive	False Positive	False Negative	True Negative
Classification Model	73	4	9	117
Segmentation Model	79	5	16	75

Table 4 : Confusion Matrix Values from Classification and Segmentation Models

The following evaluation matrix values are gained from the classification model. 0.9753 in accuracy, 0.9690 in precision,

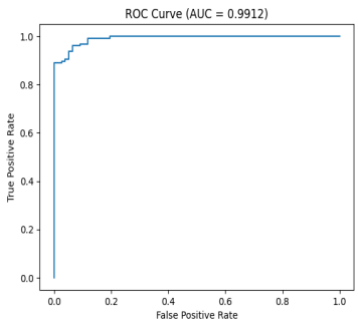


Figure 5 : AUC & ROC curve for the Trained Classification Model

0.9821 in sensitivity/recall, and 0.9804 in F1 score. Also, ACU (Area Under the Curve) got a value of 0.992 gained from the classification model as Figures 4 and 5 are presented. According to the confusion matrix gained following evaluation matrix values are gained from the segmentation model. 0.94 and 0.83 precision values for boxes and masks, 0.97 and 0.83 recall/sensitivity values for boxes and masks, 0.94 and 0.83 precision values for boxes and masks, and 0.96 and 0.83 F1 score values for boxes and masks.

3. CONCLUSION

The author conducted the benchmarking along with the following matrixes. *Accuracy, F1 score, Sensitivity/Recall, AUC (Area Under the ROC Curve).* All the matrix values for existing work are referred from their published research papers respectively. The previous approaches gained the benchmarking matrixes with the following datasets as listed below.

Deep CNN Approach [5] – 160 microscopic images and 297 microscopic field photographs are used to train the model.

Ensemble-based Approach [4] – 49,567 of collected onychomycosis-infected data from different university laboratories is used to train the ensemble models as follows as VGG19 and ResNet152. VGG19 is tested with 194 images and ResNet152 was tested with 100 images.

U-Net-based Approach [6] – 664 histologic whole slide image dataset from different laboratories is used to train the model.

VGG13-based Approach [7] – trained with 528 onychomycosis-infected data images.

Onyno / Author's Approach – 1512 onychomycosis diseased data and 1301 healthy nail data are used to train the model and 203 data were used to test and another 203 data were used to validate the model.

All the gained and existing results can be found in Table 5. The Accuracy and the Sensitivity/Recall matrix gained higher values than the existing approach values. But increasing the dataset will make an impact on existing results. Overall, the conducted research and implementation can have a significant impact in detecting Onychomycosis fungal infection. Taking an approach to improve the strategy of preventing the *Onychomycosis* that gives mentioned complications is worth spending time on.

Approach		Acc.	F1 sc.	Recall	AUC
Deep CNN based	VGG 19	-	-	0.95	0.98
	ResNet 152	-	-	0.947	0.98
Ensemble based	VGG 16	0.9598	0.9599	0.9598	0.9930
	Inception V3	0.9590	0.9850	0.9550	0.9917
Clinician		0.7280	0.7469	0.6100	0.8700
U-Net based		0.8649	-	0.9400	0.9810
VGG13 based		-	-	0.940	0.960
Onyno		*0.9753*	*0.9804*	*0.9821*	*0.9912*

Table 5 : Benchmarking of the Classification Model

4. LIMITATIONS

Since the unavailability of the related datasets, the author had to limit the testing and evaluation process to only the divided test data component from the used dataset. By increasing annotated number of datasets, the segmentation process to detect fungal growth on the nail can be made more accurate and efficient. Chosen classification models are based on the previously conducted research. Using a newly introduced classification model will make a different impact. The segmentation model only detects the fungal spread only on the nail. Not the insider growth of the fungal. Further approaches can be taken by considering nail color to predict fungal insider growth.

5. ACKNOWLEDGMENTS

I am incredibly grateful to my supervisor for her constant leadership, direction, and knowledge, all of which have helped this research to be successful. In addition, I appreciate all the academic professors for their knowledge and guidance. The gracious contribution of the previous researcher, which provided the author with access to the large dataset, was crucial in determining noteworthy discoveries. I also want to thank the data scientists and medical domain experts for their continual guidance and support; their expertise opened the path for us and ensured the accuracy and relevance of the research findings.

6. REFERENCES

[1] "Detection of onychomycosis using convolutional neural networks," by C. Y. Chen, K. H. Huang, and Y. C. Chen, published in IEEE Journal of Biomedical and Health Informatics, 2020.

[2] Pan, S.J. and Yang, Q (no date) A survey on Transfer Learning | IEEE Journals & Magazine | IEEE Xplore, IEEE Xplore. Available at: https://ieeexplore.ieee.org/document/5288526

[3] Bhavsar, P. (2021) Transfer learning in NLP, Medium. Modern NLP. Available at: https://medium.com/modern-nlp/transfer-learning-in-nlp-f5035cc3f62f

[4] Han, S.S. et al. (2018) "Deep neural networks show an equivalent and often superior performance to dermatologists in onychomycosis diagnosis: Automatic construction of onychomycosis datasets by region-based Convolutional Deep Neural Network," PLOS ONE, 13(1). Available at: https://doi.org/10.1371/journal.pone.0191493

[5] Yilmaz, A., Göktay, F., Varol, R., Gencoglan, G., & Uvet, H. (2022). Deep convolutional neural networks for onychomycosis detection using microscopic images with KOH examination. Mycoses. doi 10.1111/myc.13498

[6] Jansen, P., Creosteanu, A., Matyas, V., Dilling, A., Pina, A., & Saggini, A. et al. (2022). Deep Learning Assisted Diagnosis of Onychomycosis on Whole-Slide Images. Journal Of Fungi, 8(9), 912. doi:10.3390/jof8090912

[7] Decroos, F., Springenberg, S., Lang, T., Päpper, M., Zapf, A., & Metze, D. et al. (2021). A Deep Learning Approach for Histopathological Diagnosis of Onychomycosis: Not Inferior to Analogue Diagnosis by Histopathologists. Acta Dermato Venereologica, 101(8), adv00532. doi: 10.2340/00015555-3893

AUTOMATED DECEPTION DETECTION SYSTEM USING HUMAN FACIAL CUES AND FACIAL CHIRALITY

Chamudi Gayanika Vidanagama
Department of Computing
Informatics Institute of Technology
Colombo, Sri Lanka
chamudi.vidanagama@gmail.com

Abarnah Kirupananda
Department of Computing
Informatics Institute of Technology
Colombo, Sri Lanka
abarnah.k@iit.ac.lk

***Abstract*—Deception detection is an important task carried out in high skate situations like suspect interrogations. Criminal investigations are solely dependent on a suspect's truthful confession for a case to be solved. The only way to make a suspect confess is by conducting a suspect interrogation with evidence. Detectives face difficulties in detecting deception manually due to various reasons which have made the process hard. Therefore, researchers have identified different automated deception detection approaches. These approaches have many limitations which leads to inaccurate and invalid results. There is a necessity of an accurate, valid and high performing automated deception detection system. For this reason, Caught in a Lie system has been introduced. The system takes nonverbal indicators of humans such as facial cues and facial chirality into account and detect deception of interrogation videos. To the best of the author's knowledge, a Facial Chirality deep learning model has not been used for any automated deception detection systems up to this point. The proposed system detects deception of interrogation videos accurately and efficiently by outputting the deceptive and truthful percentage of the suspect as well as which indicators caused the deceptiveness or truthfulness of the suspect. Therefore, the Caught in a Lie system is a novel automated deception detection system which is the first to focus on facial chirality when detecting deception. This facial chirality model will assist future researchers of the domain or different domains which require facial chirality**

Keywords—Deception Detection, Facial Chirality, Facial Cues, Interrogations, Non-Verbal Indicators, Criminal Investigations, Deep Learning Model

1. INTRODUCTION

Deception detection task is executed in the interrogation process to identify the truthfulness and deceptiveness of a suspect. According to [1] 'Deception detection refers to the investigative practices used to determine a person's truthfulness and credibility'. Fight or flight, cognitive complexity and perception management are three different deception detection approaches [2]. These approaches are done by analysing the key indicators of deception which are verbal, non-verbal, and non-linguistic communication of the subject. Identifying what exactly truth or deceit is associated with psychological background and body language. In the investigation process, a lie is a statement intended to make people believe something that the speaker does not believe to be true, or not the whole truth [3]. There is an assumption that liars will exhibit stress-based cues due to the fear of being caught and guilty feeling about lying. Criminal investigators should identify the deception of a criminal when carrying out a criminal investigation in order to bring justice to the innocent and punish offenders [4]. Currently the deception detection process is done manually in most countries including Sri Lanka. Manual deception detection has major limitations and remains a difficult task.

- Task requires excessive analysis time
- Lack of highly trained and skilled detectives specialized in deception detection
- Judgements can be human biased
- Poor classification of deception

In order to get better and accurate results from the manual deception detection in interrogation, at least two skilled detectives who are specialized in human deception detection, are needed to interrogate a single suspect. The accuracy level of manual deception detection has the average of 54% [5] which is only slightly above from the decision of a coin toss. Because of these limitations, deception detection hasn't been carried out well in the interrogation process which leads to inaccurate end results.

As a solution for this problem researchers have identified automated deception detection systems. These automated deception detection systems can be categorized into three approaches which are,

1. Nonverbal indicator analysis approach
2. Physiological phenomena analysis approach
3. Multimodal approach

The author will be focusing on the facial and behavioral analysis approach in this research. These facial and behavioral analysis approaches are based on different aspects of human behavior. Some of them are human facial expression analysis, facial action unit analysis, blink rate analysis, emotion analysis, eye gaze analysis, body posture analysis. Mostly image processing and video analysis technical approaches are used to analyze the above mentioned.

2. METHODOLOGY

System consists of three main tiers which can be identified as Data Tier, Logic Tier and Client Tier. Data Tier and the Client Tier are connected by the Logic Tier which is in between. Client tier includes the end-user interacting system. Data Tier contains the required data for the system.

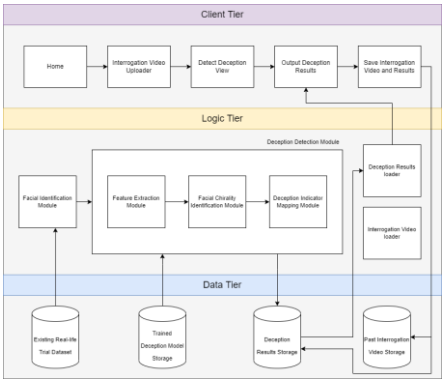


Fig. 1. Three Tier Architecture Diagram

2.1 Dataset

Since the Caught in a Lie system detects deception in interrogation videos, the data can be identified as videos. On the subject of deception detection, system identifies the human facial cues and facial chirality to detect deception in interrogation videos. Therefore, a collection of deceptive and truthful videos was required for the training and testing purposes of the model. Real-Life Trial dataset which consists of video clips of public court trials. This real-life trial dataset consists of 121 videos which includes 61 deceptive video clips and 60 truthful video clips. The average length of the truthful video clips can be identified as 28.3 seconds and the average deceptive clip length can be identified as 27.7 seconds. The dataset also includes the audio data, text data and micro expression data for the deceptive and truthful video clip sets. From these included data, micro expression data was focused when implementing the system Caught in a Lie, as the system mainly focus on the non-verbal cues when detecting deception.

2.2 Modules

- Facial Detection Module was used to identify the human face in order to proceed with the deception detection process. Retina Face which is a deep learning-based cutting-edge facial detection library was used to detect face by identifying facial landmarks.
- Emotion Cue Identification Module was used to identify the 7 universal emotions which are Happy, Sad, Angry, Fear, Contempt, Surprise and Disgust. A pretrained emotion cue recognition model was used in this module.
- Facial Chirality Module was used to identify the chiral action units occurs in the human face. 12 facial action units have been used to train and test this model.

2.3 Pre-Processing

Facial Detection Module and Emotion Cue Identification Module were created using the existing pretrained models. Retina Face library was used for facial detection while a pretrained ResNet18 model was used for emotion cue identification. Facial Chirality module was created by the author by training a ResNet50 model. This model detects the chirality action using pretrained resnet50 model. 12 main action units were used to detect action chirality which are AU01 – inner brow racer, AU02 – outer brow racer, AU03 – brow lower, AU04 – upper lid racer, AU05 – cheek racer, AU06 – nose wrinkle, AU07 – lip corner puller, AU08 – lip conner depressor, AU09 – chin raiser, AU10 – lip stretcher, AU11 – lip part and AU12 – jaw drop.

Author has written a system which takes image frames from the uploaded video and that images were preprocessed to identify the Chirality Action Units. These 12 important chirality action units and how them being true or false have affected the video to be deceptive or truthful with the deception percentage have been considered when training and testing this Facial Chirality Model.

2.4 Confusion Matrix

The confusion matrix of this research includes classes which are identified as follows,

- Positive - Deceptive
- Negative - Non deception
- True Positive - Actually deceptive, detected deceptive
- False Negative - Actually deceptive, detected non deceptive
- False Positive - Actually non deceptive, detected deceptive
- True Negative - Actually non deceptive, detected non deceptive

n=121	Predicted: NO	Predicted: YES
Actual: NO	TN = 53	FP = 7
Actual: YES	FN = 7	TP = 54

Fig 2. Confusion Matrix

3. RESULTS AND DISCUSSION

Real-life trial dataset was used to train the newly created deception detection model Caught-in-a-Lie which specialized in identifying human facial cues and chirality actions using action units. Training the model with well refined data is important for the model to produce finest results. Testing is conducted to test and verify the model has trained properly and produce expected results. Resnet50 convolutional neural network with 50 layers was used for human facial image classification to predict facial cues and chirality action unit. Resnet50 and SVM were together used for the deception detection model.

Accuracy and AUC of the model is depicted in below table.

TABLE I. ACCURACY AND AUC PERCENTAGES

Model	ACC(%)	AUC(%)
Deception Detection	88.6%	92.4%

Accuracy percentage of the deception detection model is approximately 88% which is a good accuracy compared to currently available research. AUC also known as the area under the ROC curve is the probability that model ranks a random positive example. Performance measure across all possible classification entry is provided by AUC [13]. Value of AUC is ranges from 0-1. Deception detection model's AUC percentage is displayed in the above figure. AUC percentage is approximately 91% which is better AUC percentage.

3.1 Competitve Benchmarking

Automated deception detection systems are not commercially available. Therefore, competitive benchmarking was carried out on the research in the area of automated deception detection. Since there are limited number of publicly available datasets for deception detection, most of the researches have been using the same dataset which is the ***Real-life trial dataset***. Below table compares accuracy of competitive research models with **Caught in a Lie** model.

TABLE II. COMPETITIVE BENCHMARKING

Research	Model Algorithms	Accuracy
Automatic Deception Detection in RGB videos using Facial Action Units - Facial Action Units	Logistic Regression	54.86%
	Random Forest	58.42%
	SVM - Linear	65.56%
	SVM - Sigmoid	71.99%
	SVM - RBF	76.84
Deception Detection by 2D-to-3D Face Reconstruction from Videos	Time-CNN	69%
	RNN Model + CNN	73%
Facial Expression Based Imagination Index and a Transfer Learning Approach to Detect Deception	WFM	70%
	SVM	61%
Deception Detection in Videos using the Facial Action Coding S ystem	FACs with LSTM	89.49%
	Decision Tree and Random Forest	75.2%
	SVM-RBF	76.84%
	ADABOOST	88%
LieToMe: An Ensemble Approach for Deception Detection from Facial Cues	Random Forest	83.61%
	L-SVM	91.87%
	Linear Regression	92.01%
	MLP	91.87%
	RBF-SVM	90.75%
Toward End-to-End Deception Detection in Videos	Vocal Model	74.16%
	Visual Model	75%
	Hybrid Model	84.16%
Caught in a Lie	ResNet50 + SVM	88.6%

Benchmarking results illustrates that **Caught in a Lie** system outperforms most of the researches carried out in the area of deception detection as the system has an accuracy of 88%. Therefore, it concludes that the system satisfies the requirement of deception detection well.

3.2 Functional Benchmarking

Functional benchmarking was conducted to compare **Caught in a Lie** model with the manual deception detection. Manual deception detection conducted by humans, only has an accuracy of 52% which is similar to the value achieved from a simple coin toss guess. Since the system model accuracy has a better value of 88% which outperforms the 52% of manual deception detection accuracy, it is proven that the **Caught in a Lie** model is a better system for the deception detection process.

3.3 Limitations of the Project

Identified limitations of the research are as follows:

- Project scope was only focused on non-verbal indicators of human face which are human facial cues and facial chirality. Project could be considering other non-verbal cues such as head posture, body posture and body movements in detecting deception.
- Currently system detects deception in interrogation videos nevertheless system could be implemented for real-time deception detection of suspects while interrogating.
- Model was only trained using real-life trial dataset which is a video dataset of 121 videos of court trials. System could be more effective if it was trained with interrogation videos in Sri Lanka but due to legal issues interrogation videos cannot be exposed to anybody therefore model was trained using publicly available dataset.

3.4 Future Enhancements

Since Sri Lankan Police Department currently does not use automated deception detection systems, **Caught in a Lie** will be the first system to be used. Therefore, author intends to overcome the limitations identified through evaluation, with enhancements in the future. Future enhancements are discussed below,

- Converting the system to a real-time deception detection system will be a major future enhancement.
- Considering non-verbal indicators such as head posture, body posture and body movements which covers flight or fight, and perception management aspects of the suspect will be a necessary future enhancement for better deception results.
- Training the model with interrogation videos in Sri Lanka will provide effective results as the input videos of the system will be similar to training data.

It is better to have facial chirality model results being displayed in the application with the chiral detection of human face which will be clearer for detectives to identify deception and even learn from the system.

4. CONCLUSION

Caught in a Lie research was successfully completed by accomplishing research aim, objectives while fulfilling the research gap and contributing to the body of knowledge. Author was able to present a novel method for deception detection using human facial cues and facial chirality. This proposed method results 88% accuracy which out performs most of the currently available systems for deception detection. Importance of **Caught in a Lie** was identified by conducting evaluations and getting feedback from domain experts. Since the accuracy of the system is high, this automated deception detection system will be a huge impact on criminal investigation in Sri Lanka. It is evident that, system will be a great help for detectives to detect deception in interrogations with less time and effort. Author is well pleased about succeeding the project, **Caught in a Lie** since it contributes to improve one of the busiest domains in Sri Lanka, criminal investigation.

5. ACKNOWLEDGEMENT

First and foremost, I would like to express gratitude to my supervisor, Ms. Abarnah Kirupananda, for

guiding me through the whole project research and encouraging to complete the project on time. Her guidance and encouragement were a great support for the successful completion of the project. I owe many thanks to domain experts from Criminal Investigation Department, Sri Lanka Police for participating in the interviews and providing necessary information for the requirement elicitation and evaluation of the project. I wish to express my appreciation for police officers who provided domain information which assisted throughout the whole project. Besides I would like to thank all the lecturers at Informatics Institute of Technology for supporting and guiding me through academic years which assisted in completing the project. Finally, I would like to give my special thanks to my family for assisting and facilitating me with the necessities for the successful completion of the research.

6. REFERENCES

[1] Paul Ekman Group. 2021. *Deception Detection | How to Tell If Someone is Lying*. [online] Available at: <https://www.paulekman.com/deception/deception-detection/> [Accessed 27 October 2021].

[2] Youtube.com. 2021. *Former FBI Agent Explains Criminal Profiling | Tradecraft | WIRED*. [online] Available at: <https://www.youtube.com/watch?v=LHqbfJ3tq4Y&list=WL&index=13> [Accessed 27 October 2021].

[3] Plato.stanford.edu. 2021. *The Definition of Lying and Deception (Stanford Encyclopedia of Philosophy)*. [online] Available at: <https://plato.stanford.edu/entries/lying-definition/#:~:text=A%20lie%20is%20an%20untruthful,involves%20an%20intention%20to%20deceive.> [Accessed 27 October 2021].

[4] Lakshan, I., Wickramasinghe, L., Disala, S., Chandrasegar, S. and Haddela, P., 2019. Real Time Deception Detection for Criminal Investigation. *2019 National Information Technology Conference (NITC)*,

[5] Burgoon, J., 2015. Deception Detection Accuracy. The International Encyclopedia of Interpersonal Communication, pp.1-6.

[6] Belavadi, V., Zhou, Y., Bakdash, J., Kantarcioglu, M., Krawczyk, D., Nguyen, L., Rakic, J. and Thuriasingham, B., 2020. MultiModal Deception Detection: Accuracy, Applicability and Generalizability. *2020 Second IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications (TPS-ISA)*,.

[7] Jaiswal, M., Tabibu, S. and Bajpai, R., 2016. The Truth and Nothing But the Truth: Multimodal Analysis for Deception Detection. *2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW)*,.

[8] n.d. 2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII). Institute of Electrical and Electronics Engineers.

[9] Khan, W., Crockett, K., O'Shea, J., Hussain, A. and Khan, B., 2021. Deception in the eyes of deceiver: A computer vision and machine learning based automated deception detection. *Expert Systems with Applications*, 169, p.114341.

[10] Avola, D., Cinque, L., Foresti, G. and Pannone, D., 2019. Automatic Deception Detection in RGB videos using Facial Action Units. *Proceedings of the 13th International Conference on Distributed Smart Cameras*,.

[11] Ngô, M., Mandira, B., Yılmaz, S., Heij, W., Karaoglu, S., Bouma, H., Dibeklioglu, H. and Gevers, T., 2021. *Deception Detection by 2D-to-3D Face Reconstruction from Videos*. [online] arXiv.org. Available at: <http://arxiv.org/abs/1812.10558> [Accessed 25 October 2021].

[12] Avola, D., Cascio, M., Cinque, L., Fagioli, A. and Foresti, G., 2020. LieToMe: An Ensemble Approach for Deception Detection from Facial Cues. *International Journal of Neural Systems*, 31(02), p.2050068.

[13] Google Developers. 2022. *Classification: ROC Curve and AUC | Machine Learning Crash Course | Google Developers*. [online] Available at: <https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc#:~:text=AUC%20represents%20the%20probability%20that,has%20an%20AUC%20of%201.0.> [Accessed 24 February 2022].

iCIIT 2023

THE 4TH INTERNATIONAL CONFERENCE ON
INNOVATIONS IN INFO-BUSINESS & TECHNOLOGY

EXTENDED ABSTRACTS

TECHNOLOGY OF BUSINESS



CricPlayerTracker: An Application for identifying cricket talents in Sri Lanka

Gajanen Karunananthan
Department of Business School
Informatics Institute of Technology
Colombo, Sri Lanka
gajanen.2018093@iit.ac.lk

Austen Mascranghe
Department of Business School
Informatics Institute of Technology
Colombo, Sri Lanka
austen@iit.ac.lk

Alqa Husni
Department of Business School
Informatics Institute of Technology
Colombo, Sri Lanka
alqa.h@iit.ac.lk

ABSTRACT

Cricket is a popular sport in Sri Lanka, having a long history and an enthusiastic following. Over the years, the country has produced numerous exceptional cricketers, including some of the world's best players. Cricket is played at all levels in Sri Lanka, from school and club cricket to international competitions. Cricket has a tremendous impact on the culture and economy of the country, with many businesses and sectors dependent on it for revenue and employment possibilities. Despite its popularity, cricket in Sri Lanka suffers a number of obstacles, such as a lack of resources and facilities, unequal access to chances, and questions over the legality of some players' bowling actions. Cricket Player Tracker which is a cricket talent scouting system with legal bowling action prediction in Sri Lanka is a novel system that aims to identify and nurture young cricketing talent in Sri Lanka. The system utilizes machine learning and computer vision techniques to predict the legality of a player's bowling action and store and analyze data over time to track the progress of individual players. The system's objective to identify a player's bowling action and help create a fair and transparent system for talent identification and selection, while its ability to prevent the development of illegal bowling actions can help improve the overall quality of cricket played in Sri Lanka. This thesis explores the development, implementation, and evaluation of the cricket talent scouting system with legal bowling action prediction in Sri Lanka, including its technical specifications, UI, and effectiveness in identifying and nurturing young cricketing talent. The suggested solution's legal, ethical, and societal challenges have been thoroughly examined, as have the product's design, implementation, and testing.

Keywords.

Cricket, Cricket Stats, Illegal Action Prediction, Player-Coach relationship, Scouting Platform, Sri Lanka Cricket.

1. INTRODUCTION

The aim of the project is to identify the difficulties faced by the Players/Coaches/Teams in spotting the cricket talent available in the country and to design, develop and evaluate a system which will assist the Male / Female players to showcase their talents and the Coaches/Teams to pick the right talent to their teams and guide them for the success with fewer dependencies.

2. PROBLEM STATEMENT

Cricket is the most popular sport in Sri Lanka, but in recent years, cricket fans have become less excited due to the Sri Lankan National men's cricket team's losing momentum [1]. Criticisms have arisen regarding the lack of talent production, even when potential players are identified [2]. Nevertheless, recent success in the Asia Cup showcased Sri Lanka's cricketing talent, albeit managing and getting the maximum output from them remains a challenge [3]. To address this, Sri Lanka conducts scouting programs across provinces, but limitations in information gathering and economic difficulties for players pose obstacles [4]. The role of coaches is crucial in young players' growth and success, leading to a high demand for quality coaches [5]. Despite Sri Lanka's reputation for producing talented players, the lack of a consistent scouting scheme has hindered their progress [8].

2.1 Inability to identify large pool of players.

Since Sri Lanka has only four high performance centers located to train the players, it will be inaccessible to most outstation players.

2.2 Lack of development in cricket

In Sri Lanka, many schools/ clubs don't have turf wickets. So that many players from outstations struggles while playing in the turf wicket for the first time. Many players do not know how to make use of the opportunities available far from their home which hinders them from making use of their talent [6].

2.3 Ineffective player management

Each player is unique. Since we don't have much qualified coaches, it is a very tough task to manage the players individually.

2.4 Transportation Problem for the players

Players from outstation more than often tend to miss the scouting sessions. Also, regular long-distance traveling will affect the player's fitness

2.5 Unable to scout players on regular basis.

Since Sri Lanka has lots of talented players all over the island, it is also a very tough task to arrange and organize scouting program regularly, which will cost more money and require more human resources. However, by scouting there are so many outstation players who have emerged and got opportunities [9].

2.6. Unable to find the players with legitimate bowling actions.

Illegal Bowling Action is when a player throws the ball rather than bowls it, defined by the ICC. It can lead to players being banned from bowling due to their bowling actions. Proper coaching is key to preventing this from happening [7].

3. METHODOLOGY

The CricPlayer Tracker project employed an iterative and incremental approach in its software development lifecycle (SDLC) to achieve timely and cost-effective development with well-organized deliverables. The project management methodology, PRINCE2, ensured control and risk mitigation in handling changing project requirements. For software development, the Rational Unified Process (RUP) allowed for frequent revisions, addressing essential functional requirements early on, and proved adaptable and effective for small-scale projects, leading to successful outcomes.

4. RESULTS & DISCUSSION

4.1 FINDINGS FROM INTERVIEW, QUESTIONNAIRE AND SURVEY

According to the study conducted among 112 players found that only 40.2% of the players who had played hard ball cricket at any age group went on to represent professional clubs or teams. This supports the coach's observation that a significant drop occurs in the number of students participating in cricket as they advance in age groups. Out of 20 students in the U13 category, fewer continue in the U15 level, and the numbers decrease further for U19. Only a few (less than 5 players) progress to the U19 team, with most dropping out in the early stages of their cricketing journey.

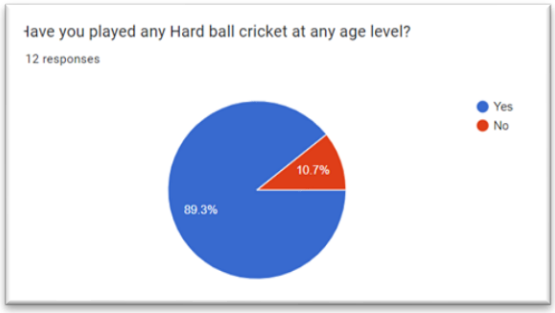


Figure 1: Survey responses from 112 players

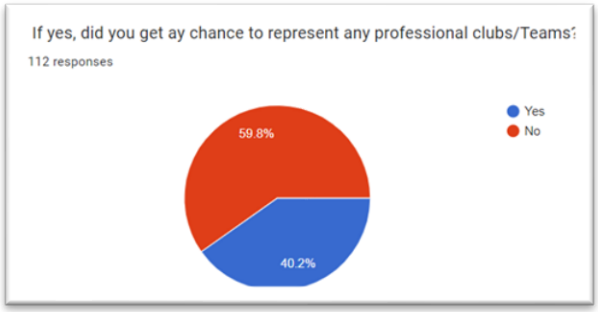


Figure 2: Players who got chance to proceed further in their career

4.2 REVIEW OF APPLICATIONS

Features	Ludimos	Pitch Vision	SLC Scorers.com	Batsman.com	Cric Clubs	Proposed Solution
Adding Stats	☑	☑	☑	☑	☑	☑
One on one chat	☑	☑	☒	☒	☒	☑
Illegal bowling action Predication	☒	☒	☒	☒	☒	☑
Sorting the players/Coaches	☒	☒	☒	☑	☒	☑
Measuring Speed of the ball	☑	☒	☒	☒	☒	☒
Sharing cricket videos/images	☑	☑	☒	☒	☑	☑
Search other Players/coaches	☑	☑	☑	☑	☑	☑
Adding match Scorecards	☒	☑	☑	☑	☑	☑
Removing Junk Data	☒	☒	☒	☒	☒	☑
Mobile Version	☑	☑	☑	☑	☒	☒
Focused on Sri Lanka	☒	☒	☑	☑	☑	☑

Table 1: Review of Applications

Applications	Advantages	Disadvantages
Ludimos	• Having Mobile App • Predicting the speed of the delivery • Having advanced features to record the video clips. • User Friendly application • In-build slow motion feature • Scheduling feature • Measuring the line & Length of deliveries. • Having a centralized dashboard.	• Unable to publicly share videos/images • Only paid version is available • Can't identify the illegal bowling action • Filtering the players' option is not available. • Not focused on Sri Lanka
Pitch Vision	• Live match streaming is available • Livescoring • Interactive scorecards • Auto Highlights • Having mobile app • In-build slow motion feature • Fitness management feature	• Can't identify the illegal bowling action • Filtering the players option is not available. • Not focused on Sri Lanka • Too many junk data
SLC Scorers.com	• Live match scorecards available • Focused on Sri Lanka • Having Mobile App • Player Stats available • Having tournament dashboard	• Can't identify the illegal bowling action • Filtering the players option is not available. • Unable to connect with coaches • Too many junk data • Unable to share images/videos • Only site owner can create users
Batsman.com	• Live match scorecards available • Focused on Sri Lanka • Having Mobile App • Player Stats available	• Can't identify the illegal bowling action • Filtering the players option is not available. • Unable to connect with

Table 2: Advantages & Disadvantages of Similar Applications

5. CONCLUSION

The uniqueness of the solution, CricPlayerTracker (CPT), is discussed, emphasizing its potential benefits for cricket coaches and teams in scouting talented players in Sri Lanka. However, certain limitations of the solution are acknowledged, such as the need for manual verification of uploaded videos and potential challenges with accuracy in certain scenarios. Future implementations are proposed, suggesting potential enhancements to the system, such as implementing image processing for user authenticity verification, automatic verification of certificates, improved action prediction, enhanced pose estimation techniques, recommendation models for nearby coaches, and expanding the system to cater to other sports.

6. ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my supervisor, Austen Mascrenghe sir, and co-supervisor Alqa Husni for their invaluable guidance, support, and encouragement throughout the course of this thesis. Their expertise, insights, and patience have been instrumental in shaping and refining my research. I am also grateful to my module leader, Abarnah Kirupananda miss, and course leader, Mrs. Janice Abeykoon miss, for their unwavering support and encouragement throughout my academic years, which has been vital in helping me navigate through the challenges of this research. I would like to thank my parents and my family, especially my mother, for their unconditional love, support, and sacrifices, which have enabled me to pursue my academic goals. Their unwavering belief in me has been a constant source of inspiration and motivation. Finally, I would like to extend my heartfelt appreciation to my colleagues for their camaraderie, support, and motivation throughout this journey. Their enthusiasm, insights, and encouragement have been a valuable source of learning and growth for me.

7. REFERNENCES

[1] “Sri Lanka Sports - The Most Popular Sports in Sri Lanka,” kevmrc.com, Apr. 11, 2022. https://www.kevmrc.com/most-popular-sports-in-sri-lanka (accessed Mar. 13, 2023).

[2] Paper, T., Editor, R., Page, O., 2022, B., Quotes, S., Sports, O., gardens, H. and Decline in talent and lack of passion in Sri Lankan cricket saddens me, s., 2022.

[3] N. Eranda, “Losing competitiveness in Sri Lankan Cricket: An Analysis based on Resource-Based View,” Jul. 2019.

[4] NDTV, Can’t Even Go To Practice": Sri Lanka Cricketer On Massive Fuel Crisis, https://www.ndtv.com/world-news/sri-lanka-fuel-crisis-chamika-karunaratne-asia-cup-2022-fuel-crisis-became-an-obstacle-for-sri-lankan-cricketers-says-chamika-karunaratne-3162978.

[5] Ludimos, Top 5 problems cricketers face during training, https://www.ludimos.com/blog/top-5-problems-cricketers-face-during-training .

[6]Decline in talent and lack of passion in Sri Lankan cricket saddens me, says Muralitharan. [online] Thehindu.com. Available at: https://www.thehindu.com/sport/cricket/decline-in-talent-and-lack-of-passion-in-sri-lankan-cricket-saddens-me-says-muralitharan/article26224611.ece

[7] “Sri Lanka take unorthodox route to success,” ESPNcricinfo. https://www.espncricinfo.com/story/sri-lanka-take-unorthodox-route-to-success-409410.

[8] D. Collins, Á. MacNamara, and A. Cruickshank, “Research and Practice in Talent Identification and Development—Some Thoughts on the State of Play,” Journal of Applied Sport Psychology, vol. 31, no. 3, pp. 340–351, Jul. 2019, doi: 10.1080/10413200.2018.1475430..

[9] S. L. Cricket, “Viyaskanth emerged through Sri Lanka Cricket’s talent hunt programme - Sri Lanka Cricket,” Dec. 06, 2020. https://srilankacricket.lk/2020/12/viyaskanth-emerged-through-sri-lanka-crickets-talent-hunt-programme/, https://srilankacricket.lk/2020/12/viyaskanth-emerged-through-sri-lanka-crickets-talent-hunt-programme

WorkPromo - Performance Evaluation System for Employee Promotion Eligibility of the IT Industry Sri Lanka

Anuji Senevirathna
Undergraduate
Informatics Institute of Technology
57, Ramakrishna Road, Colombo 6, Sri Lanka
Anuji.2018479@iit.ac.lk

Shyani Siriwardene
Senior Lecture/ Level Coordinator
Informatics Institute of Technology
57, Ramakrishna Road, Colombo 6, Sri Lanka
shyani.s@iit.ac.lk

ABSTRACT

Employee performance appraisal plays a key role in a company. However, most IT companies in Sri Lanka rely on 6 months or year-end evaluation process, and such organizations are reluctant to evaluate employee progress continuously in a regular manner. Since appraisals are not a regular process, difficulties often arise in evaluating performance in relation to identifying employees for promotion. The aim of this project is to identify and analyze the inefficiencies in current performance appraisal systems in the IT sector in Sri Lanka and design, develop, and evaluate a progressive web-based application to conduct regular employee performance appraisals with a view to facilitating employee promotion. in a regular manner to offer employee promotions. A preliminary pilot study was carried out among a biased convenience sample of employees who work in the IT sector, in a wide range of IT companies to understand the current inefficiencies in the performance appraisal process with a view to offering promotions to employees. A literature review further assessed the problem area to identify the related issues that contributed to the problem domain. A requirements-gathering process involving a survey and interviews was conducted to identify the necessary specifications for the solution to the identified problem. The platform "WorkPromo" was developed as a solution that helps IT organizations to conduct employee performance appraisals in a more efficient and accurate manner. It utilizes a machine learning model to predict employee promotion eligibility based on employee sprint-wise and annual review rating data. The software was created to overcome the inefficiencies of the current appraisal procedure and offer a comprehensive solution for employee promotions in Sri Lanka's IT sector.

Keywords

Employee performance evaluation, performance appraisal methods, agile methodology, machine learning, employee promotion eligibility, employee promotions, IT industry Sri Lanka.

1. INTRODUCTION

Performance appraisal is a crucial pillar for the success of any organization. It is a way of measuring the technical and soft skills of employees to ensure that their performance is maintained at an acceptable level [1]. Since the industrial revolution, labor rights and workplace standards were improved to benefit the employees. In the modern world, employee satisfaction and recognition are required to be maintained for the growth of the organization and employee retention [2]. Performance evaluation plays a major role and should aim at the career improvement of employees as well as offering them promotions. Most IT companies in Sri Lanka rely on a 6 months or year-end evaluation process, and such organizations are reluctant to evaluate employee progress continuously in a regular manner [3]. Performance appraisal at the end of the PR cycle is often not an efficient way for HR managers to evaluate employees. It could be stressful and challenging for the employees at the end of the PR cycle [4]. 90% of HR leaders claim that performance appraisal reviews do not provide accurate results to evaluate employees as the evaluation is conducted based on unwritten rules and a managerial perspective [5]. A pilot study among IT sector employees in Sri Lanka revealed that most employees interviewed are unsatisfied with the performance appraisal methods used in their organization. The main escalating factors for the proposed solution are inadequate performance appraisal methods, lack of guidance for career development, lack of clarity regarding job roles, and negative impact on the managers [5].

This study aims to address obstacles in employee performance evaluation and to explore solutions for a more efficient platform to conduct appraisals and enable employee promotion. By addressing these research objectives, the study seeks to identify ways to overcome these challenges and improve the overall effectiveness of employee performance appraisal.

2. LITERATURE REVIEW

This literature review focuses on analyzing current research and existing solutions in the area of employee performance appraisal, in order to identify limitations and gaps in the process and potential solutions.

2.1 Concept of Performance Appraisal

Performance appraisal is a common management practice used in organizations to assess employees' work behavior and its impact on overall company performance. It fosters a positive work environment, encouraging healthy competition and a sense of achievement for both employees and the company [6]. The process involves evaluating employees' qualities and job performance to achieve three main objectives [7].

It aims to provide each employee with pertinent performance feedback [7].

It intends to serve as an entry point for behaviour change or improvement in the direction of more efficient work practices [7]. It makes an effort to provide managers with data they may use to evaluate forthcoming job promotions [7].

2.2 Challengers of performance appraisal in the IT sector

Demotivation of staff members and inequitable evaluations can occur due to the manager's personal opinions and prejudices in performance evaluation in the IT sector. Performance standards along with appraisal techniques must be regularly updated due to the IT sector's fast rate of growth [8]. The intricacy and caliber of IT tasks may be difficult to assess using conventional evaluation approaches. It is necessary to create evaluation processes that can accurately gauge the individual's technical knowledge, capacity for resolving issues, and creativity [8]. Project-oriented employment and agile processes are prevalent in the IT sector, necessitating regular efficient feedback methods [9]. The continual development of agile projects might not prove compatible with the traditional performance evaluation processes, requiring an embrace of agility-focused performance evaluation techniques [10].

2.3 Evaluation of similar applications

"AssessTEAM" is identified as the most comparable software to the proposed web application, serving as a comprehensive solution for managing employee performance. It focuses on enhancing overall performance rather than solely conducting performance appraisals for the purpose of employee promotion. It is capable of setting and tracking employee goals, self-evaluation, providing feedback, and monitoring employee performance to increase employee performance. However, the following feature gaps have been identified in this application and other existing applications. Based on them, a solution was proposed. The solution proposed in this research attempts to address these gaps.

Gaps identified include the inability to conduct task-wise or sprint-wise reviews and ratings of employees and provide sprintwise feedback and learning recommendations; and the absence of task-wise rating details for better performance appraisal. It also includes the lack of specific criteria tailored to evaluating employee performance, and inadequate attention being given to promotion procedures.

By addressing these gaps research the study seeks to identify ways to overcome the challenges and improve the overall effectiveness of employee performance appraisal.

3. METHODOLOGY

Interviews and questionnaires as primary methods. Interviews were conducted with IT sector employees and included employees, supervisors, managers, and professionals involved in performance appraisal. The goal was to gather diverse perspectives and firsthand insights, uncovering new requirements and alternative viewpoints. Questionnaires were also distributed to IT sector employees to gather their opinions and identify the optimal requirements for the proposed solution. 45 responses were received for the industry survey that was sent out.

In addition to primary research methods, a literature review and background reading were conducted. The literature review involved researching and analyzing existing literature on the research area, providing valuable knowledge of the challenges, gaps, and proposed solutions in the field. Background reading involved utilizing various sources such as reports, journal articles, and websites, to gain a deeper understanding of the problem and desired solution. Furthermore, existing systems were studied to identify areas for improvement and ensure the proposed solution effectively addressed the identified gaps. These research methods collectively contributed to a comprehensive understanding of the subject and informed the development of the proposed solution.

4. FINDINGS

The main findings derived from the interviews and questionnaires can be summarized as follows.

I. Most IT companies are currently following a 6 month or year-end performance evaluation process.

II. Most IT companies' goals and objectives are not clear and specific to their job role.

III. The current 6-month or year-end employee performance appraisal process does not effectively enable employees to understand their strengths and weaknesses and thereby enhance their career growth.

IV. A significant number of employees claimed that their company's evaluation process and promotion or reward system lacked objectivity and fairness.

5. PROPOSED APPROACHED

5.1 Proposed Application

WorkPromo is a web application created to conduct employee performance appraisals in the IT industry in Sri Lanka. The application primarily focuses on companies that utilize Agile methodologies, which have become widely adopted in the IT sector. The WorkPromo application fills the gaps of 6-month and year-end evaluations with its unique features and provides a comprehensive solution to effectively evaluate employee performance and make informed decisions regarding employee promotions.

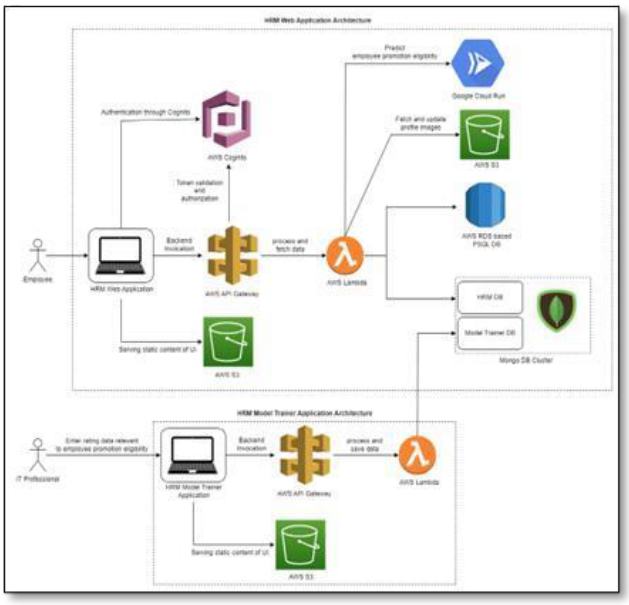


Figure 1 – High- Level Diagram

6. CONCLUSION

The research project has made significant contributions to the field of employee performance appraisal in the IT industry in Sri Lanka. However, there are several limitations that need to be acknowledged.

The WorkPromo application currently lacks mobile compatibility with IOS and Android platforms. Additionally, its evaluation process is limited to a specific set of IT roles, catering primarily to IT companies that adhere to Agile methodologies. Furthermore, the application's administrative functionality is constrained, as the admin cannot register multiple users simultaneously or disable users within the application. The determination of job promotion eligibility relies on the accuracy of the ML model employed. When a manager is responsible for more than five employees, evaluating them sprint-wise becomes an overwhelming task. Another limitation is the absence of a feedback-integrated dashboard, which prevents the admin from conveniently viewing and comparing user feedback.

For future work, the pre-processing techniques employed in the ML model will be enhanced. Additionally, the admin portal will be upgraded to allow the simultaneous addition of multiple users to the application, incorporating upload and download functionality. Moreover, an email notification service will be implemented to inform managers when employees submit their sprint-wise and annual self-evaluations. Furthermore, the admin will be granted the capability to disable and delete users. Furthermore, integration with JIRA will be explored to streamline task tracking for employees.

7. ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisor, Mrs. Shayani Sririwardene, for her invaluable support and guidance throughout my final year project. I am also thankful to Mr. Kavindu Somadasa, Mr. Rochana Lakshitha, and Mrs. Binari Smarasinghe for their insightful feedback and trust in the potential of my solution. I extend my appreciation to the faculty members at IIT for their mentoring and expertise. Special thanks to Mrs. Janice Abeykoon and Mrs. Abarnah Kirupananda for their guidance and motivation. I am grateful to my friends for their unwavering support, and most importantly, I want to express my profound appreciation to my parents for their love, sacrifices, and endless belief in me.

REFERENCES

[1] 4 Employee Performance Management Challenges and How to Fix Them. (2021). Bplans Blog. Available from https://articles.bplans.com/employee-performance-management/ [Accessed 28 October 2022].

[2] 5 major Challenges of Performance Review & Appraisals in
Organizations. (no date). Available from https://talygen.com/blogdetail/5-major-challenges-ofperformance-review-- appraisals-in-organizations [Accessed 3 May 2023].

[3] Acikgoz, Y. (no date). How to make performance reviews less terrible – especially given the challenges of supervising remote workers. The Conversation. Available from http://theconversation.com/how-tomakeperformancereviews-less-terrible-especiallygiven-the-challenges-ofsupervising-remote-workers182009 [Accessed 4 November 2022].

[4] Advantages & Disadvantages of Gantt Charts. (no date). Analysistabs. Available from https://analysistabs.com/ganttchart/advantages-disadvantages/ [Accessed 4 November 2022].

[5] Bacal, R. (2018). Ten Stupid Things Managers Do To Screw Up Performance Appraisal. Medium. Available from https://rbacal.medium.com/ten-stupid-things-managers-dotoscrew-up-performanceappraisalc87dc6244a6a [Accessed 5 October 2022].

[6] Bacal, R. (2018). Ten Stupid Things Managers Do To Screw Up Performance Appraisal. *Medium*. Available from https://rbacal.medium.com/ten-stupid-things-managers-doto-screw-up-performanceappraisalc87dc6244a6a [Accessed 5 October 2022].

[7] Benefits of Wireframe - Reasons to Use Wireframe - Edraw. (no date). Available from https://www.edrawsoft.com/benefits-of-wireframe.html [Accessed 2 May 2023].

[8] Cornerstone Talent Experience: One platform. Limitless potential. (no date). Available from https://www.cornerstoneondemand.com/ [Accessed 5 November 2022].

[9] Daoanis, L.E. (2012). PERFORMANCE APPRAISAL SYSTEM: It's Implication To Employee Performance. *International Journal of Economics and Management Sciences*, 2 (03).

[10] De Andres, R., García-Lapresta, J.L. and González-Pachón, J. (2010). Performance appraisal based on distance function methods. *European Journal of Operational Research*, 207 (3), 1599–1607.

Warranty Wallet: A Platform to Integrate Manufacturers, Retailers, and Consumers for a Pitfall-Free Warranty Process

Madushi Munasinghe
School of Business,
Informatics Institution of Technology,
Ramakrishna Road, Colombo 06.
madushi.2019765@iit.ac.lk

Shyani Siriwardene
School of Business,
Informatics Institution of Technology,
Ramakrishna Road, Colombo 06.
shyani.s@iit.ac.lk

Salitha Dinushika
School of Business,
Informatics Institution of Technology,
Ramakrishna Road, Colombo 06.
salitha.d@iit.ac.lk

ABSTRACT

Warranty programs are becoming increasingly common in the modern world, providing consumers with the assurance that the manufacturer intends to maintain a long-term relationship with the consumers and provide after-sales service. Warranty programs also provide consumers an opportunity to have defective items repaired or replaced with a new one at a no cost or at a reduced cost. Retailers play an important role by facilitating the warranty process.

Although an effective warranty procedure adds value to consumers, retailers, and manufacturers, flaws in the process due to a lack of integration drive stakeholders away. The aim of this project is to analyze the factors that cause the lack of integration among the stakeholders in the current warranty process in Sri Lanka and to design, develop, test, and evaluate an application that will address the pitfalls in the warranty process.

In order to achieve the above-mentioned aim, a detailed literature survey was conducted to identify pitfalls in the current warranty process. Furthermore, findings were confirmed through surveys and interviews conducted with stakeholders involved in the warranty system. Addressing identified requirements, the "Warranty Wallet" application is introduced. The application was further reviewed by technical and industry experts.

Keywords

Warranty Tracking, Warranty Claim Management, Warranty Data Analysis, Smart Contract, QR Integration

1 INTRODUCTION

Warranties are commonly issued for goods in the modern world, providing consumers with assurance of long-term support and manufacturers with valuable product data. However, deficiencies in the warranty process can have a negative impact on consumer experience, leading to a decreased likelihood of repeat business. Integration among manufacturers, consumers, and retailers is crucial for a smooth warranty process.

Consumers face difficulties in warranty registration and tracking, with low completion rates for warranty cards and the need to remember various warranty details and document locations. Claiming warranties also poses challenges, as consumers require specific proof such as purchase receipts or warranty cards, and contacting warranty providers can be cumbersome.

Retailers struggle with tracking purchase history, making it difficult to authenticate warranty claims, particularly for identical products sold by multiple retailers. The two-week grace period for claims puts additional pressure on retailers, who may have to bear the cost if the claim falls within this timeframe.

Privacy concerns and low response rates contribute to difficulties in verifying consumer identities and certifying warranty claims. This lack of confirmation hampers the warranty process for retailers and manufacturers alike.

Manufacturers may be unaware of product defects unless retailers inform them, potentially impacting decision-making, and brand reputation. Moreover, the absence of a direct connection between consumers and manufacturers opens the door for counterfeit goods and counterfeit warranties, enabling fraudulent activities by retailers.

Overall, addressing these challenges and promoting collaboration among stakeholders is essential for improving the warranty process and enhancing the consumer experience.

2 BACKGROUND

2.1 Role of Warranty

For consumers, a warranty assures the consumer that the faulty item will be either repaired or replaced at no cost or at reduced cost[4]. Warranty is considered as a signal of quality. A longer warranty period reduces purchase risk[2].

A warranty provides protection to manufacturers since warranty policies specify the coverage of warranty. It also helps in terms of promotion of sales of goods since it gives consumers a sense of reliability. In addition, a good warranty program will add value to the product[6], which is beneficial for both manufacturers and retailers. In addition, data gathered through warranty will help the manufacturer to take decisions, especially in terms of improving the quality of a product [3].

2.2 Current Warranty Process

The warranty process involves three main components: registration, tracking, and claiming, which require active consumer engagement. Warranty registration can be done through paper forms or electronic methods, and some companies use

purchase bills as proof of warranty. Consumers face challenges in keeping track of warranty proofs and often misplace them, leading to difficulties in claiming warranties. Loss of warranty proof can result in consumers being unaware of warranty expiration dates. There are various routes for warranty claims, depending on the type of failure reported, and it is important to attend to claims promptly and effectively.

Difficulties in contacting manufacturers and retailers, complicated claiming processes, and the lack of consumer databases pose challenges in the claiming process. Manufacturers and retailers struggle with authorizing claims when warranty proof is not presented and with maintaining synchronized communication among stakeholders. Information and communication technology (ICT) can help improve the warranty process by providing electronic warranties, remote assistance, cooperative warranty services, and predictive warranty services.

Warranty data analysis plays a crucial role in the management process. It involves setting objectives, strategies, and priorities to improve customer satisfaction, minimizing waste and unnecessary effort, conducting assessments of product performance, and making necessary improvements to the warranty program. Warranty data includes claims data collected during servicing and supplementary data for effective management. It is used for early detection of production issues, observing product performance targets, understanding relationships between different product levels, predicting future claims and costs, and estimating product reliability for warranty policy decisions. However, there is a risk of low data quality due to various factors, so steps should be taken to improve data quality.

3 Methodology

3.1 Questionnaires

A questionnaire was distributed among consumers to understand their opinion and issues faced in the current warranty process.

The survey was distributed covering consumers who are above 18 years of age. The Snowball sampling technique was used [5], and survey was distributed among individuals who are passionate about warranty process and encouraged them to distribute it with others.

3.2 Interviews

Interviews were conducted to understand different warranty processes companies follow as manufacturers/retailers and to identify issues associated with them. Interviews were conducted covering small scale businesses to large scale businesses [1].

4 RESULTS

During the requirements process, 5 staff members from manufacturing companies and 10 staff members from retail companies were interviewed to identify the processes and pitfalls in the current warranty process. Apart from the interviews, 141 responses were gathered from the survey which was distributed among consumers via google form.

The survey had 141 respondents, with a gender distribution of 55.5% females and 44.7% males. Most respondents (68.1%) considered warranties to be crucial.

Of the interviewees, all 15 stated that warranties are vital for their businesses, as they increase customer trust, aid in customer retention, add value to products, serve as a marketing tool, and help evaluate products. Additionally, 14.2% of respondents reported experiencing warranty fraud.

Regarding warranty registrations, only 14.2% of the survey respondents frequently followed the process due to difficulties and lack of awareness. Manufacturers' warranties were offered by most manufacturing companies (8 out of 10 interviewed), while two retail companies offered both manufacturer's and retailer's warranties.

The majority of respondents used physical methods (such as purchase bills and paper-based warranty cards) to track warranties, with only a small percentage using software solutions. Most manufacturers had their own warranty tracking systems, while many retailers lacked such systems, making it challenging for consumers to claim warranties without their warranty cards.

Warranty claiming predominantly involved visiting the retailer of the product (85.4%), highlighting the crucial role of retailers in the process. Difficulties during the claiming process were reported by some respondents, including the lack of required documents and failure to follow pre-defined methods.

As a conclusion, the survey revealed challenges in warranty registration and tracking, with low completion rates for warranty cards. Retailers faced difficulties in authenticating claims and tracking purchase history. Privacy concerns and low response rates hampered the verification of consumer identities. Lack of integration also led to counterfeit goods and fraudulent activities.

5 SOLUTION

5.1 Overview of solution

Warranty Wallet addresses the challenges identified in the current warranty process by providing a seamless platform for stakeholders to integrate efficiently. Through the use of QR codes, retailers and consumers can easily register and track their warranties, eliminating paperwork and data management difficulties. The platform streamlines the claiming process, enabling direct communication between manufacturers, retailers and consumers reducing privacy concerns and ensuring authentic warranties through blockchain technology.

With real-time data analysis, manufacturers can make informed decisions to improve product quality, while consumers receive timely alerts about warranty expirations. Overall, Warranty Wallet enhances collaboration among manufacturers, retailers, and consumers, resulting in a more efficient and trustworthy warranty management experience. The High-level architecture of the system is included in figure 1.

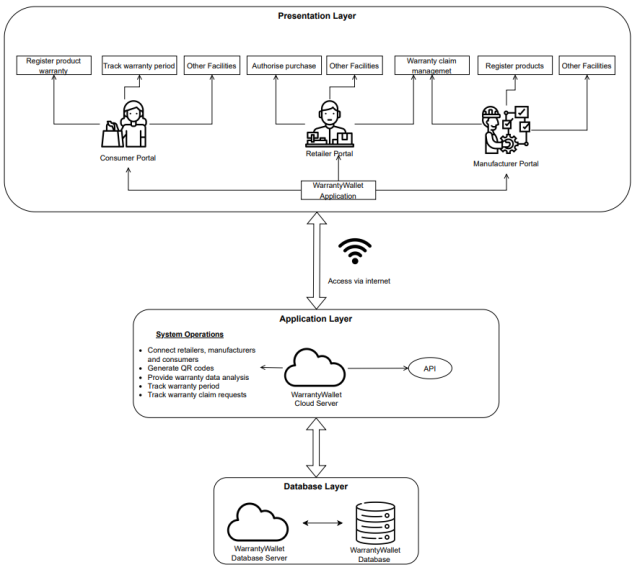


Figure 1: High-level Architecture of Solution

5.2 Technologies

The application was developed as a responsive web application using React JS, Node JS and Mongo DB as the main technologies. Apart from these, blockchain was used to securely store data such as the purchase date and warranty period, to avoid any alterations.

6 RECOMENDATIONS AND FUTURE WORK

Recommendations and future work suggested for the further development of the app is discussed under the section 06.

6.1 Recommendations

- Prompt if there are any warranty claims in the "in progress" status when a consumer files multiple warranty claims, to avoid duplications.
- Provide the ability to attach pictures or videos when creating a warranty claim and in chat messages.
- Develop a mechanism to handle consumer warranty registrations if a retailer skips warranty registration.
- Enhance data analysis for manufacturers to view both manufacturer and retailer claim data.
- Send consumers alerts when the product warranty is nearing expiry.
- Add search bars to increase usability of the platform.

6.2 Future Work

- Integrate the warranty management system with a point of sales system and inventory system.
- Permit multiple users from three portals to register under one warranty, when a transfer of a product among those stakeholders takes place. Provide the option to skip the retailer when there is a direct purchase with a manufacturer.
- Permit retailers to register a warranty for a consumer if the consumer is agreeable to this.
- Permit role-based access for manufacturer and retailer portals providing access for different job roles under manufacturer or retailer logins, using their credentials.

7 ACKNOWLEDGMENTS

The author, extended her gratitude to the final year project module team, Ms. Abarnah Kirupananda and Ms. Janice Abeykoon, and staff at IIT for their insights and feedback. Additionally, she acknowledged the invaluable contributions of industry professionals and expressed her appreciation for the unwavering support of her family and friends.

8 REFERENCES

[1] Hamza Alshenqeeti. 2014. Interviewing as a Data Collection Method: A Critical Review. *ELR* 3, 1 (March 2014), p39. DOI:https://doi.org/10.5430/elr.v3n1p39

[2] Kaiying Cao and Ping He. 2018. Price and warranty competition in a supply chain with a common retailer. *INFOR: Information Systems and Operational Research* 56, 2 (April 2018), 225–246. DOI:https://doi.org/10.1080/03155986.2017.1363590

[3] Guillermo Gallego, Ruxian Wang, Julie Ward, Ming Hu, and Jose Luis Beltran. 2014. Flexible-Duration Extended Warranties with Dynamic Reliability Learning. *Prod Oper Manag* 23, 4 (April 2014), 645–659. DOI:https://doi.org/10.1111/poms.12178

[4] D.N.P Murthy and I Djamaludin. 2002. New product warranty: A literature review. *International Journal of Production Economics* 79, 3 (October 2002), 231–260. DOI:https://doi.org/10.1016/S0925-5273(02)00153-6

[5] P. Sedgwick. 2013. Snowball sampling. *BMJ* 347, dec20 2 (December 2013), f7511–f7511. DOI:https://doi.org/10.1136/bmj.f7511

[6] Xiaolin Wang and Wei Xie. 2018. Two-dimensional warranty: A literature review. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability* 232, 3 (June 2018), 284–307. DOI:https://doi.org/10.1177/1748006X17742776

Influence of Supervisory Support and Work-Life Balance (WLB) on Organizational Citizenship Behavior (OCB) in the Apparel Industry in Sri Lanka

Sandaruwan Udayanga
sandaruwan.2019685@iit.ac.lk

Gayashini Rathnayake
gayashini.r@iit.ac.lk

ABSTRACT

Keywords

1. Chapter 1: INTRODUCTION

1.1 Background

In 2020, the garment sector generated export earnings of US$3938.85 million, accounting for 39.74% of the nation's product exports (EDB Sri Lanka, 2020). A total of 350,000 people are employed directly in the industry in Sri Lanka, according to Sri Lanka Apparel, the website of the Joint Apparel Association Forum of Sri Lanka (JAAFSL 2020).

1.2 Research Gap

Employee turnover is high among machine operators in the apparel sector. OCB reduces employee turnover. However, OCB has received limited research attention in the apparel sector in Sri Lanka. Supervisor support and Work-Life Balance lead to less employee turnover. However, the influence of supervisor support and Work-Life Balance on OCB has received limited research attention. This is the gap that this study attempts to make in the research literature in the post-pandemic era.

1.3 Research Questions

RQ1: What is the influence of supervisory support on Organizational Citizenship Behavior?

RQ2: What is the impact of Work-life balance on Organizational Citizenship Behavior?

1.4 Research Objectives

RQ1: What is the influence of supervisory support on Organizational Citizenship Behavior?

RQ2: What is the impact of Work-life balance on Organizational Citizenship Behavior?

2.Theoretical Background and Models

2.1 Social Exchange Theory

This study's guiding principle is the social exchange theory (SET), which was created by Homans in 1958. The idea holds that people enter into relationships in anticipation of specific benefits, which often manifest as either material or intangible resources (Homans, 1958; Thibaut & Kelley, 1959; Blau, 1968; Emerson, 1962) (Sani P 2020).

2.2 Boundary Theory

This theory emphasizes the establishment and management of boundaries between work and personal life. It suggests that individuals can maintain work -life balance by creating and maintaining clear boundaries that separate their professional and personal spheres.

3.Conceptual Model

Supervisory support is the independent variable (Li et al., 2021). Work-life balance is the independent variable (Kurniasih et al., 2021). Organizational citizenship behavior (OCB) is the dependent variable (Kurniasih et al., 2021).

Organizational Citizenship Behaviour is defined as " discretionary, extra-role behaviors on the part of a salesperson that have been shown to influence managers' evaluations of performance" (Organ, 1995)
The term "organizational citizenship behavior" (OCB) describes people's actions that support efficient organizational performance. By creating a supportive social and psychological environment where the task influences to achieve the organizational objectives effectively. (Newton & Kim, 2016) Employees value OCB because it fosters social relationships that have an impact on their ability to do their jobs. (Newton & Kim, 2016) (Gnankob, 2021). Organizational citizenship behavior (OCB) is a fast-emerging term that examines how and why people positively contribute to their organizations outside of clearly defined professional duties. (Gulliver, 2020). OCB has five dimensions including conscientiousness, sportsmanship, Courtesy, civic virtue, and altruism (Littooij, 2020).

Supervisory support is defined as "Employee perceptions of how much their supervisors regard their efforts and are concerned about their well-being are referred to as supervisor support."(Eisenberger et al., 1986; Kottke & Sharafinski, 1988; M Mohamed, 2016). Supervisor support (SVS), by organizational support theory (Eisenberger et al., 1986), enables changes in employees' levels of affective commitment (Sadiya, 2015). Mohamed (2016) in the seminal article argued that Supervisors are tasked with directing and assessing employees' work performance on behalf of the company.
H1: There is a significant positive influence between supervisory support and OCB

Work-Life Balance is defined as a " strategy that aids employees in juggling their personal and professional obligations"(Shirmohammadi et al., 2022). A person who chooses to prioritize their personal and professional obligations equally has achieved a work-life balance. A person who lacks this balance has more duties at work and home works longer hours and has less free time. There is an internal and external urge to find a work-life balance. (McLean, 2022). increase employee satisfaction, companies place high importance on

incorporating positive work-related emotions (such as job satisfaction and organizational responsibility)

4.Methodology

4.1 Research Philosophy

The study's hypotheses and research questions may have also been affected by positivist presumptions, such as the idea that knowledge can be based on observable and quantifiable data. In the Sri Lankan apparel industry, for instance, the researchers may have predicted that there is a positive relationship between supervisory support, work-life balance, and organizational citizenship behavior and tried to test this hypothesis by quantitative data analysis. (Thevanes, N. and Harikaran, S. 2018) (Vashisht, S. and Vashisht, R. 2022).

4.2 Research Approach

This research uses the deductive approach to Test the relationship between key variables (Bougie and Sekaran, 2019).

4.3 Research Method

The study utilized quantitative research techniques influenced by the positivistic paradigm. Quantitative techniques were employed, and previous studies (Tinh et al., 2021; Bougie and Sekaran, 2019) were referenced as they are considered to have less bias. Data collection was conducted through a survey, which included validated scales and indicators based on prior empirical research. The survey instrument underwent pilot testing before its implementation

4.4 Sampling Strategy

Case study approach was used. Cluster sampling technique was used and the garment factory workers were selected.

4.5 Research Instrument

Supervisory support was operationalized using a scale used by Tsui et al (1997). Work life balance was operationalized using measures from WLB scale by Hayman (2005).

OCB was operationalized using five indicators included in the OCB scale by Podsakoff (1990) employing a 1-5 ordinal scale ranging from strongly disagree to strongly agree.

4.6 Data Collection

A pilot study was conducted using 30 sets of data. The questionnaire was distributed among 350 employees of the sewing machine operators in the factory and 210 usable responses was collected and the rest was removed. The response rate of the service is 60% which is good. The data was collected from factory level workers in a single, leading apparel factory. The questionnaire was translated to Sinhala since the respondents do not have an adequate understanding of English. The questionnaire was back-translated to minimize errors in translation.

5.Analysis

The data gathered were analyzed using the SPSS version 26 the first step of the analysis part is to test the reliability to identify the suitability of the data gathered. To analyze the associations between the variables, normality, correlation, regression, reliability, validity testing, and hypothesis testing were used.

Discussion of research objective 1

In the Sri Lankan apparel business, supervisory support had a median positive influence on corporate citizenship behavior ($P< 0.05$). According to the empirical data from Cheng Chang & Feng Uen's journal articles from 2020 (Cheng Chang & Feng Uen, 2022) and Cheng Chang H and Uem J's articles from 2022 (Eliyahu, 2022)

According to regression analysis, supervisory support has a greater impact on OCB than Worklife balance. This finding supports previous research findings by (Shakshi and Ravi in their journal article from 2022)

There is a strong positive relationship between work-life balance and organizational citizenship behavior. This supports prior empirical findings by N. Thevane and S. Harikaran, lecturers at Jaffna University, the correlation between the work-life balance and organizational citizenship behavior is 0.68, demonstrating a strong correlation between the dependent and independent variables (Thevanes & Harikaran, 2018).

As the normality of the data set shows a bell curve, we can conclude that the data set is normal and the Kurtosis data chain should lie between -3 and +3 to become normalized distribution (Thevanes & Hariharan, 2018).

We can therefore conclude that there is a positive relationship between the independent factors, such as work-life balance and supervisory support, and the dependent variable organizational citizenship behavior based on the previously mentioned study instrument and empirical data. Sri Lanka's apparel sector.

According to the demographic data, the majority of the workers in the appeal sector are women who have only completed the GCE OL, and most of them make between 20,000 and 50,000 per month. This supports prior empirical findings by EDB Sri Lanka, 2020 official website (EDB Sri Lanka, 2020)

Conducting training programs for supervisors will help to develop a group of motivated workers in the advisory sector, reducing conflicts between supervisors and team members and lowering labor turnover. (Faerman, 2019)

The relationship between the team members and the supervisors will be strengthened if the workplace is made nice and if the team members are given the materials they need for competitions. (Faerman, 2019)

putting up initiatives to foster communication between the team members and the supervisors (Faerman, 2019)

Conducting training programs for supervisors will help to develop a group of motivated workers in the advisory sector, reducing conflicts between supervisors and team members and lowering labor turnover. (Faerman, 2019)

The relationship between the team members and the supervisors will be strengthened if the workplace is made nice and if the team members are given the materials they need for competitions. (Faerman, 2019)

putting up initiatives to foster communication between the team members and the supervisors (Faerman, 2019)

Flexible working conditions: Offering flexible working conditions, including remote work or flexible schedules, can aid employees in better juggling their professional and personal obligations. Additionally, it may increase employee retention and morale. (Daily News, 2020)

Regular breaks: Employers might urge staff to take regular breaks to prevent burnout and sustain productivity. This can involve frequent, brief pauses as well as longer ones for lunch or exercise. (Daily News, 2020)

Training in time management: Employers can offer time management training to assist staff in setting priorities for their work and improving time management. Employees may be able to better manage their job and personal responsibilities as a result. (Daily News, 2020)

Building a culture at work that is helpful and values work-life balance can make employees feel appreciated and supported. This can entail offering services for mental health, promoting open communication, and emphasizing the value of work-life balance. (Daily News, 2020)

Clear expectations: To assist workers in better time management, employers can establish clear expectations and limitations around work hours and tasks. Additionally, this can lessen the risk of burnout and raise general job satisfaction. (Daily News, 2020)

A combination of rules, practices, and cultural adjustments that place a priority on employee welfare and support their professional and personal goals are needed to improve work-life balance in the garment business.

Based on the enrichment theory we can say that there is a bond between work-life balance and OCB based on (Farh 2004; Cooper 2013) which OCB will be dependent on work-life balance.

Based on the aforementioned conclusions, it is possible to infer that numerous additional factors affect individual behavior as well as organizational citizenship behavior at all types and levels of the organization. As a result, additional research is required to complement the current understanding of the additional factors that may influence individual behavior as well as extra-role behavior (OCB).

1.5 Limitations of the Research

Since this research instrument focused on MAS Bodyline the values generated from other companies can be different from others because of the size of the factory and the team member count

Other than work-life balance and Supervisory support there are lots of independent factors which influence the OCB such as leadership motivation etc.

The depth of insights gained from quantitative research is often constrained by the use of organized data collection techniques like surveys and experiments. (Radu, 2023) The thorough exploration of complicated events or the capture of subjective sensations may therefore not be achievable. (Radu, 2023)

.

5 References

- Abdullah, M.R. and Marican, S., (2020). Supervisory support and organisational citizenship behaviour: the role of self-efficacy as mediator. *International Journal of Public Sector Performance Management*, *6*(6), pp.854-875.
- Anushi, S.K., Priyanath, , H.M.S. and Tennakoon,, W.D.N.S.M. (2022). Nexus among psychological apital, work-life alance and job satisfaction ... *Nexus among Psychological Capital, Work-life Balance and Job Satisfaction of Employees in Apparel Industry*. Available from https://www.researchgate.net/profile/Niranjala-Tennakoon/publication/358909210_Nexus_among_Psychological_Capital_Work-life_Balance_and_Job_Satisfaction_of_Employees_in_Apparel_Industry/links/621f2f6c7106690c08538879/Nexus-among-Psychological-Capital-Work-life-Balance-and-Job-Satisfaction-of-Employees-in-Apparel-Industry.pdf [Accessed 30 January 2023].
- Bell, E. (2022). *Business research methods*. Oxford university press.
- Bello, Z. and Tanko, G. (2020). (PDF) review of work-life balance theories - researchgate. *Review of Work-Life Balance Theories*. Available from https://www.researchgate.net/publication/347811070_Review_of_Work-Life_Balance_Theories [Accessed 22 April 2023].
- Bismala, L. (2019). Factors affecting organizational citizenship behavior: A literature review. *EUDL*. Available from https://eudl.eu/doi/10.4108/eai.8-10-2018.2288740 [Accessed 19 April 2023].
- Bismala, L. (2019). Factors affecting organizational citizenship behavior: A literature review - [scite report]. *scite.ai*. Available from https://scite.ai/reports/factors-affecting-organizational-citizenship-behavior-vJdYNMrN [Accessed 20 April 2023].
- Bougie, R. and Sekaran, U. (2019). *Research methods for business: A skill building approach*. John Wiley and Sons.

Influence of Customer Engagement and Brand Attachment on Brand Loyalty in the Fashion Industry in Sri Lanka

ABSTRACT

This paper will explore the effect of customer engagement and brand attachment on brand loyalty in the Fashion Industry in Sri Lanka. Findings revealed that to foster brand loyalty, fashion brands should enhance customer engagement and encourage brand attachment. It exposed a positive relationship between Customer engagement, brand attachment, and brand loyalty. Customer engagement on social media significantly impacts brand loyalty positively (Nash, 2019; Lim and Rasul, 2022). Brand attachment serves as a mediator between customer engagement and brand loyalty (Li et al., 2020; Kumar and Nayak 2019; Diallo et al., 2021; Goyal and Verma, 2022). Studies show that brand loyalty is influenced by brand attachment (Yu, 2020). Customer engagement is a crucial driver of brand loyalty (Samarah et al., 2021; Singh and Srivastava, 2019; Adhikariet al., 2019). Engaging with brands on social media leads to positive outcomes like loyalty (Vivek et al., 2012; Park et al., 2021; Bergel et al., 2019). Only focused on Facebook (Gupta and Sharma, 2022) that having 7.69 million users in Sri Lanka and limited to Kelly Felder, GFlock, Avirate, Arienti, Kathyusha Exclusivewear, and Lurreli (Statista, 2021) brands.

Key Words

Customer engagement, Brand attachment, Brand loyalty, Fashion Brands, Facebook

1. Research Gap

Previous studies (Li et al., 2020; Shetty et al., 2021; So et al., 2022; Ting, Abbasi, and Ahmed, 2020; Hu et al., 2019; Datta et al., 2021; Molinillo et al., 2020; Kietzmann et al., 2020; Ghani et al., 2020; Bolukbasi and Koc, 2021; Kim and Woo, 2019; Zaidun, Muda, and Hashim, 2021; Choi, Lee, and Lee, 2021) have identified several research gaps and limitations in understanding the relationships among customer engagement, brand attachment, and loyalty. But there is a lack of understanding regarding the influence of customer engagement on social media platforms and the specific dimensions of brand loyalty. In the context of Sri Lanka, empirical inquiries have examined customer engagement, brand attachment, and brand loyalty in various industries, yet a distinct research void remains within the fashion industry (Nishantha, 2020; Jayaweera and Abeysekera, 2020; Kodithuwakku, 2021; Wickramanayake and Weerasinghe, 2020).

2. Research Questions and Research Objectives

RQ1: What is the influence of Customer Engagement on Brand Loyalty?

RQ2: What is the impact of Brand Attachment on Brand Loyalty?

RQ3: What is the effect of Customer Engagement on Brand Attachment?

RQ4: How does brand attachment mediate the effect of customer engagement on brand loyalty?

Based on the above Research Questions, the Research Objectives below were identified.

RO1: To understand the influence of Customer Engagement on Brand Loyalty

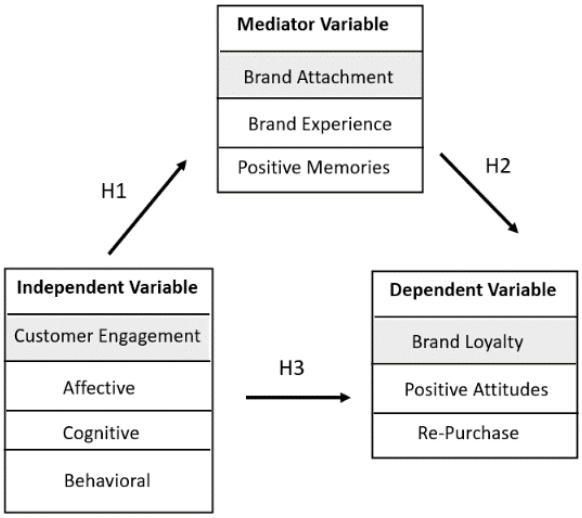
RO2: To assess the impact of Brand Attachment on Brand Loyalty

RO3: To identify the affect of Customer Engagement on Brand Attachment

RO4: To identify the mediating effect of brand attachment on customer engagement and brand loyalty.

3. Conceptual Model

Customer engagement is an independent variable. Brand loyalty is the dependent variable and brand attachment is the mediator variable.



Customer engagement is a complex phenomenon that involves emotional, psychological, behavioral, affective, and cognitive components (Jayalath and Galdolage, 2021). It is defined as a state where customers feel comfortable and actively participate with a brand beyond makinga purchase (Doorn et al., 2010, p. 254; De Silva, 2021).

H1: There is a significant positive impact of customer engagement on brand loyalty.

Li (2019) defined brand attachment as "The process by which individuals establish a connection with a brand through the creation of a self-connection to positive memories." (Li et al., 2019; Marsasi and Yuanita, 2023; Pedeliento et al., 2016; Hwang et al., 2019).

H2: There is a significant positive impact between customer engagement and brand attachment.

Brand loyalty is experiencing significant growth and is consideredа crucial factor for a company's success. It refers to the unwavering commitment of customers to repeatedly purchase a preferred product (Li et al., 2019; Marsasi and Yuanita, 2023).

H3: There is a significant positive impact between brand attachment and brand loyalty.

The brand attachment has a strong, positive relationship with brand loyalty ($p<0.05$). The supported empirical research are Centeno, Mandagi, Indrajit, (2022); Li, Teng, and Chen, (2020); Rathnayake, (2015); Cornelia and Pasharibu, (2020); Garanti, (2019). This is supported by social exchange theory.

H4: Brand attachment mediates the relationship between customer engagement and brand loyalty.

Brand attachment significantly mediates the relationship between customer engagement and brand loyalty. The supported research findings by Li et al., (2020); Kumar and Nayak (2019).

4. Theoretical Background and Models

4.1 Self-congruence Theory (Japutra et al, 2019)

The alignment of a brand's personality with a consumer's self-image (Tan et al., 2019;Hakmanage and Kumara, 2021) can lead to greater loyalty (Japutra et al., 2019; Hwang et al., 2019; Japutra et al., 2019; Hakmanage and Kumara, 2021).

4.2 Customer Engagement Theory (Kumar et al., 2019)

Loyalty drives positive emotions, leading to enhanced customer engagement (Pansari and Kumar, 2017) along with three dimensions; affective, cognitive, and behavioral (Kumar et al., 2019; Lim and Rasul, 2022).

4.3 Attachment Theory (Gilal et al., 2020)

Attachment theory proposes that brand attachment is a crucial determinant of consumer-brand outcomes (Gilal et al., 2020). Brand attachment has been redefined as a relationship between individuals and brands (Park et al., 2010; Jahn, Gaus, and Kiessling, 2012; Gilal et al., 2020).

4.4 Habit theory (Kim, 2019)

Brand loyalty may be a consequence of habit, as customers may repurchase a particular brand due to its familiarity and the reduced cognitive effort required (Kim and Koo, 2019).

5. Methodology

5.1 Research Philosophy

The objective is to answer the question of how customer engagement influences brand loyalty. The aim is to identify and verify relationships by using empirical evidence, indicating a positivistic research methodology (Singh and Srivastava, 2020; Jayalath and Galdolage, 2021).

5.2 Research Approach and Research Method

The study uses the deductive approach to recognize the relationship between existing theories based on concepts and by using hypotheses the relationships are empirically tested (Bougie and Sekaran, 2019; Pandey, 2019; Bandara, 2021). The study used quantitative research techniques influenced by the positivistic paradigm (Tinh et al., 2021; Bougie and Sekaran, 2019). Data was collected through a survey with validated scales and indicators based on prior research. The survey instrument was pilot tested before implementation.

5.3 Research Instrument Development

Customer engagement was operationalized using three indicators drawn from the scale used by Bergen & Brock (2019) using a Likert Scale ranging from Strongly Disagree (1) to Strongly Agree (5). Brand attachment was operationalized using two indicators used by Gilal et al (2020) using a Likert Scale ranging from Strongly Disagree (1) to Strongly Agree (5). Brand loyalty was operationalized using two indicators used by Wijekoon and Fernando (2020) using a Likert Scale ranging from Strongly Disagree (1) to Strongly Agree (5).

5.4 Sampling Design

The chosen sampling method is a simple random sampling method that is easy to eliminate bias and administer (Bell, 2022). The research targeted individuals above the age of 18 residing in the western province of Sri Lanka. A pilot test was conductedwith 27 respondents and the survey was distributed to 574 participants, resulting in 325 valid responses. The response rate of 56% was deemed acceptable based on prior research (Garritty, Gartlehner, and Streit, 2021).

6.Data Analysis and Representation

6.1 Normality and Missing Values

To test for the normality of the data, both the Shapiro-Wilk test and Kolmogorov-Smirnov tests were performed. The results of the statistical tests indicated that the distribution of the data can be considered normal, as the p-value was found to be greater than 0.05 (Souza et al., 2023).

6.2 Reliability Analysis

By analyzing the Cronbach Alpha value, the research instrument is assessed for Reliability (Schrepp, 2020). The research instrument is reliable because Cronbach Alpha value is 0.773 (Shrestha, 2021).

6.3 Validity

6.3.1 Sample Adequacy

Variable	Threshold Value	Cronbach Alpha
Pilot Study	$\alpha > 0.7$	0.842
Customer Engagement		0.713
Brand Attachment		0.721
Brand Loyalty		0.725

KMO test: KMO value = 0.589 (>0.5 threshold) (Jayalath and Galdolage, 2021; Goni et al., 2020), confirming sufficient sample. Bartlett's test: p-value < 0.05 (Jayalath and Galdolage, 2021), further supports sufficient sample size.

6. 3.2 Convergent Validity

Brand loyalty exists since the correlations of inter-item exceed 0.3 (Ibrahim and Ahmed, 2019; Berdugo et al., 2020; Hair et al., 2016; Aslam and Luna, 2021).

6. 3.3 Construct Validity

For construct validity exist Cronbach Alpha value of 0.842 met the threshold value of 0.7 that the indication successfully met the criteria (Jayalath and Galdolage, 2021; Aslam and Luna, 2021; Hair et al., 2016).

6.3.4 Discriminant Validity

Discriminant Validity exists since eigenvalues are less than 1.

7. Hypothesis Testing

H1: There is a significant positive impact of customer engagement on brand loyalty.

The regression analysis provides an estimation of the relationship among the key variables while using R values are used to do the analysis (Shrestha, 2020). Since the p-value is significant (p=0.000), there is a significant positive relationship between customer engagement and brand loyalty. However, the coefficient table shows a p value which is more than 0.05 and hence the relationship is not significant. Hence H1 is not supported.

H2: There is a significant positive impact between customer engagement and brand attachment.

Customer engagement is having a medium, positive relationship with brand attachment ($p < 0.05$). The supported empirical researches are Li, Teng, and Chen, (2020); Rabbanee, Roy, and Spence, (2020); Rabbanee, Roy, and Spence, (2020). This is supported by attachment theory. The coefficient table shows a p-value which is less than 0.05 and hence the relationship between customer engagement and brand attachment is significant. Therefore, H2 is supported.

H3: There is a significant positive impact between brand attachment and brand loyalty.

Brand attachment has a strong, positive relationship with brand loyalty ($p<0.05$). The supported empirical research is Centeno, Mandagi, Indrajit, (2022); Li, Teng, and Chen, (2020); Rathnayake, (2015); Cornelia and Pasharibu, (2020); Garanti, (2019). This is supported by social exchange theory. The coefficient table also reveals that the p value is statistically significant and hence there is a significant and positive impact of brand attachment on brand loyalty. Therefore, H3 is accepted.

ANOVA[a]						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	23.856	1	23.856	113.489	.000[b]
	Residual	67.685	322	.210		
	Total	91.541	323			
a. Dependent Variable: Brand_Loyalty						
b. Predictors: (Constant), Brand_Attachment						

H4: Brand attachment mediates the relationship between customer engagement and brand loyalty.

The P value is significant. Therefore, hence Brand Attachment significantly mediates the relationship between Customer engagement and Brand Loyalty. H4 is supported.

Input:			Test statistic:	Std. Error:	*p*-value:
a	0.476	Sobel test:	5.91336862	0.04894131	0
b	0.608	Aroian test:	5.89340263	0.04910711	0
s_a	0.063	Goodman test:	5.93353893	0.04877494	0
s_b	0.064	Reset all	Calculate		

8. CONCLUSION & RECOMMENDATION

8.1 Discussion of Research Objectives

RO1: To understand the influence of Customer Engagement onB rand Loyalty

Customer engagement is having a weak, positive relationship with brand loyalty in the fashion clothing industry which established mean value is greater than the Mean value ($p < 0.05$). Much of the research is done in industries such as banking, hotel, smartphone, skincare, and hospitality. The supported empirical research is Li, Teng, and Chen, (2020); Shouk and Soliman, (2021); Kaur et al., (2020); Guizon and Magnoni, (2019); Abror et al., (2019); Monferrer, Moliner, and Estrada, (2019); Dissanayaka and Ragel, (2023) that supported by customer engagement theory. The regression analysis showed that brand attachment and customer engagement predict brand loyalty 26.2% of the time. The purpose of buying a product through digital platforms is affected by price, culture, design, and quality (Pedeliento et al., 2016; Whan Park et al., 2010; Hwang et al., 2019) that supported by self-congruence theory.

RO2: To assess the impact of Brand Attachment on Brand Loyalty

The brand attachment has a strong, positive relationship with brand loyalty which established mean value is greater than the Mean value ($p<0.05$). The supported empirical research is Centeno, Mandagi, Indrajit, (2022); Li, Teng, and Chen, (2020); Cornelia and Pasharibu, (2020); Garanti, (2019).

RO3: To identify the affect of Customer Engagement on Brand Attachment

Customer engagement is having a medium, positive relationship with brand attachment which established mean value is greater than the Mean value ($p < 0.05$). The supported empirical researches are Li, Teng, and Chen, (2020); Rabbanee, Roy, and Spence, (2020); Rabbanee, Roy, and Spence, (2020).

RO4: To identify the mediating effect of brand attachment onc ustomer engagement and brand loyalty.

Brand attachment significantly mediates the relationship between customer engagement and brand loyalty. The supported research findings by Li et al., (2020); Kumar and Nayak (2019).

Brand attachment generates the maximum contribution to the dependent variable, Brand loyalty based on the beta value of 0.494. The supported past findings are Tseng, (2020); Chieng and Roy, (2022). There is a moderate correlation and supported by regression findings as well. Regression analysis exposed customer engagement and brand attachment and predicted brand loyalty by 51%. The supported past finding by Lia, Teng, and Chen, (2020) that exposed in a study among 298 Taiwanese respondents.

The bell curves the data distribution is normal which indicates the data set was appropriate for further analysis. The KMO value is greater than 0.5 and hence the sample collected the data was considered adequate.

The below regression equation recommended is derived based on conducted analysis.

Brand loyalty = 1.322 + 0.608 * (Mean of brand attachment - 0.062 * (Mean customer engagement)

9. Conclusion

This study explored customer engagement's impact on brand attachment and brand loyalty in Sri Lanka. Results revealed a weak, positive link between customer engagement and brand loyalty, mediated by brand attachment. Brand attachment had a strong, positive association with brand loyalty. Additionally, customer engagement had a moderate, positive relationship with brand attachment. These findings are supported by theories. To foster brand loyalty, fashion brands should enhance customer engagement and encourage brand attachment. Limitations include geographical scope and time constraints.

10. REFERENCES

[1] Bandara, K. B. T. U. K. 2021. Brand Romance: An Upright Predictor of Brand Loyalty in Mobile Phone Market, based on University Academics in Sri Lanka. *J. ICBI.* DOI: http://dx.doi.org/10.2139/ssrn.3862940

[2] Bell, E. (2022). Business research methods. Oxford University Press.

[3] Bougie, R. and Sekaran, U. (2019). Research methods for business: A skill building approach. John Wiley and Sons.

[4] Harsandaldeep Kaur, Mandakini Paruthi, JamidUl Islam, and Linda Hollebeek. 2020. The role of Brand Community Identification and reward on consumer brand engagement and brand loyalty in Virtual Brand Communities. *J.46.* DOI: https://www.sciencedirect.com/science/article/pii/S0736585319308135

[5] Khyati Shetty and Jason R. Fitzsimmons. 2021. The effect of brand personality congruence, Brand Attachment and Brand Love on loyalty among Henry's in the luxury branding sector. *J. FMM,* (March 2022), 1361-2026. DOI: https://doi.org/10.1108/JFMM-09-2020-0208

[6] Sebastian Molinillo, Rafael Sánchez, and Francisco Cabanillas. 2020. Analyzing the effect of social support and community factors on customer engagement and its impact on loyalty behaviors toward social commerce websites. *J. Elsevier.* DOI:https://www.sciencedirect.com/science/article/pii/S074756321930144X

[7] Sudarsan Jayasingh. 2019. Consumer brand engagement in social networking sites and its effect on brand loyalty. *J CBM 6.* DOI: https://www.tandfonline.com/doi/epdf/10.1080/23311975.2019.1698793?needAccess=true&role=button

[8] Tharoosha Jayalath and Badra Galdolage. 2021. Customer Engagement in Private Sector Healthcare: How Does it Affect Customer Loyalty. *J. EMT 27, 2,* (June 21), 10 pages. DOI: 10.9734/JEMT/2021/v27i230328

[9] Thuseethan, S. and Kuhanesan, S. 2015. Influence of facebook in academic performance of Sri Lankan University students. *J. SSRN* 14, 4 (August 2014), 9 pages. DOI: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2478336

[10] Way Lia, Yu Tengb, and Yu Chena. 2020. Unlocking the customer engagement-brand loyalty relationship in tourism social media: The roles of brand attachment and customer trust. *J. HTM* 44, (2020), 184-192. 9 pages. DOI: https://e-tarjome.com/storage/btn_uploaded/2020-11-15/1605418866_11581-etarjome%20English.pdf

iCIIT 2023

THE 4TH INTERNATIONAL CONFERENCE ON INNOVATIONS IN INFO-BUSINESS & TECHNOLOGY

Organised by

INFORMATICS INSTITUTE OF TECHNOLOGY

Platinum Sponsor

SLTMOBITEL The Connection

Gold Sponsor

AXIATA DIGITAL LABS

Silver Sponsor

ZONE24x7

Bronze Sponsors

NCINGA ENGINEERED FOR HUMANITY

Limark

Technical Co-sponsors

ETLTC

acm Chapter

Event Partners

bcs Sri Lanka Section

CSSL COMPUTER SOCIETY OF SRI LANKA

FITIS

ISACA Sri Lanka Chapter

SLASSCOM THE KNOWLEDGE AND INNOVATION CHAMBER

Print Media Sponsor

Daily FT

the SUNDAY TIMES

the SUNDAY TIMES EDUCATION TIMES THE WISDOM UPGRADE

Electronic Media Sponsor

Hi TV The Luxury Channel

Hospitality Patner

GALADARI COLOMBO

9 773021 603000

Scan the QR to Connect with ICIIT 2023