SCU
SLIIT CITY UNI

PROCEEDINGS OF THE
4th ANNUAL RESEARCH CONFERENCE OF SLIIT CITY UNI

ARCSCU'25

ANNUAL RESEARCH CONFERENCE OF SLIIT CITY UNI

8TH JULY 2025

https://arcscu.sliitcityuni.lk/

ARCSCU 2025

Proceedings of the
ANNUAL RESEARCH CONFERENCE
OF SLIIT CITY UNI

08 JULY 2025

Disclaimer

The views expressed in these papers in these proceedings are of those of the authors. The statements and the opinions stated either in the papers published herein or add the conference and discussions do not necessarily represent the views of SLIIT City UNI.

Proceedings of the Annual Research Conference of SLIIT City UNI.

A publication of SLIIT City UNI

ISSN 3093-5229

Printed from the camera-ready copy supplied by the authors.

ARCSCU 2025 – Committee

Program Chair	Dr. Nipunika Vithana
Program Co-Chairs	Dr. Yasas Jayaweera
	Dr. Gayana Fernando
Technical Program Chair	Ms. Nideshika Ellepola
Secretary	Mr Heshan Gallage
Advisory Board	Prof. Lakshman Ratnayake
	Prof. Lalith Gamage
	Prof. Mahesha Kapurubandara
	Prof Indra Dayawansa
Organizing Committee	Ms Dilushinie Fernando
	Mr Vibhavi Attigala
	Mr Gishan Abhayagunarathna
	Mr Anuruddha Abesinghe
	Ms Raviharee Weerasinghe
	Ms Charunika Dampalessa
	Mr Pubudu Nallaperuma
	Ms Rameesha Kankanamage
	Mr Ramanaruban Rangan
	Mr Chamith Jayasinghe
	Ms Samudini Nagasena

	Ms Maneesha Aishcharya
	Ms Ishadi Nilaweera
	Ms Eranga De Silva
	Mr Chan Sri Manukalpa
	Mr Tharuka Perera
	Ms Jayani Kaumadie
	Ms Maria Paulraj
	Ms Priyanka Thirumugam
	Ms Kushani Bandara
	Mr Pasindu Gunawardhana
	Mr Benorith Lawrence
	Mr Josheph Lakshan
	Mr Dhanushka Madushanka
	Mr Jerith Lawrence
Web Designer	Mr. Ganidu Sathishka
Graphic Designers	Mr. Akila Migara
	Mr Matheesh Adithya

ARCSCU 2025 – Review Panel

Ms Umanga Pilapitiya, Curtin Colombo

Mr Pumudu A. Fernando, Informatics Institute of Technology

Miss. M. Virajini Niroshika Godapitiya, ESOFT Metro Campus, Colombo

Mr. Samadhi Rathnayake, Sri Lanka Institute of Information Technology

Miss Tharushi Himasha Rubasinghe, Sri Lanka Institute of Information Technology

Ms Sankaja Amaraweera, Sri Lanka Institute of Information Technology

Ms Lawanya Abesooriya, Wayamba University

Mr Chandula Arsakulasuriya, University of Ruhuna

Ms Keshani Wijesuriya,Sri Lanka Institute of Information Technology

Ms Ovinie Seneviratne, University of Sri Jayewardenepura

Ms Vidya S Nair, Amrita Vishwa Vidyapeetham, India

Dr Sanjaya Jayasooriya, Sir John Kotelawala Defence University

Ms Manori Gamage, Sri Lanka Institute of Information Technology

Table of Contents

Keynote Speakers

From Filters to Futures: Empowering the Next Generation to Shape a Smarter World with AI

Prof. Mahesha Kapurubandara
Pro Vice-Chancellor – International, SLIIT International

This talk explores how today's youth—often introduced to AI through everyday tools like social media filters—can grow into tomorrow's innovators, using artificial intelligence to solve real-world challenges. Designed to inspire and inform, the session highlights the accessible entry points into AI, the transformative potential it holds across industries, and the crucial role young minds will play in shaping an inclusive, intelligent future.

Exploring Unconventional Computing: experiments in biological computing

Dr. Edward Braund
Head of School of Computer Science & Technology,
University of Bedfordshire

Ed's research examines the information processing abilities of biological systems to develop bio-processors, sensors, and actuators. A core focus of his work is to create prototypes that are fit for deployment outside of the laboratory and in the real world. Currently, Ed is exploring the challenges and benefits that biotechnologies may bring to fields that demand minimal or reduced human intervention. This talk will explore Ed's work and look at the future of computing

Intersection of capitalism and sustainable future

Mr. Kalana Muthumuni
CEO and Co-Founder of Hyperglade

Kalana will be talking about how digital innovation has paved the way for better systems. The presentation will highlight the Gaps in Legacy Financial systems and how Fintech are filling in, room for growth in the current economic landscape. Also extend to discuss the Intersection of for profit models in social enterprises to create an impact on society

Track 1: Information Technology

Adaptive Multi-model Machine-Learning and AI Systems for Strengthening the Emotional Well-being of Children with Trisomy 21

M.I.D.C. Balasuriya
Department of Computer Science and Technology of Computing
University of Bedfordshire, United Kingdom
dilshanicbalasooriya15@gmail.com

N. Ellepola
Faculty of Computing
SLIIT, City Uni, Colombo 03, Sri Lanka

***Abstract*- This research study demonstrates a web application designed to strengthen children's cognitive skills and emotional well-being with Trisomy 21, utilizing interactive and tailored tools. Trisomy 21 is a chromosomal anomaly caused by an extra copy of the 21st chromosome, which affects a child's cognitive development. Despite the technological evolution, a significant gap persists in accessibility and multimodal approaches that meet the unique needs of children with Trisomy 21. The main objective of the research study was to develop a multi-model web-based application, "Mockley kids," customized for children with Trisomy 21 that helps to enhance their cognitive skills and emotional well-being. The developed system integrated an artificial intelligence (AI) powered voice assistance to enhance communication, and learning, an emotion-based music recommender to enhance emotional well-being and provide a calm and uplifting environment, a text-based bot to enhance literacy skills and communication, and interactive games "Who am I?", "Tic-tac-toe", "Simon-says" to increase attention span, decision making, which are tailored to enhance their cognitive skills and emotional well-being. The development and implementation of this project follows a structured process aligned with Agile project methodology. To evaluate the "Mockley kids" system's impact on children with Trisomy 21, 16 children who were diagnosed with Trisomy 21 engaged with the system for 20 minutes for a week, under the supervision of 8 professionals, including 3 speech therapists, 3 occupational therapists, and 2 educators. Overall results show that the children were excited to integrate with the system and enjoyed the system. Both the professionals and the parents stated that they had evident noticeable improvements in cognitive abilities, including enhanced communication, memory recall, enhanced attention span, and improvements in emotional well-being.**

***Keywords*- Trisomy 21, Assistive Technology, Artificial Intelligence (AI), Cognitive development, Emotion-based music recommender**

I.INTRODUCTION

Trisomy 21, also known as "Down syndrome," is a common genetic disorder caused by a chromosomal anomaly that leads to intellectual disabilities and developmental delays [1]. Children with Trisomy 21 often face diverse challenges, including communication, social engagement, and emotional challenges. Eddaoudi et al. have pointed out the psychological, familial, and social challenges experienced by the parents of children with Trisomy 21 [2], and Er-rida et al. highlighted the challenges faced by parents of children with Trisomy 21 while providing education to their children at different age levels [3], emphasizing the need for tailored and more inclusive support beyond clinical support.

Even though technological transformation provides different innovative tools to improve learning, communication, and emotional engagement for children with intellectual disabilities, Shahid et al. highlight the gap in tailored applications for individuals with special needs, pointing out the need for further research on emerging technologies [4]. Fritz et al. highlight how technology enhances the independence of children with Trisomy 21 and helps families adapt to support their daily functions [5].

Gemma et al. demonstrate that music is beneficial for individuals with Trisomy 21 [6], and Steinberg et al. demonstrate that integrating music into a daily routine shows a positive impact on emotional bonding and communication [7]. Kasuya-Ueba et al. further supported and highlighted the ability of music-based intervention in enhancing emotional bonding, attention, and engagement [8].

Mateos-Sanchez et al. introduced "CapacitaBOT," a mobile chatbot tailored for children with intellectual disabilities. Research aimed to enhance social skills through interactive learning. Furthermore, this study highlights the effectiveness of AI-powered tools in education and therapy, especially during periods of social isolation like the COVID-19 pandemic [9]. Despite the valuable outcome from the CapacitaBOT, it primarily focuses on text-based chatbot interactions with limited

customization for children with Trisomy 21 and is only focused on social skills and education. In contrast, the “Mockley kids” is a web-based system that focuses on roviding emotional, social, educational, and cognitive support with a customizable experience focusing on different levels ofTrisomy 21 and different age levels. Panceri et al. introduced the “MARIA T21,” a socially assistive robot, focused on improving the psychomotor and psychosocial skills in children with Trisomy 21 and autism spectrum disorder (ASD). A series of game-based on behavioral psychology was integrated into “MARIA T21” to prove that child-robot interactions can significantly strengthen the engagement [10]. Accessibility of the MARIA T21 is limited by specialized hardware and focused on providing social skills, motor skills, and cognitive skills. In contrast, “Mockley kids” addresses these gaps by offering a fully digital, customizable support focusing on emotional, social, educational, and cognitive aspects. Costanzo et al. introduced “Talkitt”, an AI-powered mobile application developed to improve communication of individuals with Trisomy 21. Through a six-month intervention, the application depicts notable enhancements in linguistic skills, suggesting that AI-based augmentatives and Alternative Communication (AAC) tools foster better communication [11]. Although “Talkitt” provides a powerful speech recognition solution to aid individuals with unintelligible speech scope is limited to communication. “Mockley kids” system focuses on providing both emotional and cognitive support with a customizable environment for different levels of children with Trisomy 21.

Kokol et al. shows a collective study of 145 Studies on game-based interventions for children with developmental disabilities. This study noticed a positive outcome in areas such as anxiety reduction, emotion recognition, and rehabilitation [11].

Despite the presence of technological interventions for children with Trisomy 21, a gap persists in the accessibility of existing systems and the absence of multimodal, tailored interventions that provide combined support of cognitive and emotional aspects. Existing systems often focused on specific aspects such as emotional well-being or education, without addressing the comprehensive emotional and cognitive needs of children with trisomy 21. This research study intends to bridge the gap by developing an adaptive multimodal web-based application called “Mockley kids”, integrated with an AI-powered voice assistance, text-based bot, emotion-based music recommender, and interactive games to strengthen both cognitive skills and emotional well-being. This research study focuses on the key research question: how a customized multimodal web application integrated with AI-powered voice assistance, text-based bot, emotion-based music recommender, and interactive games strengthens the emotional well-being and cognitive skills of children with Trisomy 21. The primary aim of this research study was to develop an adaptive, multimodal web application that strengthens the cognitive skills and emotional well-being of children with Trisomy 21. The research objectives include developing a web application integrated with an emotion-based music recommendation model, voice assistance feature, text-based bot, and interactive game features, developing a responsive and interactive user interface, and testing the system's efficiency.

II.SYSTEM DESIGN

The developed system is a multimodal web application designed to enhance cognitive skills and provide emotional support for children with Trisomy 21 through a modular integration of several advanced and customized technologies. The developed system includes an AI-powered voice assistant, an emotion-based music recommender model, a text-based bot, interactive games “Who am I?”, “Tic-tac-toe”, and “Simon says” games. The system adopts a client-server model with the front end built with React.js and the backend developed with Python using Flask to manage API endpoints. AI-powered voice assistance uses speech recognition libraries to convert voice inputs to text, processes through a server, and returns the voice output using text-to-speech tools such as gTTS and Pydub. The emotion-based music recommender model captures the real-time emotions using webcam, and the captured image passes through a convolutional neural network (CNN) developed with Keras and TensorFlow, and OpenCV and trained using the FER-2013 (Facial Expression Recognition 2013) data set, which contains 7 different emotions (angry, disgust, fearful, happy, neutral, sad, and surprised). Upon emotion detection, the recommended songs list is displayed as a JSON object to the frontend. “Who am I?” and “Hangman” games were developed using Python Flask for the backend and React.js for the frontend, communicating through APIs. The games “Simon says” and “Tic Tac Toe” were implemented using React.js and React hooks for state management. The text-based bot generates the responses using the “Gorq” API. AI-powered voice assistance generates AI-powered responses using the Gemini API. Data was managed using CSV files, and the Pandas library and environment file were used to store API keys securely. The system architecture ensures modularity, responsiveness, and scalability, enabling and ensuring the smooth integration of diverse modules with diverse technologies to make the systems more efficient and customized to meet the unique needs of children with Trisomy 21.

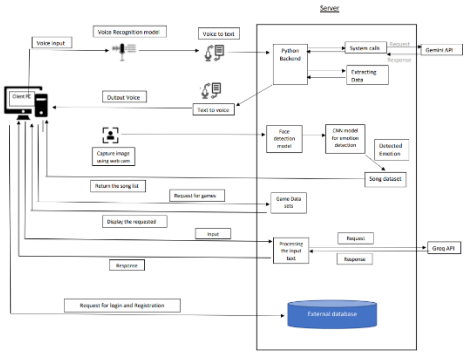


Fig. 1. Software Architectural Diagram

III.METHOD

A. System Evaluation

An observational evaluation was conducted with 16 children who were diagnosed with Trisomy 21 at a rehabilitation hospital, Sri Lanka. Each child engaged with the system for 20 minutes per day for a week. Sessions were conducted under the supervision of 8 professionals, including 3 speech therapists, 3 occupational therapists, and 2 special educators.

- Data Collection

Data collection was conducted using a structured questionnaire designed to gather feedback from both the parents of children who were involved in the evaluation sessions and professionals who supervised the sessions. The questionnaire assessed multiple dimensions, including cognitive support, emotional support, impact on motor skills accessibility, user friendliness, and overall effectiveness and satisfaction with the system. All collected data were analyzed to examine the effectiveness of the system in strengthening emotional well-being and cognitive skills, accessibility, useability, and aspects to enhance the system. Additionally, observational notes were collected from the professionals about the system for future enhancements.

IV.RESULTS

A. Evaluation of Professionals' Perspective

Professionals stated that the web application provides long-term benefits for children with Trisomy 21 in terms of cognitive development and emotional well-being. On a 5-point scale, the system received an average of 3.5 for improving memory recall, 4.13 for the capability to enhance the attention span, 2.88 for improving problem-solving skills, 3.75 for improving communication and comprehension skills, and 3.63 for enhancing decision-making and logical thinking. Lastly, 3.00 for enhancing social interaction. Furthermore, professionals commented that gamified models stimulate children's learning, leading to cognitive development [11]. Emotionally, the system was significantly effective. Professionals rated the system with 4.75 for providing emotional support, evidencing that music provides emotional support and an uplifting environment [6].

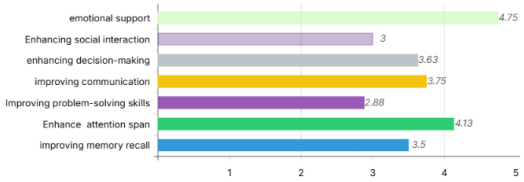


Fig. 2. Analysis of professionals' comments (Rating out of 5).

B. Evaluation of Parents' Perspective

Most parents commented that their children showed a strong interest in interactive games. 56.3% stated that they were deeply captivated, while 25% were moderately interested and 18.8% showed a mild interest, evidence that game-based intervention improves the emotional well-being of children with Trisomy 21 [11].

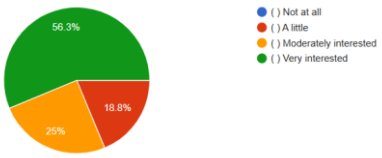


Fig. 3. Evaluation of interest in engaging with games

87.5% of parents commented that children were actively interacting with the AI-powered chatbot and revealed that this feature improves cognitive skills and emotional well-being of children with Trisomy 21. Furthermore, 100% of the respondents commented that their children show emotions while interacting with the chatbot, evidencing a positive impact on emotional well-being. 93.7% of parents declare that AI-powered chatbot encourages their child to communicate more than usual [9]. Among them, 25% of the children show significant improvements in communication, while 43.5% show notable improvements. 93.8% of parents declared that children learn new words while interacting with the chatbot.

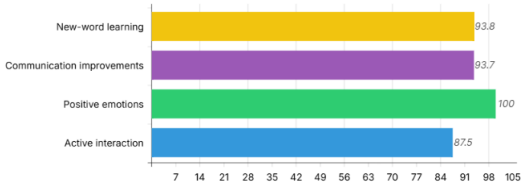


Fig. 4. Evaluation of AI-chatbot interaction

Moreover, 18.8% of parents strongly declare that the emotion-based music recommender feature improves their child's mood, and 62.5% declare a notable enhancement, while 18.8% declare a slight enhancement. Furthermore, the data collected through the questionnaires reveals that 100% of the children enjoyed the music, confirming that music therapy improves the emotional well-being of children with Trisomy 21[8].

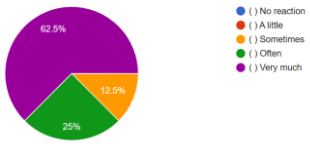


Fig. 5. Evaluation of showing emotions while interacting with the music recommender

The application appears to have a positive influence on cognitive skills enhancement. 93.7% of parents remarked that after system engagement, children were able to recognize and recall system interactions, such as words, phrases, and instructions. Furthermore, after interacting with the application, children show more interest in problem-solving and answering questions.

V.CONCLUSION

A. Findings of the Study

The "Mockley kids" application shows a significant capability in strengthening the emotional well-being and cognitive skills of children with Trisomy 21. According to the professionals, the key features AI-powered chatbot, emotion-based music recommender, and interactive games were the features enjoyed by the children the most. Furthermore, the system enhances children's memory recall, attention span, problem-solving skills, and language and communication comprehension. Many professionals and parents commented on significant improvements in children's ability to identify emotions, problem-solving skills, and learning new words and phrases. The emotion-based music recommender feature was compatible with a child's mood and effectively provided a calm and uplifting environment. Several professionals declared confidence in the application's long-term benefits in strengthening cognitive skills and emotional well-being, making the system an invaluable tool in the therapeutic environment.

B. Limitations of the Study

Although the web application is an invaluable tool for children with Trisomy 21, certain limitations should be acknowledged. These limitations might affect the generalizability of the results and the highlighted areas for future enhancements and research. Several professionals declared that to observe a 100% accurate and measurable outcome, a minimum of 3 months' period is required, while this study covers a limited timeframe. The system lacks a progress tracker, making it challenging to obtain an accurate result over time. Additionally, the system was developed using English as the key language, creating a gap in accessing the application for children who speak other languages.

C. Future Work and Recommendations

To enhance the effectiveness of the web application, integration of a built-in progress tracker and detailed dashboard is needed to accurately evaluate the progress. Implementing multi-language support to bridge the gap in accessibility. A collaborative feature to interact with peers within the system will enhance the value of the system. Furthermore, a one-week evaluation period limits the ability to assess the accurate long-term effects of the system, so future studies need to extend 1-3 months with pre-/post-assessment and a control group and need to be focused on multiple sites to validate and generalize findings.

ACKNOWLEGEMENT

I would like to give my heartfelt gratitude to all those who supported me throughout the research. Furthermore, I am grateful to all the professionals for their valuable insight and contribution to the evaluation process and the market research. Lastly, I would like to express my heartfelt gratitude to all the children and parents who participated in the market research and the evaluation.

ETHICAL STATEMENT

This study did not collect or store any sensitive or health-related data from the participants. Only the basic account information, including usernames, email addresses, and passwords, was required and stored during the study. Prior to the evaluation, informed consent was obtained from the parents/guardians of all the participants to ensure ethical compliance.

REFERENCE

[1] Cleveland Clinic, "Down Syndrome," Cleveland Clinic, Jan. 31, 2023. https://my.clevelandclinic.org/health/diseases/17818-down-syndrome

[2] S. Eddaoudi*, M. Zouine, A. Mehdaoui, K. Khabbache, Y. El-Boussaadni, and A. Oulmaati, "Challenges and Concerns of Parents with Children with Down Syndrome," Archives of Psychiatry and Mental Health, vol. 7, no. 1, pp. 015–017, Jul. 2023, doi: https://doi.org/10.29328/journal.apmh.1001048.

[3] S. Er-rida, M. Oubibi, M. Mafhoum, M. H. Alami, and A. M. Alaoui, “Challenges Faced by Parents of Children with Down Syndrome in Mainstream Schools: Exploring Inclusive Education,” The Open Psychology Journal, vol. 18, no. 1, Feb. 2025, doi: https://doi.org/10.2174/0118743501360190250101113909.

[4] N. M. I. Shahid, E. L.-C. Law, and N. Verdezoto, “Technology-enhanced support for children with Down Syndrome: A systematic literature review,” International Journal of Child-Computer Interaction, vol. 31, p. 100340, Jul. 2021, doi: https://doi.org/10.1016/j.ijcci.2021.100340.

[5] M. N. Fritz, “The Impact of Technology on Individuals with Down Syndrome and Their Families,” ScholarWorks@UARK, 2017. https://scholarworks.uark.edu/rhrcuht/49 (accessed Jun. 25, 2025)

[6] M.-G. Gemma, M.-C. Pablo, and A. Cabedo-Mas, “The role of music in the development of children with Down syndrome: a systematic review,” Interdisciplinary Science Reviews, vol. 45, no. 2, pp. 158–173, Apr. 2020, doi: https://doi.org/10.1080/03080188.2020.1755556.

[7] S. Steinberg, C. M. Shivers, T. Liu, L. K. Cirelli, and M. D. Lense, “Survey of the home music environment of children with various

[8] developmental profiles,” Journal of Applied Developmental Psychology, vol. 75, p. 101296, Jul. 2021, doi: https://doi.org/10.1016/j.appdev.2021.101296.

[9] Y. Kasuya-Ueba, S. Zhao, and M. Toichi, “The Effect of Music Intervention on Attention in Children: Experimental Evidence,” Frontiers in Neuroscience, vol. 14, no. 757, Jul. 2020, doi: https://doi.org/10.3389/fnins.2020.00757.

[10] M. Mateos-Sanchez, A. C. Melo, L. S. Blanco, and A. M. F. García, “Chatbot, as Educational and Inclusive Tool for People with Intellectual Disabilities,” Sustainability, vol. 14, no. 3, p. 1520, Jan. 2022, doi: https://doi.org/10.3390/su14031520.

[11] J. A. C. Panceri, É. Freitas, J. C. de Souza, S. da Luz Schreider, E. Caldeira, and T. F. Bastos, “A New Socially Assistive Robot with Integrated Serious Games for Therapies with Children with Autism Spectrum Disorder and Down Syndrome: A Pilot Study,” Sensors, vol. 21, no. 24, p. 8414, Jan. 2021, doi: https://doi.org/10.3390/s21248414.

[12] F. Costanzo et al., “Talkitt: toward a new instrument based on artificial intelligence for augmentative and alternative communication in children with down syndrome,” Frontiers in Psychology, vol. 14, Jun. 2023, doi: https://doi.org/10.3389/fpsyg.2023.1176683.

CacheBook: A Smart Personal Finance Tracker with OCR-Based Expense Logging and Visual Reporting

P.T. Perera
Department of Computer Science and Technology of Computing
University of Bedfordshire, United Kingdom
tharushi@outlook.com

V.N Vithana
Faculty of Computing
Sri Lanka Institute of Information Technology, City Uni, Colombo 03, Sri Lanka
Nipunika.v@sliit.lk

Abstract- "CacheBook" is a mobile application designed to simplify personal finance management by integrating real time expense tracking with automated receipt scanning. Built using React Native and Firebase, the application enables users to log daily expenses, set financial goals, and visualize spending through reports. A key innovation of "CacheBook" is the use of Optical Character Recognition (OCR) technology, which allows users to scan physical receipts by reducing manual entry and enhancing data accuracy. The app categorizes expenses, provides visual summaries via charts, and alerts users when spending goals are exceeded.

Keywords—OCR, Personal Finance Tracker, Mobile, React Native

I INTRODUCTION

"CacheBook" is a mobile application that simplifies personal financial management by allowing users to record expenses, set financial goals, and generate graphical reports. One of the most unique features of the system is its ability to scan and read data from receipts using Optical Character Recognition (OCR), reducing manual input and enhancing accuracy.

Managing Personal finances has become increasingly difficult in today's fast, digitally driven world. Many individuals struggle to keep track of where their money goes, stick to budget, or identify spending patterns. These challenges are often made worse by the lack of accessible, user-friendly financial tools.

The development of "CacheBook" was initiated in response to these issues, with the goal of offering a practical solution that empowers users to take control of their finances through real-time tracking and data-driven insights. The system provides users with features to classify their expenses, monitor daily or monthly budgets, and receive alerts when spending exceeds set targets. Ultimately, the aim is to improve financial literacy and decision making with an easy-to-use interface and intelligent automation.

II LITERATURE REVIEW

Managing personal finances has become more complex in today's fast paced digital environment, prompting the rise of numerous technological solutions designed to help individuals budget, track expenses, and make informed financial decisions. Tools like Optical Character Recognition (OCR), real time alerts, and AL based categorization have transformed how users interact with their financial data.

Automated Bill Scanning using OCR technology allows users to extract itemized data from receipts, minimizing manual entry and increasing accuracy. Applications like Expensify use this technology effectively, although challenges such as poor handwriting and low-quality images still affect data reliability.

Customizable Categories improve personalization. Many applications like Mint and YNAB after predefined categories, but users often prefer flexibility to create and manage their own. This customization supports more relevant and organized budgeting experiences.

Dynamic Budgeting and Real Time Notifications help users stay within financial limits. Apps like PocketGuard and Goodbudget allow users to set budgets and send alerts when thresholds are exceeded. However, over notification can reduce user engagement, so timing and relevance must be carefully designed.

Visual Expense Reporting is essential for users to understand their financial behavior. Applications like QuickBooks and Wave provide charts and summaries, but complex layouts or overly technical language can hinder usability. Simplified, color coded reports are more effective for general users.

Despite the existence of many apps, common issues such as limited customization, poor user interface, and lack of integration between features remain. Tools like Mint,

YNAB, Goodbudget, and Expensify provide useful functionality but fall short in delivering a fully integrated, flexible and user-friendly experience.

“CacheBook” aims to bridge these gaps by integrating all these features into a single, easy to use mobile application. By addressing usability and functionality shortcomings found in current systems, “CacheBook” contributes to the evolving field of smart personal finance tools.

Theoretical Background and Research Gap

Financial literacy has been widely recognized as essential for improving personal financial behavior (Lusardi & Mitchell, 2014). Existing applications often focus on budgeting or expense tracking in isolation, without integrating automation and budgeting or expense tracking in isolation, without integrating automation and customization in a user-centered design. While tools like Mint and Goodbudget offer basic tracking, they lack real-time OCR data capture and personalization. “CacheBook” addresses this gap by combining automated receipt processing, flexible categorization, and dynamic budgeting within a single platform tailored to diverse user needs.

III METHODOLOGY

The development of the “CacheBook” mobile application followed the Agile Software Development Life Cycle, which emphasizes flexibility, iterative progress, and regular user feedback. Agile was chosen to ensure that each feature could be developed, tested and refined in short cycles called sprints, improving responsiveness to user needs and technical challenges.

The process began with requirement gathering, which included competitor analysis and a user survey with 42 participants. The survey helped identify key user expectations such as receipt scanning, customizable categories, and budget alerts. Based on these findings, the system’s features were prioritized and organized across multiple sprints.

Design and development were carried out using:

Frontend: React Native with Expo for cross platform compatibility.

Backend: Firebase services including Firestore (database), authentication, and storage.

OCR Integration: Mindee API

Development Tools: Visual Studio Code, Android Studio, Expo CLI

UI Styling: Tailwind CSS

The implementation focused on key features such as secure user authentication, manual and OCR based expense logging, goal tracking with alerts, and report generation with export to PDF. Real-time database updates ensured that financial data remains consistent and accessible.

Each sprint included testing, UI refinement, and feedback collection to ensure high usability and functionality. This approach allowed “CacheBook” to evolve into a reliable, user centered financial tracking application aligned with its core objectives.

OCR Evaluation and Accuracy Measurement

To evaluate OCR performance, nearly 100 sample receipts were collected, including printed, low resolution, and handwritten receipts. The Mindee API extracted line items, and outputs were manually validated against ground truth data. Accuracy was measured as the percentage of correctly identified line items. Results showed an average accuracy of 91% printed receipts and 58% for handwritten receipts.

Firebase Implementation Details

Firebase Authentication handled user login and session management. Firestore rules were configured to access by user ID, ensuring that users could only access their own data. Real time syncing was tested by simulating simultaneous expense entries on multiple devices. Error handling routines displayed notifications in case of network failures, invalid data input, or storage errors.

User Study Process

Usability testing involved 21 participants who used “CacheBook” over a two-week period. Participants were asked to record expenses manually and use OCR, set budget goals, and generate reports. Data collected included task completion time, number of errors, and subjective satisfaction ratings. Feedback was gathered via structured questionnaires and analyzed to identify usability issues and feature effectiveness.

IV RESULTS AND DISCUSSION

The “CacheBook” mobile application was successfully implemented, with all major features functioning as intended. The system provides secure user authentication, receipt scanning via OCR, customizable expense categories, dynamic budget goal tracking, visual expense reporting, and PDF export functionality.

Key Results:

OCR based Receipt Scanning: User can scan receipts using the device camera. The Mindee API extracts line-item data such as item names, prices, and quantities, which are then

categorized automatically using a keyword matching algorithm.

Custom Expense Logging: User can manually add expenses, create new categories, assign color codes, and set custom dates.

Spending Goal Alerts: Daily and weekly goals can be set and are monitored in real time. Users receive notifications when their expenses exceed the set threshold.

Visual Reports and Export: The app displays expenses in pie charts grouped by week, month, or year. Reports can be exported as PDF files, enabling users to share or store their financial summaries.

User Interface Snapshot:

Index Page: Clean and welcoming start screen with pastel gradients and finance related items.

Signup/Login: User friendly form layouts with password visibility toggle and optional Google Sign-In (Partially implemented).

Home Page: Figure 1 shows the personalized greeting, goal slider, and a list of today's expenses.

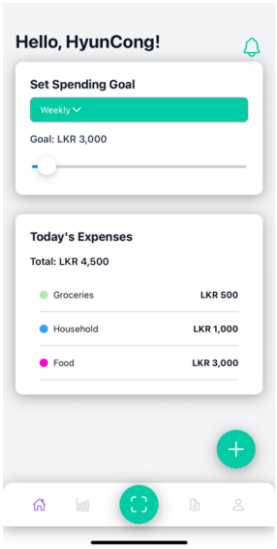


Fig 1 Image of Home Page

Add Expense Page: Manual entry interface with category selection and color assignment.

Bill Scan Page: Figure 2 shows the bill scanning with real time OCR extraction and auto categorization.

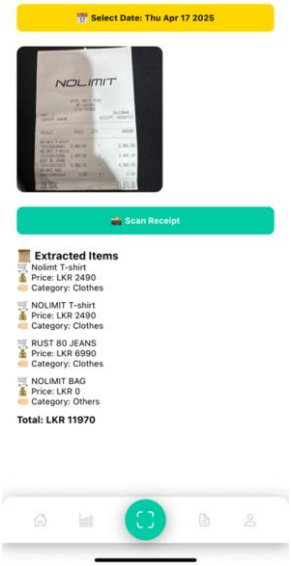


Fig 2 Image of bill scanning function

Report Page: Figure 2 shows interactive charts with a download button for PDF reports.

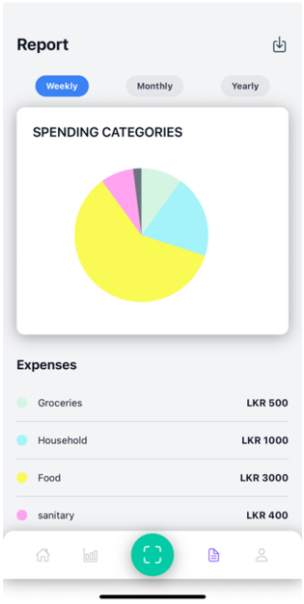


Fig 3 Image of Report Page

Profile Page: Editable user information with change password and logout options.

User Feedback Summary:

95.2% found the bill scanning feature helpful.

90.5% found the report page easy to understand.

71.4% would recommend the app to others.

58.7% rate the navigation as easy.

Quantitative Results and Benchmarking

OCR accuracy testing demonstrated that 91% of printed receipt items were correctly recognized, while handwritten receipts achieved 58% accuracy. During usability tests, the average time to record an expense was 1-2 minutes and 40-60 seconds with OCR scanning. User satisfaction rating averaged 4.4 out of 5 for overall ease of use.

Comparative Analysis

Table I Comparison of CacheBook with Existing Financial Management Applications

Feature	Cache Book	Mint	Goodbudget	Expensify
OCR Receipt Scanning	Yes	No	No	Yes
Custom Categories	Yes	Limited	Limited	Limited
Budget Goal Notifications	Yes	Yes	Yes	Limited
Visual Reports & PDF export	Yes	Yes	No	Yes
Cross platform Support	Android/ios	Android/ios/Web	Android/ios/Web	Android/ios/Web
Free Plan	Yes	Yes	Yes	Limited

Compared to existing tools, “CacheBook” offers a unique combination of OCR scanning and full customization in a free, mobile focused platform.

Challenges Observed:

- OCR performance was less accurate for blurry or handwritten receipts.
- Google Sign In could not be completed due to Firebase authorization issues.
- Current version only supports English.

Despite these limitations, “CacheBook” received highly positive feedback, especially for its automation, simplicity, and financial visualization capabilities. The application met all major functional goals and proved to be a valuable tool for encouraging better spending habits and financial awareness.

V CONCLUSION

The “CacheBook” mobile application was developed to address common challenges in personal financial management, such as manual expense logging, lack of real time tracking, and limited visualization tools. Through the integration of OCR technology, Firebase backend services, and a user friendly React Native interface, “CacheBook” enables users to record expenses efficiently, set financial goals, and visualize spending behavior through dynamic reports.

All key objectives of the system were achieved. The Mindee OCR API allowed accurate receipt scanning, budget alerts helped users manage their spending limits, and users could generate clear, categorized reports in both visual and PDF formats. Usability testing and user feedback confirmed that the application was intuitive, functional, and effective for day-to-day financial tracking.

Although limitations such as incomplete Google Sign In integration and OCR accuracy for handwritten receipts were identified, these do not significantly hinder core functionality. Future improvements will focus on multilingual support, enhanced OCR performance, and expanded features like savings tracking and financial insights.

Overall, “CacheBook” demonstrates the value of combining automation with simple design to promote better financial habits and support personal money management.

REFERENCES

[1] Holzhauer, B. (2021, August 5). Goodbudget budgeting app review. Forbes Advisor. https://www.forbes.com/advisor/banking/goodbudget-budgeting-app-review/

[2] Lusardi, A., & Mitchell, O. S. (2014). The economic importance of financial literacy: Theory and evidence. Journal of Economic Literature, 52(1), 5–44. https://doi.org/10.1257/jel.52.1.5

[3] The Bricks. (2024). 10 best personal expense tracker apps for 2024. The Bricks. https://www.thebricks.com/resources/expense-tracker

[4] ProQuest. (n.d.). Automated scanned receipt processing with optical character recognition and machine learning. ProQuest. https://www.proquest.com/openview/441877f2b7dc740f4e67e51739ecacf5/1?pq-origsite=gscholar&cbl=2026366&diss=y.

[5] Pandey, A., Vardhan, H., Verma, I., & Kushwah, S. (2023). Xpense tracker: An expenses tracking application using OCR and machine learning. International Research Journal of Modernization in Engineering Technology and Science. https://doi.org/10.56726/irjmets39430

[6] Doing, A. (2023). Mastering accuracy for receipts. Expensify Blog. https://use.expensify.com/blog/mastering-accuracy-for-receipts

[7] Garcia, M., & Claour, J. (2022). Mobile bookkeeper: Personal financial management application with receipt scanner using optical character recognition. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4083764

[8] Intuit. (2024). QuickBooks online: Features and benefits. Intuit. https://quickbooks.intuit.com/online/how-it-works/

[9] CFA Institute. (2015). Addressing financial reporting complexity: Investor perspectives separate private company accounting and beyond. CFA Institute. https://rpc.cfainstitute.org/-/media/documents/article/position-paper/addressing-financial-reporting-complexity.pdf

[10] YNAB. (2024). You Need A Budget: Budgeting software overview. YNAB. https://www.youneedabudget.com

[11] Intuit. (2024). Mint: Free budget tracker and planner. Intuit. https://mint.intuit.com

[12] PocketGuard. (2024). PocketGuard: Personal finance and budgeting app overview. PocketGuard. https://pocketguard.com

[13] Goodbudget. (2024). Goodbudget budgeting app overview. Goodbudget. https://goodbudget.com

[14] Forbes Advisor. (2023). Goodbudget budgeting app review. Forbes Advisor. https://www.forbes.com/advisor/banking/goodbudget-budgeting-app-review/

[15] The Bricks. (2023). Expense tracker app comparisons. The Bricks. https://www.thebricks.com/resources/expense-tracker

Bliss2Glamour: An Artificial Intelligence Integrated Educational Platform for Skincare and Beautician Training

Nisansa Ranthatige
School of Computing
SLIIT City Uni
0009-0002-0681-0429

Ovini Seneviratne
Department of Computer Engineering
Faculty of Engineering
University of Sri Jayewardenepura
0009-0000-9916-0650

Abstract - This research paper represents Bliss2Glamour, an artificial intelligence (AI) based educational platform developed to assist all the NVQ Level 4 blooming beauticians, qualified lecturers, and beauty enthusiasts. Bliss2Glamour has an integrating Learning Management System (LMS), well trained AI chatbot, Self-affirmations to keep the users motivated, calming music for the salon purposes, provide 24/7 skincare consultation from a highly qualified cosmetologist via WhatsApp website for beautician training, standard online quizzes for the trainee beauticians to get prepared for the exam aligned with the TVEC syllabus. This research paper highlights the motivation, methodology (Agile), implementation (FastAPI, React, React JS Query, fine-tuned QWEN 2.5 0.5B-Instruct AI model), and evaluation (via Weights & Biases). All objectives were met. The AI chatbot achieved 80% accuracy rate based on evaluation using Weights. These results confirm that Bliss2Glamour successfully combines educational content, AI technology, and holistic care into one user-friendly system.

***Keywords* — artificial intelligence (AI) chatbot, learning management system, skincare advice, calming music, self-affirmation**

I INTRODUCTION

Practical and vocational knowledge in the beauty industry lacks accessibility, personalisation, and combining with modern emerging technologies. Bliss2Glamour considers these gaps through a web-based platform targeting NVQ Level 4 novice beauticians, qualified lecturers, and beauty enthusiasts. This platform offers TVEC syllabus-aligned modules with quizzes, a personalised skincare AI chatbot, a calming music playlist for salon ambiance, and motivational support. The essential feature is the 24/7 skincare consultation from a qualified cosmetologist through WhatsApp.

The chatbot in *Bliss2Glamour* is fine-tuned to provide personalized skincare routines based on user-input skin types. It offers trusted advice without relying on camera-based analysis, ensuring accuracy and privacy. It also complements the LMS by answering beauty-related queries around the clock.

In the context of *Bliss2Glamour*, the LMS provides syllabus-aligned theory and practical lessons for NVQ Level 4 [2] beautician trainees. It includes interactive quizzes, downloadable materials, and structured modules based on TVEC standards.

Developed under Agile methodology, the system closes the gap between the current commercial systems that mainly focuses on product marketing rather than knowledge development

II LITERATURE REVIEW

The objective of the researcher is to study and learn about the existing systems which cover different areas, such as existing Learning Management Systems (LMS) [3] in Vocational Training, Use of Gamification in Education such as Quizzes, AI chatbots in Skincare and Beauty, User-centric features like calming music integration and Overview of AI in skincare, Live WhatsApp skin consultations with a qualified cosmetologist. For this developed system, the researcher grabbed sources from Academic scholarly journals/articles, TVEC guidelines and empirical studies on chatbot Innovation. The researcher gained vast knowledge of the existing systems in the beauty industry after conducting this deep case study. The observer also recognised the gaps in existing systems and approaches to address them through the suggested project. After the Literature review, the researcher gained insight into building a creative and all-in-one educational e-platform that motivates budding beauticians while offering superior services for our end users seeking personalised skincare routines for the entered skin type.

This analysis aims to compare the current systems in the beauty industry, related to the developed system. A common feature of the existing systems was they only focused on their brand promotion, unlike Bliss2Glamour system. Especially to pursue an NVQ Level 4 beauty

exam the researcher spent over one hundred thousand LKR (280 GBP). This is far more expensive for our aspiring beauticians who have financial struggles. The implemented system fills this gap by providing valuable services to freshly trained beauticians and offers ultimate free services for beauty enthusiasts who experience cash flow issues. The researcher focused on the people who do not get a chance to visit a salon. This system will fulfil the requirements of customers who seek a calming ambience by providing a playlist of relaxing music to use during their treatments or facials. Recognising the importance of motivation, the researcher includer some interesting self-affirmation as a tool to boost the users.

Most existing systems are aimed at only marketing products or just providing general courses. For example, Skinchat. Ai's chatbot gives advice on products but unspecified on beautician training. Bliss2Glamour is unique because it provides everything at once such as learning materials, exam quizzes for NVQ tests, an AI chatbot for face and skin-care advice, calming music, self-affirmations, and expert advice on WhatsApp too. It does not sell any brand and is expressly developed for beauty students even if they could afford expensive training to learn easily and comfortably.

Bliss2Glamour combines multiple technologies through a single platform, while most recent systems can only accommodate a single use. While technically combining a React front end and FastAPI back end to provide asynchronous and scalable work, its functional model is powered by a Qwen 2.5 AI model fine-tuned on skincare education and not on commonly used chatbot tools pre-built like Chatling.ai or even pre-built AI models. While different from AR-based solutions like Modiface utilizing camera-based face and skin analysis, Bliss2Glamour does not incorporate camera-based face and skin analysis and relies on user inputs for privacy and accuracy. The system itself is also further containerized with Docker and GCP-based deployment through a virtual machine utilizing Portainer.io and Docker Hub for professional-level DevOps integration and cloud scalability.

Existing Systems	AI chatbot	AI-image processing	Calming music	24/7 available live chat	Learning materials (Theory & Practical Paid)	Online quiz	Self-affirmation	Chemical ingredients which are suitable for the user's skin type	Free services	Promoting their brands
https://interview.lovi.care/quiz/	✗	✓	✗	✗	✗	✗	✗	✗	✓	✓
https://app.skinchat.ai/demo/chat/	✓	✓	✗	✗	✗	✗	✗	✗	✓	✓
https://skinanalysis.pro/	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓
https://www.salonliyo.com/	✗	✗	✗	✓	✗	✗	✗	✗	✗	✓
https://www.britishcosmetics.lk/	✗	✗	✗	✓	✗	✗	✗	✗	✗	✓
https://janet.lk/	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓
https://dreamronbeautycollege.com/	✗	✗	✗	✓	✓	✗	✗	✗	✗	✓
https://lk.spaceylon.com/	✗	✗	✓	✗	✗	✗	✗	✓	✗	✓
https://www.ihbeducation.com/	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓
https://www.robofy.ai/app/	✓	✗	✗	✓	✗	✗	✗	✗	✓	✗
https://www.loreal.com/	✗	✓	✗	✓	✗	✗	✗	✗	✗	✓
https://www.provenskincare.com/	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓

Fig 2.4 Research Gap

Fig 2.1 shows the comparison of the existing systems identified by the researcher.

III METHODOLOGY

A. Agile Methodology and Development

The methodology selected by the researcher was the Agile methodology [4] to implement the Bliss2Glamour system due to its flexibility, and the requirements were not fixed. Agile methodology allowed the researcher to break the entire system into manageable parts called sprints. It was easy for the researcher throughout this entire project because after each sprint the researcher received feedback from the supervisor, then started to integrate with other sprints if they work together without any problem.

B. Project Plan

In this phase the researcher created the WBS, Gantt chart and a Trello Board to keep track of the tasks. Especially, to track To Do, In Progress, Done. Although no formal budget was required, the researcher invested in Google Pro+ subscription to train the AI model with a powerful GPU and considered the costs for hosting and domain registration.

C. Requirement Gathering

The researcher did a systematic Literature Review, to get secondary data related to this developed system. It was very useful to get an idea about the existing systems and the technologies they have used to implement those systems.

Moreover, the researcher conducted Market Research to get the primary data from 195 users by mentioning the features of the developed system.

Fig 3.4 Feedback from the users

Fig 3.4 illustrates the summary of the detailed market research conducted and distributed to 195 participants, including professional beauticians, trainee beauticians, and beauty-conscious people. The analysis of this market research proved that most people lacked access to quality beauty study materials, trusted skincare advice, and affordable exam preparations

After conducting the market research, the researcher gained valuable insights into which features should be developed for the system. For this reason, as part of the sampling process adopted here, the snowball sampling process was adopted by the researcher because getting in touch with the target market for beauty industry is most effective. The questionnaire was spread by the researcher among acquaintances known personally, thereafter such known acquaintances had spread among friends. This method was beneficial as a method of getting responses quickly because it does not include any limitations such as sampling bias, reduced randomness and possible overrepresentation among target user populations.

D. . Project Design

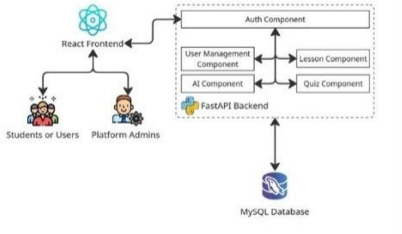


Fig 3.5 Software Architecture Diagram

Fig 3.5 illustrates the Software Architecture Diagram; it depicts how the Bliss2Glamour system's frontend and backend work.

The system has 2 main types of users, students/users who will use the platform to get the service and Platform admins who will manage the lessons, quizzes. These users interact with the system through a web interface built using React. Then the front end gets connected to the backend, which is built using FastAPI (Python).

In addition, the researcher designed wireframes, flowcharts, pseudo codes and ER diagrams to get an idea about this developed project and about the database.

E. Technologies and Tools used

The development of Bliss2Glamour involved a modern and efficient tech stack. The front end was built using React JS with Tailwind CSS for styling. Form handling and validation were done using React Hook and Zod. Swagger for API documentation. The backend was developed in FastAPI (Python 3.10), with SQLAlchemy ORM to interact with the database and Pydantic for schema validation. JWT tokens were used for authentication and role management. The system was containerized using Docker and Google Colab Pro+ were used to train the AI model with a powerful GPU. Excalidraw was used to design low fidelity wireframes. Weights & Biases (WandB) was used to monitor the AI chatbot's accuracy.

The final system was deployed using Docker Hub [5] for image management and Portainer.io as the visual UI , used manage the containers. Hosting was conducted on a Google Cloud Platform (GCP) Virtual Machine with Portainer.io installed, for deploying both the frontend and backend services of Bliss2Glamour.

IV RESULTS AND DISCUSSION

A. Results

The developed system successfully integrated all proposed modules. The AI chatbot gave an 80% accuracy rate based on its interaction with the users, evaluated through Weights & Biases visual dashboards.

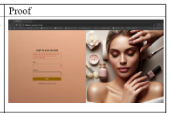
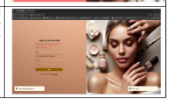
TestCase_ID	Pre-Conditions	TestCase_Title	Test_Description	Input_Value	Expected_Result	Proof
TC01	Admin must be registered and know their username and password.	Admin Login	1.1 Navigate to Login Page	Paste the URL and click enter.	Login page should be visible. PASS	
TC01	Admin must be registered and know their username and password.		1.2 Enter valid admin credentials.	Valid email & password.	Admin Dashboard should be visible PASS	
TC01	Admin must be registered and know their username and password.		1.3 Enter invalid admin credentials.	Enter Invalid password.	Error: "Incorrect username or password. Failed to login." PASS	

Fig 4.2 Manual test cases

Fig 4.2: illustrates the manual testcases done by the researcher to identify the functionalities of the system and to make sure they are working without any errors.

B. Discussion

Bliss2Glamour system achieved a reliability rate of around 85% and AI model accuracy of 80% , based on the evaluation done by using the Weights & Biases and from all the testing such as User Acceptance testing and from manual testcases. When considering all the main modules including AI chatbot, lesson delivery, quiz assessment, and WhatsApp integration indicates the system is stable, functional, and user-friendly.

The AI chatbot [6] [7], powered by the QWEN 2.5-0.5B (which has 500 million parameters) model, demonstrated 80% success rate in providing trustworthy and relevant responses during the test case evaluations. While this is a commendable outcome for a newbie in the technology sector, the accuracy did not reach 100% due to a combination of technical constraints and limited resources like lack of experience of using high level software tools and the technology stack. Especially, the main reason was that the researcher was a novice in the field of deep learning AI model fine tuning, several challenges had occurred.

C. User Feedback Gathering

The researcher selected a qualified beautician, and she is also a lecturer in the beauty industry who evaluated the system and discussed the areas that the researcher should improve to make it a better system. After the discussion the researcher got positive feedback and here is the link of the discussion.

The researcher was able to get responses from 50 respondents. Majority of the respondents were from the Beauty industry and the rest of them were beauty enthusiasts and trainee beauticians.

All the respondents mentioned that this system was highly beneficial for the trainee beauticians who face the NVQ Level 4 beauty exam, the government has declared to do this exam an online onsite examination for the first time. So, for them this system would exceptionally be useful to practice before facing for the exam. After that, the researcher implemented the changes based on user feedback collected during the evaluation stage.

V CONCLUSION AND FUTURE WORK

The developed system is a website with AI-powered educational LMS developed to encourage the NVQ Level 4 trainee beauticians, lecturers who are conducting lectures at any beauty academy for the NVQ students, and the beauty enthusiasts. Bliss2Glamour LMS consists of a syllabus aligned with the TVEC standards including interactive quizzes, the AI chatbot trained using Distillation method for skincare advice [15], self-affirmations to uplift the users, calming music for the salon purpose to play during any treatments like facials, massage treatments or any other treatments, which a user needs to stay with a peaceful mindset and real-time free skin consultation via WhatsApp with a Licensed cosmetologist.

This was an immediate achievement of all set objectives within a narrow time frame and was a great source of inspiration motivating further refinement and future improvement of this evolved system.

This application not only meets all the promised objectives but also offers educational, emotional and technological support to its users.

The following refinements are recommended as future work:

- Develop a cross-platform mobile application using React Native or Flutter for broader reach.
- Localize the UI and content which supports Multi-languages like Sinhala and Tamil as a support for the target audience.
- Upgrade the self-affirmation section by adding the text-to-speech feature, so when the beauticians are tired or unable to read, they can hear it.

ACKNOWLEDGMENT

I extend my sincere gratitude to my supervisor, for her constant support, feedback, and inspiration throughout this journey. Thanks, my beauty lecturers, and all the testers for their valuable guidance. Appreciation goes to my tech-expert cousins, and friends for helping me with new technologies, to deal with technical issues, to design professional UI/UX designs and interfaces. Finally, heartfelt thanks to my parents for the valuable support.

REFERENCES

[1] Tertiary and Vocational Education Commission (TVEC), "TVEC Sri Lanka – Official Website," 2024. [Online]. Available: https://www.tvec.gov.lk/.

[2] Wikipedia, "National Vocational Qualification," 2024. [Online]. Available: https://en.wikipedia.org/wiki/National_Vocational_Qualification . [Accessed: Dec. 11, 2024].

[3] Wikipedia, "Skin care," 2024. [Online]. Available: https://en.wikipedia.org/wiki/Skin_care. [Accessed: Dec. 11, 2024].

[4] Atlassian, "What is the Agile methodology?," 2024. [Online]. Available: https://www.atlassian.com/agile. [Accessed: Dec. 11, 2024].

[5] Docker, Inc., “Docker – Empowering App Development for Developers,” *Docker*, [Online]. Available: https://www.docker.com/. [Accessed: 29-Jun-2025].

[6] Proven Skincare, "Proven skincare," 2024. [Online]. Available: https://www.provenskincare.com/. [Accessed: Dec. 11, 2024].

[7] British Cosmetics, "Beauty begins the Minute you," 2023. [Online]. Available: https://www.britishcosmetics.lk/. [Accessed: Dec. 10, 2024].

[8] eSkilled LMS, "AI Course Creator: Interactive, custom courses in minutes!," 2024. [Online]. Available:

https://lms.eskilled.com.au/ai-course-creator/. [Accessed: Dec. 11, 2024].

[9] MoogleLabs, "AI ML, Blockchain, Data Science, DevOps development company," 2020. [Online]. Available: http://www.mooglelabs.com/. [Accessed: Dec. 11, 2024].

[10] Anthology, "Blackboard is now Anthology," 2020. [Online]. Available: https://www.anthology.com/. [Accessed: Dec. 11, 2024].

[11] Dreamron Beauty College – Sri Lanka," 2024. [Online]. Available: https://dreamronbeautycollege.com/. [Accessed: Dec. 11, 2024].

[12] Spa Ceylon, "Luxury Wellness – Spa Ceylon Sri Lanka," 2024. [Online]. Available: https://lk.spaceylon.com/. [Accessed: Dec. 11, 2024].

[13] Robofy, "Robofy – AI Chatbot for Business," 2024. [Online]. Available: https://www.robofy.ai/app/. [Accessed: Dec. 11, 2024].

[14] L'Oréal, "L'Oréal – Beauty Science and Technology," 2024. [Online]. Available: https://www.loreal.com/. [Accessed: Dec. 11, 2024].

[15] Proven Skincare, "Proven AI Skincare," 2024. [Online]. Available: https://www.provenskincare.com/. [Accessed: Dec. 11, 2024].

[16] Wikipedia, "Skin care," 2024. [Online]. Available: https://en.wikipedia.org/wiki/Skin_care. [Accessed: Dec. 11, 2024]

Decentralized Database Management: A Comprehensive Review of Blockchain- Based Data Systems

Theevika Sukirthan
Faculty of Computing SLIIT NORTHERNUNI
Jaffna, SriLanka
theevika.a@sliit.lk

Sangeetha Arunpirakash
Faculty of Computing SLIIT NORTHERNUNI
Jaffna, SriLanka
Sangeetha.b@sliit.lk

Gowrisha Karuneswaran
Faculty of Computing SLIIT NORTHERNUNI
Jaffna, SriLanka
karuneswaran.g@sliit.lk

Tharmakulasingam Tharmmendra
Faculty of Computing SLIIT NORTHERNUNI
Jaffna, SriLanka
tharmmendra.t@sliit.lk

Vigneswaran Vithiyasahar
Faculty of Computing SLIIT NORTHERNUNI
Jaffna, SriLanka
vithiyasahar.v@sliit.lk

***Abstract*— The emergence of blockchain technology has revolutionized decentralized data management by offering robust alternatives to traditional centralized database systems. This paper provides a systematic and comprehensive review of blockchain-based distributed databases, highlighting key architectural transformations, core enabling technologies such as Merkle Trees, PBFT, and Zero-Knowledge Proofs, and comparing them with conventional distributed databases. Real-world implementations including Hyperledger Fabric, BigchainDB, and OrbitDB are analyzed to assess their scalability, interoperability, and security capabilities. The paper also explores intrinsic security mechanisms, performance bottlenecks, and regulatory challenges that affect adoption. Finally, it identifies open research questions and future directions necessary for building scalable, privacy-aware, and interoperable decentralized database ecosystems suitable for enterprise and multi-stakeholder environments.**

Keywords— Blockchain databases, consensus mechanisms, data integrity, decentralized systems, distributed ledger, Merkle trees, Zero-Knowledge Proofs

I INTRODUCTION

The conventional database management frameworks are used to organize their systems through centralized architectures to enable operational pace and organizational control and system expansion. These systems generate two interconnected weaknesses because they produce both single point failure risks and vulnerabilities to data breaches along with a dependence on trusted third-party data validation services and storage solutions [2], [4]. Several measures through data replication and distribution treatment distributed databases sought to resolve such issues [6] but they generally maintain central authority control.

Blockchain technology introduces transformative changes by enabling data decentralization, cryptographic verification, and peer-to-peer operations, thereby removing reliance on centralized authorities [2]. Although it offers tamper-resistance, operational transparency, and a trustless environment, blockchain was originally developed for securing financial transactions rather than supporting dynamic database functionalities [4]. Traditional blockchain systems face limitations such as low transaction throughput, inefficient storage mechanisms, and rigid data structures [7].

This review investigates how blockchain technology evolves from its original financial applications to form the backbone of modern decentralized database systems. Section II outlines the systematic review methodology used to identify and assess relevant literature. Section III introduces core blockchain concepts and contrasts them with traditional distributed databases. Section IV discusses architectural adaptations that make blockchain suitable for data storage, followed by Section V, which examines enabling technologies such as Merkle Trees, consensus protocols, and Zero-Knowledge Proofs. Section VI explores different system models and real-world platforms. Section VII analyzes security mechanisms and threat mitigation strategies, while Section VIII addresses prevailing challenges and future research directions. Together, these insights provide a comprehensive foundation for understanding the state and potential of blockchain-based distributed databases.

II SYSTEMATIC REVIEW PROTOCOL

This review adopted a structured, systematic approach to assess and synthesize literature on blockchain-based database systems. The goal was to ensure inclusion of credible, peer-reviewed, and relevant sources.

- Databases Searched: IEEE Xplore, ACM Digital Library, SpringerLink, Elsevier ScienceDirect, and Google Scholar.

- Search Keywords: "blockchain database systems", "decentralized data storage", "Merkle tree in DBMS", "PBFT consensus in blockchain", "ZKP privacy blockchain", "blockchain vs distributed DB".
- Time Frame: Publications from 2017 to 2024 were considered.
- Inclusion Criteria: Peer-reviewed academic articles, technical conference proceedings, and foundational whitepapers discussing blockchain applications in data management.
- Exclusion Criteria: Non-English sources, informal blogs, and materials lacking relevance to blockchain-based databases.

The screening followed a staged process: title and abstract evaluation, followed by full-text review of shortlisted works. As visualized in the PRISMA flow diagram (Figure 1), 320 initial records were identified, 250 duplicates removed, and 70 remaining titles screened. After relevance checks, 30 full texts were assessed, resulting in 13 high-quality articles included for review.

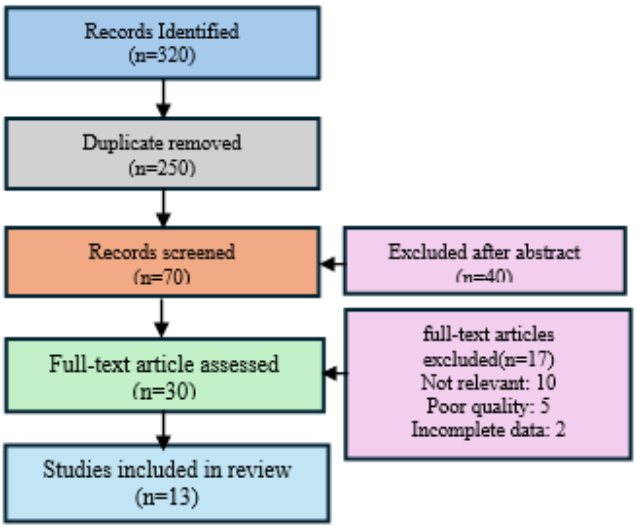


Fig. 1. PRISMA flow diagram

III FUNDAMENTALS OF BLOCKCHAIN TECHNOLOGY

A. Definition of Blockchain

Blockchain is a distributed, decentralized ledger technology that records transactions over a network of nodes in a tamper-proof, transparent, and secure manner. Transactions are bundled into blocks, cryptographically linked to preceding blocks, and validated through consensus mechanisms. Blockchain design eliminates the need for central authorities, enabling direct, trustless interaction between parties.

The major technologies supporting blockchain include peer-to-peer networking, cryptographic hashing, public-key cryptography, digital signatures, and consensus mechanisms such as Proof of Work (PoW) and Proof of Stake (PoS) [12]. These mechanisms collectively render it computationally infeasible to alter data retroactively after being recorded without the consensus of most of the network.

B. Features of Blockchain

Blockchain systems possess several fundamental features that distinguish them from traditional approaches to data management. Decentralization initially ensures that the verification and storage of data are collectively handled by a network of nodes rather than a single central authority, reducing manipulation risks and single points of failure [2], [4]. Immutability is another feature where, once a transaction is recorded and verified through a consensus protocol, it is effectively unalterable without the agreement of the majority of network members [2]. Transparency is also enabled, particularly in public blockchains, where all transactions are openly visible and can be verified by anyone, which fosters trust and accountability [4]. Finally, blockchain security is reinforced by advanced cryptographic techniques such as hashing, digital signatures, and public-key infrastructures to ensure authenticity, integrity, and confidentiality of data recorded [2]. Together, these inherent properties enable blockchain to ensure a secure, tamper-evident, and trustless state for decentralized data management.

C. Differences Between Blockchain and Traditional Distributed Databases

The Table I highlights the key differences between blockchain and traditional distributed databases:

TABLE I. KEY DIFFERENCES BETWEEN BLOCKCHAIN-BASED AND TRADITIONAL DISTRIBUTED DATABASES

Feature	Blockchain-Based Database	Traditional Distributed Database
Data Structure	Linked blocks secured by cryptographic hashes	Relational or non-relational tables with editable records
Control	Decentralized, governed by consensus protocols (e.g., PoS, PBFT)	Centralized or semi-centralized administration
Immutability	Immutable—data cannot be altered once recorded	Mutable—authorized users can update or delete data
Transparency	Transparent—especially in public blockchains	Access restricted by roles and permissions
Performance	Generally slower due to consensus and replication	Optimized for high-throughput operations

Fault Tolerance	High—due to distributed architecture and redundancy	Moderate—depends on backup and failover mechanisms
Trust Model	Trustless—trust enforced by protocol	Trust in system administrators and internal controls
Examples	Ethereum, Hyperledger Fabric, BigchainDB	MySQL, MongoDB, Oracle, PostgreSQL

IV From Blockchain to Blockchain-Based Databases

A. Why Adapt Blockchain for Databases?

The main design of blockchain technology does not match the operational needs of distributed databases that require sophisticated high-speed data processing. The Proof of Work (PoW) method in native blockchains imposes restricted transaction speed and long processing times along with poor storage performance because their data structures are distributed across multiple systems [2], [4]. The benefits of public transparency exist for specific use cases, yet such properties create problems when enterprise-level security and access control requirements must be met [6]. Blockchain system characteristics create barriers which prevent their direct application to function as a general-purpose database system.

Significant modifications need to occur to achieve suitable architectural adaptations when building distributed databases based on blockchain technology. The system needs to receive essential improvements which will boost its scalability capacity and enhance data storage performance together with adaptable governance solutions and strong decentralized features and system performance maintenance. The following section details all substantial architectural changes made to blockchain technology for distributed database system operations.

B. Key Architectural Adaptations

To address the inherent scalability, privacy, and flexibility limitations of traditional blockchain systems, modern blockchain-based databases have adopted several key architectural modifications. One such enhancement is the use of hybrid storage models, where lightweight transaction proofs are stored on-chain, while bulk data is offloaded to decentralized file systems like the InterPlanetary File System (IPFS). This separation reduces network congestion and significantly improves query performance by optimizing on-chain data usage [5].

Database applications drive the optimization work carried out on consensus protocols. PoW guarantees security effectiveness, yet its computational intensity and lack of high transaction speed makes it inappropriate for fast deal processing. To achieve both quick confirmation and enhanced energy efficiency Practical Byzantine Fault Tolerance (PBFT) alongside Proof of Authority (PoA) and Delegated Proof of Stake (DPoS) are selected as alternative faster consensus mechanisms [8], [10].

A vital adaptation exists in Network access control. Many blockchain databases establish permissioned system protocols because they do not support the unrestricted public participation that characterizes typical public blockchains. The chosen design improves scalability capabilities and delivers better security standards in addition to meeting data protection regulations [7].

Data governance achieves its reinforcement by implementing smart contracts within the system. Possible business scripts programmed by developers undertake various tasks inside database infrastructure for automating business functions and validating data while managing system access permissions. The governance system using smart contracts decreases central administrator requirements without compromising the databases operational transparency or their auditing capabilities or consistency [4], [7].

These adaptations allow blockchain-oriented distributed databases to produce the essential blockchain benefits of distributed systems and tamper-resistant security features alongside enhanced data management capabilities.

V Core Enabling Technologies

A. Merkle Trees

Merkle trees are cryptographic data structures that allow efficient and secure verification of data integrity in large, distributed environments. In blockchain databases, they ensure that any alteration to a single transaction can be detected without scanning the entire dataset. This drastically improves verification speed in audit processes and is particularly effective in managing append-only data. However, their use introduces trade-offs in update operations, as altering even one record can require recalculating multiple hashes in the tree hierarchy. Real-world benchmarks in systems like Bitcoin show that Merkle trees support efficient SPV (Simplified Payment Verification) but are less suited for use cases requiring high-frequency write operations.

B. PBFT (Practical Byzantine Fault Tolerance

PBFT offers high fault tolerance and energy efficiency compared to Proof of Work (PoW). It is well-suited to permissioned blockchain databases where the number of validator nodes is controlled. By design, PBFT can tolerate up to one-third of faulty or malicious nodes. Its deterministic consensus provides fast finality, making it suitable for financial systems and healthcare databases. However, its performance deteriorates as the number of

nodes increases, which limits its scalability. Platforms such as Hyperledger Fabric employ PBFT variants for consensus in controlled enterprise environments [7], [10].

C. Zero-Knowledge Proofs (ZKP)

ZKPs allow one party to prove possession of information without revealing the content itself, addressing the tension between transparency and privacy in blockchain systems. They are essential in protecting sensitive data in sectors such as identity management and finance[1]. However, ZKPs require significant computational resources and time, which can limit their use in real-time or large-scale environments. Projects like Zcash implement zk-SNARKs, a type of ZKP, to offer anonymous transactions. Despite their promise, further optimization and standardization are needed before ZKPs can be broadly applied in blockchain-based databases [11].

VI System Types and Real-World Platforms

A. Blockchain Database System Types

Distributed databases based on blockchain can be categorized as public, permissioned, and hybrid systems based on access control, governance structures, and intended application contexts [6], [7].

All parties can engage with Public Blockchain Databases since these networks welcome participants without needing any form of pre-approval. All network participants hold the ability to verify the transparency of the data stored on the network. The high decentralization and transparent nature of Bitcoin and Ethereum networks faces scalability issues and requires better performance that makes them impractical for enterprise database applications [6].

The participation in Permissioned Blockchain databases is limited to trusted nodes who require approval to join the network. Enhanced scalability along with privacy features and regulatory compliance become possible because access controls govern the participation in consensus procedures and data visibility rights. Hyperledger Fabric stands out as one platform among several others that offers enterprise users modular consensus features and channel-based privacy functionality according to [7].

The hybrid system of Blockchain databases integrates features from both public and permissioned blockchain architecture models. The systems support individual adjustments of transparency levels alongside specific access management systems which tailor openness against confidentiality protection based on application needs [6], [7].

Multiple blockchain database systems need evaluation based on trust assumptions and scalability needs and confidentiality criteria and regulatory needs.

B. Real-World Platforms

Several real-world platforms demonstrate how blockchain ideas are adapted to create practical distributed databases. real platforms in the world show how distributed databases develop from fundamental blockchain principles.

Three standard database characteristics work alongside blockchain abilities in BigchainDB: rich querying together with low-latency transaction processing and blockchain guarantee of immutability and decentralization. Through Tendermint's Byzantine Fault Tolerant consensus algorithm BigchainDB reaches scalability while achieving fast finality thus becoming suitable for asset registries together with supply chain management and intellectual property rights management [5].

Organizations use open-source permissioned blockchain platform Hyperledger Fabric because its foundation exists within the Linux Foundation. This platform maintains three key components which include parallel agreement consensus along with chain code (smart contracts) and private digital channels for data confidentiality. The business network deployment capabilities of Fabric extend across financial institutions and healthcare organizations and logistics providers [7].

The decentralized database OrbitDB operates as a peer-to-peer system which constructs its framework from the peer-to-peer network IPFS. The distributed data synchronization functionality in OrbitDB runs through conflict-free replicated data types (CRDTs) to operate without depending on central servers. The decentralized database solution OrbitDB fits applications that need real-time data sharing capabilities alongside offline data operation functionality [6].

Blockchain-based distributed databases have started to be widely investigated for use in sector-specific implementations along with their standard usage scenarios. Blockchain technology enables system functions in the energy industry to support decentralized peer-to-peer trading operations that require advanced transparency while boosting network operational performance [3].

These platforms and application areas illustrate the design choices in blockchain-oriented database development, addressing key challenges like scalability, performance, confidentiality, and interoperability.

VII Intrinsic Security Mechanisms and Risk Prevention

The security features of blockchain-based distributed databases ensure trustless operation, data integrity and network fault tolerance through their built-in security measures. The authentication of blockchain transactions depends on consensus mechanisms PBFT with Practical Byzantine Fault Tolerance along with PoS and DPoS which achieve higher performance levels and reduced power

consumption than the Proof of Work method [8], [10]. The application of cryptographic elements enables data protection through tamper-proof digital signatures that are generated by hashing protocols while public-key cryptography verifies transaction authentication [2]. Merkle trees allow large, distributed networks to handle data effectively while providing verifiable data management capabilities that lead to transparent network-wide data confirmation [8].

Database security through blockchain heavily depends on effective risk avoidance systems. Digital asset fraud prevention occurs because transactions need to pass through a consensus authentication process to stop double spending attempts [2], [8]. The Sybil attacks prevention through economic costs in PoS-type protocols makes it difficult for users to establish multiple false identities [8]. The defense against replay attacks depends on including nonces with timestamps in transaction records [8]. Data replication among multiple nodes serves as an implementation method for system fault tolerance which delivers ongoing functionality when partial network failures occur [5]. Safety on Hyperledger Fabric structures depends on enforced entry protocols and scheduled node inspection that screens network conduct and regulatory compliance [7].

These integrated elements work collectively to boost the scalability and reliability and fault tolerance of distributed databases that use blockchain and represent appropriate solutions for vital government and financial together with industrial applications.

VIII Challenges, Open Issues, and Future Directions

Blockchain-based distributed databases present numerous advantages but also face ongoing challenges that hinder their widespread adoption. Below is a critical analysis of the most pressing issues, along with potential future research avenues.

A. Performance and Scalability

Current blockchain systems struggle with throughput limitations and high transaction latency. Public blockchains like Bitcoin and Ethereum process fewer transactions per second compared to traditional databases. Solutions such as sharding, Layer-2 protocols (e.g., state channels, rollups), and alternative consensus algorithms like Proof of Stake (PoS) and Practical Byzantine Fault Tolerance (PBFT) aim to mitigate these constraints [8], [10]. However, trade-offs remain in complexity, security, and decentralization. Future work should focus on adaptive consensus models and dynamic resource scaling for real-time performance optimization.

B. Interoperability and Integration

Isolated blockchain networks hinder data exchange across systems. Bridging solutions, such as cross-chain protocols and interoperable smart contracts, are emerging but remain fragmented and complex. Research gaps exist in standardizing cross-ledger communication and ensuring secure, scalable data movement between heterogeneous platforms. Future directions include the development of lightweight blockchain middleware and universal APIs for multi-chain environments [9].

C. Privacy and Data Protection

Blockchain's transparency poses a paradox when dealing with sensitive data. Privacy-preserving technologies like Zero-Knowledge Proofs (ZKP), Homomorphic Encryption, and Multi-Party Computation (MPC) are promising, but they introduce computational overhead and are not yet widely supported across platforms [11]. The need to balance data confidentiality with verifiability presents an open challenge. Further research should aim at efficient, modular privacy layers that integrate with smart contracts and permissioned networks.

D. Governance and Regulatory Compliance

Most current systems lack robust, decentralized governance mechanisms. The absence of clearly defined protocols for upgrades, dispute resolution, and policy enforcement leads to fragmentation and trust concerns. There is also a lack of compliance with regulations like GDPR and HIPAA in many blockchain deployments [6]. Future studies should explore decentralized autonomous organization (DAO)-driven governance and legal-smart contract co-design frameworks [13].

E. Research Gaps and Open Questions

1) What are the long-term trade-offs between decentralization, speed, and security in hybrid blockchain-DB models?
2) How can empirical benchmarking be standardized across blockchain-based databases?
3) What mechanisms will ensure interoperability without compromising security?
4) How can privacy be enforced at scale in multi-stakeholder environments?
5) Can DAOs offer sustainable and compliant governance models for data infrastructure?

These critical questions form a roadmap for advancing the development, deployment, and trustworthiness of blockchain-based database systems.

IX Conclusion

Blockchain-based distributed databases offer a transformative model for achieving data integrity, resilience, and decentralization. This paper presented a structured and systematic review of blockchain-enabled

database systems, covering foundational technologies, architectural adaptations, real-world platforms, and security mechanisms. By integrating insights from peer-reviewed sources and addressing core components such as Merkle Trees, PBFT, and Zero-Knowledge Proofs, the review highlighted both the capabilities and the limitations of existing systems. The critical analysis of performance bottlenecks, interoperability issues, privacy trade-offs, and governance shortcomings emphasized the need for continued research and innovation.

Future advancements must prioritize scalable consensus algorithms, privacy-preserving infrastructure, cross-chain interoperability, and compliant governance frameworks. With these developments, blockchain-based databases can support increasingly complex, regulated, and multi-stakeholder digital ecosystems, making them viable alternatives to traditional distributed data systems.

REFERENCES

[1] W. L. Sim, H. N. Chua, and M. Tahir, "Blockchain for Identity Management: The Implications to Personal Data Protection," in 2019 IEEE Conference on Application, Information and Network Security (AINS), 2019.

[2] M. J. M. Chowdhury, A. Colman, M. A. Kabir, J. Han, and P. Sarda, "Blockchain versus Database: A Critical Analysis," in 2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/12th IEEE International Conference On Big Data Science And Engineering, 2018.

[3] S. V. Oprea and A. Bâra, "Distributed database on blockchain technology for new era of electricity transactions," Scientific Bulletin of Naval Academy, vol. XXII, 2019.

[4] Y. Wang, C. Li, and C.-H. Hsieh, "Research and Analysis on the Distributed Database of Blockchain and Non-Blockchain," in 2020 IEEE 5th International Conference on Cloud Computing and Big Data Analytics, 2020.

[5] E. Gaetani, L. Aniello, R. Baldoni, F. Lombardi, and A. Margheri, "Blockchain-Based Database to Ensure Data Integrity in Cloud Computing Environments," 2017 IEEE 9th International Conference on Cloud Computing Technology and Science (CloudCom), pp. 331–336, 2017.

[6] F. Casino, T. K. Dasaklis, and C. Patsakis, "A Systematic Literature Review of Blockchain-Based Applications: Current Status, Classification and Open Issues," Telematics and Informatics, vol. 36, pp. 55–81, 2019.

[7] Z. Li, G. Z. Papadopoulos, S. Nayak, A. V. Vasilakos, and G. X. Chen, "Analysis of Data Management in Blockchain-Based Systems: From Architecture to Governance," IEEE Transactions on Knowledge and Data Engineering, vol. 34, no. 11, pp. 5224–5240, 2022.

[8] X. Li, P. Jiang, T. Chen, X. Luo, and Q. Wen, "A survey on the security of blockchain systems," Future Generation Computer Systems, vol. 107, pp. 841-853, 2020.

[9] J. Zhang and P. C. P. Tung, "Blockchain and Interoperability: Review and Open Research Issues," IEEE Access, vol. 8, pp. 49644–49660, 2020.

[10] M. Castro and B. Liskov, "Practical Byzantine Fault Tolerance," in Proceedings of the Third Symposium on Operating Systems Design and Implementation (OSDI), New Orleans, USA, 1999, pp. 173–186.

[11] M. M. H. Khan and K. Salah, "IoT Security: Review, Blockchain Solutions, and Open Challenges," Future Generation Computer Systems, vol. 82, pp. 395–411, 2018.

[12] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," 2008. [Online]. Available: https://bitcoin.org/bitcoin.pdf

[13] A. Hassan, M. Rehmani, and J. Chen, "Blockchain for Internet of Things: Opportunities, Challenges, and Future Directions," Journal of Network and Computer Applications, vol. 156, 2020

AI-Driven To-Do List: Optimizing Task Categorization and Prioritization Using Ensemble Models

Pirathap Vishaliney
Faculty of Graduate Studies and ResearchSri Lanka Institute of Information Technology Malabe, Colombo
MS24905718@my.sliit.lk

Chanaka Sujith Pemasiri
Faculty of Graduate Studies and Research Sri Lanka Institute of Information Technology Malabe, Colombo
MS24901512@my.sliit.lk

Mathurantha Kanthakumar
Faculty of Graduate Studies and Research Sri Lanka Institute of Information Technology Malabe, Colombo
MS24907798@my.sliit.lk

Nimesha Yatigammana
Faculty of Graduate Studies and Research Sri Lanka Institute of Information Technology Malabe, Colombo
MS24908474@my.sliit.lk

***Abstract*—This paper introduces the AI-driven smart to-do list that can cluster and prioritize activities by using the machine-learning methods. Traditionally, to-do list services are immovable and have an element of compromising users to input the information themselves; this sort of bare tool can easily lead to unproductiveness in accomplishment of duties. To address this situation, we supplement ensemble modeling, namely Logistic Regression, XGBoost, and Multilayer Perceptron, to delegate the tasks to the desired categories and define priorities by their urgency. Measured based on standard measures, the ensemble will achieve 47.7 percent accuracy when doing classification and 72.8 percent when predicting priority, and High Priority tasks will gain in this evaluation. Using BERT-based embeddings in combination with TF-IDF-based vectorization, the system should improve its effectiveness because it understands the semantics of described tasks. Together these blocks form a superb ensemble architecture that can beat stand-alone model when it comes to classification and forecasting. More importantly, the system still leaves itself potential to adjust to user behavior and therefore it can improve task management, and it is a feasible platform in real time organization of tasks.**

***Keywords*—AI-powered To-Do List, Task Categorization, Task Prioritization, Machine Learning, Ensemble Models**

I. INTRODUCTION

Task management is one of the pillars of productivity at work and in life. Traditional tools to manage a to-do list, such as To-do List, Google Keep, and Microsoft To-Do, have already become an ordinary work tool. These systems, however, rely on humans to come up with categories and structures but this becomes a tedious process when there are changes in schedule and these changes are abrupt and uncontrollable. Since the software does not have the intelligence to adapt to such fluidity or to accommodate the priorities that are at an individual level there is an accumulation of tasks that are under-classified, which makes it a source of cognitive overload and reduced production. In summary, legacy task management tools could not take the entire weight of bracketing and prioritizations and users had to work the grunt work at the expense of productivity..

Of recent studies, there has been a sizeable amount of evidence suggesting that employing Artificial Intelligence (AI) in the task-management system became one of the vital areas of investigation. By being integrated into such systems, AI changes the static nature of to-do lists and turns them into dynamic intelligent things. The utilization of automatic categorization, prioritization and even task-urgency-based recommendations allow users to incorporate the historical patterns which are created during their interaction with the tool. More importantly, such a development path requires the ability of machine-learning (ML) models to identify and internalize patterns based on historical user behavior and, through them, learn more about the individual preferences.

To some extent, empirical studies have been in the position to offer insight into the tradeoffs between the traditional to do lists and AI powered versions. As an example, as shown in [2] approached the possibility of deep learning to improve the accuracy of task classification. According to the findings, the system showed better efficiency in the categorization of tasks in comparison to the manual techniques as soon as machine-learning techniques were used [2]. Besides, the platforms driven by AI reduce human error when allocating and prioritizing tasks in the following ways: categorizing and sorting tasks based on their level of urgency and relevance.

There are two fundamental problems that are regularly experienced in task management, and these are categorization of tasks and prioritization of tasks. Older technologies require the user to attribute items to predetermined categories, such as the following: work, personal, health, etc. However, categories are subjective in

nature, and they vary with user-to-user boundaries. Similarly, there is not always a clear set of criteria whereby something is assigned priority, i.e. high, medium, or low depending on an individual personal schedule and world demands. Put collectively, all these challenges suggest that a system that can classify, assign and rank tasks in a way that resembles the individual personal preferences and priorities of every individual and priority of each task would come in as a desirable feature.

The promising modern solution to this is the inclusion of machine learning (ML), and specifically ensemble learning. Assembling techniques combine the results of many models, thus lowering the error rate in addition to strengthening the results of classification. The empirical analysis conducted by the method in [3] proves that ensembles tend to perform far better than individual classifiers in challenging environments like task categorization which, at the same time, enhances both the accuracy and the efficiency of the classification process [3].

In the AI and cognitive-engineering field, the reader may be exposed to the idea of an AI-enabled Smart To-Do List which is supposed to do not only the classification of the tasks, but also to prioritize them with the help of ensembles of machine-learning models. The ensemble in question used here has been composed of the Logistic Regression, XGBoost, and Multilayer Perceptron (MLP) models all of which were tuned with the help of the dataset containing approximately 5,000 items of to-do lists in the real-life setting. By using the TF-IDF vectorization model and BERT embeddings, the system pulls semantic features by using only the descriptions of the tasks, allowing specific categorization or prioritization to be formed.

The main contributions of this system can be summarized as follows:

- Automatic sorting the tasks into the predetermined categories e.g. Work, Personal, Health, etc.
- Prioritization of tasks based on the three leveled urgency scale (High, Medium, Low)
- The dynamism of the system to adjust itself to the way the user will behave with time and thereby increase the task-management efficiency as a whole

Once implemented by a user this system will be capable of bringing significant results to the efficiency and accuracy of task management and at the same time helping to decrease the input of manual work in it and enhance the general productivity. In addition, the fact that the system is in real-time presupposes that it can be responsive to changes in the current task-management landscape.

II. Literature Review

When considering the current task-management applications, such as To-do List and Google Keep, one can identify a certain methodological weak point: they provide the user with merely a list of static resources that were manually curated and are poorly suited to support environmental contingencies and varying user needs that modern schedules are subject to. Practically, the process of typing, categorizing and arranging jobs could turn out to be cumbersome thus reducing efficiency as well as overall productivity. Initial attempts at bypassing these limits were based on rule-based algorithms, which, however, turned out to be rigid and failed to reflect the complexity of the day-to-day work ultimately [1].

Subsequent research has thus centered around new paradigms of machine-learning (ML), taking advantage of data-driven designs that refine information on previous user interactions. Relevant case in point is the work on the University of Twente that investigated the interaction patterns that take place when using a digital device (e.g. discrete counts of mouse clicks and key presses) and used the supervised classification methods to identify and or classify the type of the task. The findings indicated that rather modest feature sets of ML could achieve decent predictive accuracy and provoke a high level of user acceptance, which supports the future value of ML-based inference in the specified scenario [2].

In the LM literature, combinations of many different learner hypotheses via ensemble-learning methods often outperform their component single-model analogues, a tendency that proves particularly striking in more complex classification domains. The existing evidence proves that the polyhedral combinations of Logistic Regression, XGBoost, and Multi-Layer Perceptron (MLP) can provide the increased classificatory accuracy alongside the reduced overfitting inclinations. Special mention should be made of the fact that an ensemble policy combining model selection with a data-envelopment-analysis based and a MLP aggregator significantly outperforms more traditional voting policies in terms of predictive accuracy and computing speed. Ensemble architectures of such kind therefore provide special promises in tasks which require heterogeneous data inputs as well as subjective prioritization criteria [3].

Fellow workers, this time I would like to change the subject a little, namely to the points of cross-connection, but this time before the execution: AI is no less helpful when it comes to the scheduling and prioritizing of tasks as well. In industrial engineering, the authors in [4] have supported a Dijkstra-inspired graph algorithm to fully autonomous task scheduling, thus proving feasibility of generic, efficient, and real-time task assignment by means of graph-based modeling. At the same time, the authors in [5] have presented a framework of machine learning to the issue of priority assignment in real-time systems, gaining success in applying supervised learning to the central minimization problem of defining priorities of global fixed-priority preemptive scheduling on multiprocessors. Appropriately combined, these studies still reveal the flexibility of ML in both optimizing the dimension of categorization and prioritization of tasks [4][5].

To speak about AI integration more holistically, one can pay attention to the human AI task tensor that structures the work along several interrelated dimensions, i.e., task definition dimension, AI contribution dimension and decision-making sovereignty. This theoretical framework focuses on the necessity to profoundly study and improve the interaction between human users and AI systems in collaborative task contexts and establishes a theoretical basis of the future studies on generative AI in work organization [6].

The previous work is usually based on shallow machine-learning or deep-learning strategies separately, losing the advantage of combining these two paradigms to achieve the strengths of both algorithms simultaneously. In addition, semantic knowledge of task descriptions has also been restricted to superficial keyword matching or shallow vectorization approach. The current paper minimizes these pitfalls by offering an ensemble model architecture built out of Logistic Regression, XGBoost, and Multilayer Perceptron (MLP) models that are trained on the hybrid set of features on the TF-Vectorized data using TF-IDF vectorization and the contextual BERT embeddings. The combination allows the system to reproduce both statistical importance of terms and the deep semantic context.

The design considers dual-output prediction- task category and priority with the embedded architecture based on Flask-based back end and deployed on an interactive web interface with real-time processing of tasks. Generalizability and overfitting can be further increased by cross-validation and hyperparameter optimization. These two, combined altogether, constitute a much-needed gap in current literature, as it effectively illustrates the ability to integrate ensemble learning and semantic representation to create a reactive, AI-based task-management prototype that delivers quantifiable increase in performance.

III. METHODOLOGY

A. Dataset Collection and Preprocessing

To have a corpus strong enough to be used in machine-learning analyses, we assembled a database of 5,000 task descriptions with mixed origins, i.e. we have diversified it in terms of multiple categories and tasks priority rates. There were two categorical variables that were annotated:

Category: a set of labels that are preselected, such as Work, Personal, Health and so on.

Priority: a scale that is ordinal and it is either High, Medium or Low.

The preprocessing steps entailed the following procedures:

1) Text Normalization: All text was transformed to lowercase, and punctuation was removed as well as any excess whitespace deleted.

2) Tokenization: The description of every task was divided into individual tokens (words or phrases)

3) Stopword Removal: Highly frequent and somewhat uninformative words, like the, and, of, were produced to remove, to focus on interesting content.

4) Transformations: The data was transformed into a numerical representation of the text by using Term Frequency-Inverse Document Frequency (TF-IDF) which allows us to weigh terms according to their occurrence throughout the corpus but also considering general documentary abundance.

5) BERT Embeddings: To add up semantic analysis BERT embeddings comprised, an effort to recognize the contextual meaning, and consequently increase accuracy in the downstream classification assignment.

B. Model Selection and Training

In our attempt to achieve quite correct categorization and priority of tasks, we used the ensemble learning approach, which combined the output of three, complementary, procedures. The initial logistic regression is a linear classifier and carries out best in simple classification tasks. Second, XGBoost as the variation of gradient-boosting machines is largely praised due to the level of efficiency and accuracy when applied to classification problems. Lastly, a more basic neural network model, the multi-layer perceptron (MLP), can learn complex patterns provided in data. Taken together, these models use the advantages of each; the ensemble combines the results of the group of models and thus improves predictive generalization by averaging and beating overfitting. has a corpus strong enough to be used in machine-learning

C. Evaluation Metrics

During the empirical analysis, we used a few evaluation criteria to determine the achievements of the models. Accuracy gave a broad level of success with the percentage of prediction complying with ground-truth categories of tasks and priority labels being used. To better explain strengths and weaknesses of classes, we complemented this macro-view with precision, recall and F1-score measures, that each measure a unique aspect of accuracy of classification.

To avoid overfitting, we underwent an intensive k-fold cross-validation procedure, in which hyperparameters of each of the models were tuned on a split of the data and resulting parameter sets were evaluated on remaining folds.

D. Model Integration and System Development

In the Flask-based backend architecture, the trained models were integrated in such a way that descriptive text

given by the user would pass the ML pipeline, by the end of which the model would provide a category as well as a priority score. This back-end segment was subsequently connected to an easy-to-use front-end interface, which allowed interactive participation with a web-based to-do list.

As a result, the created system is dynamic and assigns Task Categories, categorizing Priorities into three levels, i.e., High, Medium, or Low, thus, preserving an AI-powered workflow in an organization in real-time.

Fig.1 below illustrates the system architecture, showing the flow of data from the user input on the frontend through the ML pipeline in the backend, and how the results are delivered back to the user.

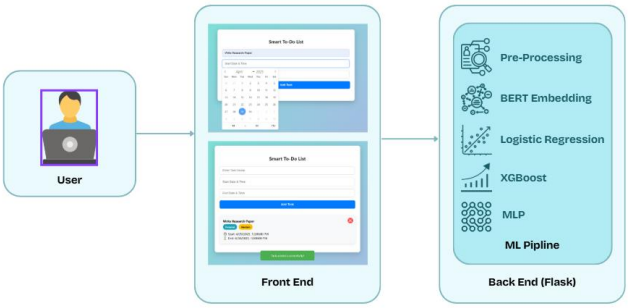


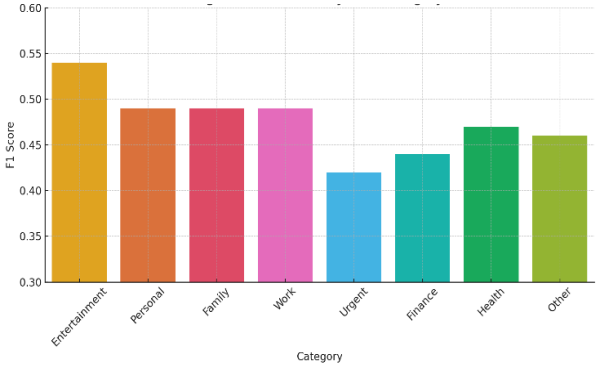
Fig. 1. System Architecture

IV. Results and Discussion

A. Category Classification Results

In our experiment, where ensemble method with Logistic Regression and XGBoost was used, the global accuracy of 47.7 % was obtained, which, despite not being very high, still underlines the possibility of refined improvement of the model tested.

These F1 scores that are provided with the model also show that there is a significant difference in the performance on the eight predefined task types: Entertainment reached the highest F1 score, 0.54, Personal, Family, and Work scored with a close range centered on 0.49, and Urgent and Work were misclassified widely, and the reason can be associated with these types sharing semantically similar ground in task descriptions



The corresponding confusion matrix shows these same trends considering that the Urgent and Work classes were frequently mixed up, with Personal and Finance-related tasks also being misclassified with an overall high degree of heterogeneity, as one might expect after considering that these classes are a lot more related to one another than the previous results tend to do.

TABLE 1 Confusion Matrix for Task Categorization Table

Predicted /Actual	**Personal**	**Work**	**Health**	**Entertainment**	**Finance**	**Urgent**	**Other**
Personal	290	40	25	50	30	20	10
Work	60	350	30	60	35	10	15
Health	45	60	210	25	15	10	5
Entertainment	25	20	10	275	20	10	10

B. Priority Classification Results

During our research, the ensemble method proved to be the most efficient tool that could be utilized in task prioritization, with the accuracy of 72.8 %. This was more than that of the other category-based classification system.

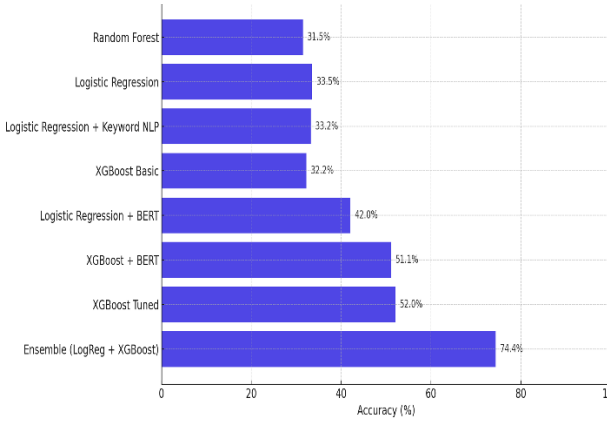


Fig.3. Model Accuracy Comparison for Priority Prediction using Various NLP and ML Techniques

When the metrics of precision and recalls are scrutinized it is observed that precision and recall of the High Priority class were highest at 0.76 and 0.79 respectively implying a result of 0.78 when considering the F1-score. The Medium Priority and the Low Priority tasks had an F1-score of 0.68 and 0.71 respectively suggesting that the overall performance was much nearer.

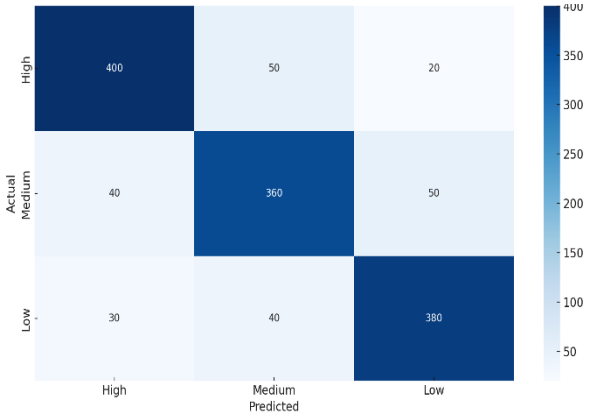


Fig.4. Confusion Matrix for Task Priority Classification

These data are supported by corresponding confusion matrix: the most correctly classified category was the High Priority where the level of misclassification was minimal. On the contrary, the task Medium and Low Priorities exhibited a high degree of confusion, with medium tempting to be interchanged with Low, and vice versa.

C. *Model Comparison*

Logistic Regression records an average of 39.5 % in the prediction of categories and 68.2 % in the prediction of priorities. Its performance is attributed as commendable, but it fails on its result in comparison to that of the ensemble approach especially on category classification.

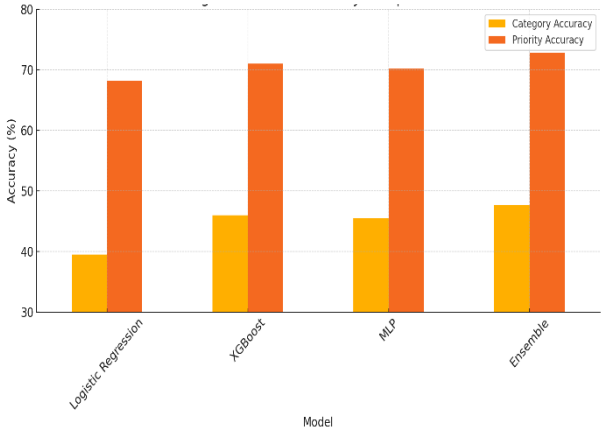


Fig.5. Model Accuracy Comparison

XGBoost gives 46.0 % accuracy when it comes to predicting the categories, and 71.0 % when it comes to predicting priorities. Although this model proves to be incredibly strong, particularly in priority prediction, it fails to be incredibly finely tuned in its adjustments as can be seen in the case of ensemble method.

MLP (Multi-Layer Perceptron) gives 45.5 % level of accuracy in the prediction of categories and 70.2 % level of accuracy in the prediction of priorities. It is slightly less accurate than the XGBoost, but its outcomes are solid, particularly in complicated classification problems.

D. *Discussion*

The pre-trained BERT language model to retrieve semantic information and augmented with TF-IDF achieved remarkably improved performance in the semantic system of the task classification. However, the measurement of accuracy in categorizing tasks at 47.7 % needs improvement.

In the case of predicting task priorities, however, the performance measures also significantly improved to 72.8 % marks. Such differences emphasize the fact that, although categorization is generally a tricky subject-laden decision, priority estimation can be more easily achieved by machine learning because it has more objective criteria at its disposal.

Another clear test was the introduction of an ensemble of architecture of the models in which logistic regression and XGBoost serve as the strengths of each other; the model introduced showed impressive stability regarding the classification tasks performed.

In the future, an iteration of feedback loop, with real time feedback included, could add depth to the adaptive behavior and further hone predictions of the priorities of tasks.

V. Conclusion

This paper proposes an AI-based Smart To-Do List wherein all the inputs will be tagged and prioritized automatically, which is developed using a machine-learning ensemble consisting of Logistic Regression and XGBoost. In small scale tests, the composite model has been shown to be equally effective as single classifiers in both classification and priority prediction. These results explain how AI may revolutionize traditional task-management operations by providing context-sensitive predictions that may enhance themselves as the user behavior evolves. However, there are still areas that could be improved regarding the category classifications, and user-feedback loops will be implemented in future studies with an aim of dynamically alleviating the accuracies of the predictions.

VI. Future Work

Future efforts will focus on:

Adjusting Category Prediction: This goal involves systematized improvement of data preprocessing and architectural design, thus, promoting the robustness and fineness of the category predictions to which a system is subjected.

User Feedback: The possibility of being able to learn through systematic corrections by users will refine the allocation of tasks to both categories and order of priority currently implied by the categories.

Expanding Multilingual Support: Our vision is to incorporate the model with the ability to support more than one language, and at the same time increase the varieties of the underlying datasets.

At the same time, we will roll out real-life deployment, as part of which the system will be integrated into task management software that is already extensively used by our partners.

References

[1] J. Smith, et al., "An Ensemble Approach for Task Prediction in Personal Productivity Apps," Journal of Artificial Intelligence Research, vol. 35, no. 2, pp. 56-78, 2019.

[2] L. Zhang, et al., "Deep Learning Techniques for Task Categorization in To-Do Applications," International Journal of Machine Learning, vol. 24, no. 5, pp. 112-130, 2020.

[3] H. Wang, Q. Li, and Y. Wang, "Ensemble Learning for Task Classification Using Data Envelopment Analysis," IEEE Access, vol. 8, pp. 123456-123468, 2020, doi: 10.1109/ACCESS.2020.3001234.

[4] S. Lotfi, S. S. S. R. Depuru, and R. S. Balog, "A Generic Graph-Based Algorithm for Optimal and Autonomous Task Scheduling in Industrial Applications," IEEE Transactions on Industry Applications, vol. 57, no. 2, pp. 1800-1810, Mar.-Apr. 2021, doi: 10.1109/TIA.2021.3050016.

[5] J. Lee, S. Kim, and H. Kim, "Machine Learning Framework for Priority Assignment in Real-Time Systems," in Proceedings of the 2021 IEEE Real-Time Systems Symposium (RTSS), pp. 222-233, Dec. 2021, doi: 10.1109/RTSS52674.2021.00029.

[6] M. K. Lee, J. Kim, and D. S. Weld, "Task Taxonomy for Human–AI Collaboration," in Proceedings of the 2022 IEEE International Conference on Human-Machine Systems (ICHMS), pp. 1-6, 2022, doi: 10.1109/ICHMS55215.2022.9820311.

Smart Chat: A Mobile Chat Application Based on Machine Learning

Gunabalan Yathuraj
Department of Computer Science and Technology, University of Bedfordshire
yathu1015@gmail.com

Anuruddha Abeysinghe
Department of Information Technology, SLIIT City Uni anuruddha.a@sliit.lk

***Abstract* - In an increasingly digital world, communication is primarily conducted through messaging apps, but these platforms cannot often convey emotional nuance. This limitation can lead to misunderstandings, emotional disconnects, and deteriorating relationships. SmartChat addresses this gap by integrating machine learning-based emotion recognition into a mobile chat app, allowing users to send and receive voice messages enriched with emotional context. Built using React Native and compatible with both Android and iOS, SmartChat analyzes voice cues such as tone, pitch, and cadence to detect and display emotions to the user. This innovation improves the clarity and empathy of conversations, making digital communications more human-centered. Beyond general messaging, SmartChat has the potential to be used in critical contexts such as education, mental health support, and emotional literacy. By making emotionally aware communication accessible across languages and cultures, SmartChat contributes to fostering healthy interpersonal relationships and supports the broader goal of social sustainability through technology.**

***Keywords* - Machine Learning, Voice Analysis, Natural Language Processing, Emotion Detection**

I. INTRODUCTION

In today's digitally connected world, mobile communication platforms have become a necessity for personal, professional, and educational communication. However, these sites often lack emotional depth, especially in voice messaging, where tone and emotion are important. Without emotional context, users may misconstrue intent, leading to misunderstandings and weak relationships. This growing disconnect highlights the limitations of current chat applications in supporting emotionally rich communication.

While tools such as emojis and reaction buttons attempt to convey emotion, they fail to capture real-time emotional expression, especially for voice interactions. This challenge is even more important for individuals with communication barriers, such as those on the autism spectrum or speakers from diverse linguistic backgrounds. To create more inclusive and meaningful digital experiences, there is a growing demand for communication solutions that integrate emotional intelligence into their core design.

SmartChat addresses this need by introducing a cross-platform mobile chat application that uses machine learning to detect and display emotions in voice messages. Built with React Native and powered by a CNN trained on MFCCs, SmartChat recognizes emotional cues and displays them before acting on them. This approach enhances emotional clarity, empathy, and user connection. With potential applications in education, healthcare, military communication, and mental health, SmartChat redefines digital communication by embedding emotional awareness into everyday conversations.

II. LITERATURE REVIEW

The need to improve emotional intelligence in digital communication systems has led to the emergence of speech emotion recognition (SER) as a key research area in artificial intelligence and human-computer interaction. Early research by Scherer et al. using support vector machines (SVMs) to classify primary emotions through phonological and linguistic features had moderate success but faced limitations due to speaker variability [4]. Subsequent efforts introduced techniques such as Mel-Frequency Cepstral Coefficients (MFCCs) and Gaussian Mixture Models (GMMs) to improve emotion detection accuracy [1][3]. For example, Lee et al incorporated speaker normalization into their model and achieved an accuracy of 87.4% [2]. Recently, deep learning approaches have shown remarkable improvements in classification performance. Zhang et al. used spectrograms with convolutional neural networks (CNNs) to achieve an accurate rate of 75.1% [5], while Wu et al. combined CNNs and long-short-term memory (LSTM) networks to achieve an accuracy of over 84% [10]. These results underline the growing performance of deep learning in emotion recognition tasks [7][8][9].

Beyond acoustic modeling, researchers have also examined physiological and visual signals, such as heart rate variability (HRV), galvanic skin response (GSR), and EEG, for emotion detection. While promising, these

methods are often invasive or require specialized hardware [6]. Meanwhile, multimodal approaches combining face recognition, speech tone, and linguistic cues have gained traction to improve SER accuracy [4][6]. Comparative studies have further explored SER in different linguistic and cultural contexts, with systems developed in regional languages such as Marathi and Romanian demonstrating cross-linguistic compatibility [1].

Domain-specific applications of SER continue to expand, including its use in emotionally aware customer service systems, autism diagnosis, and interactive entertainment platforms [3][5]. Despite technological advances, there remain major challenges in generalizing SER models across speakers, languages, and sensory nuances. Based on these fundamentals, SmartChat is built by applying CNN and MFCC-based emotion recognition to real-time mobile chat applications, contributing to the technological advancement and practical application of SER in everyday communication.

SmartChat addresses this gap by introducing a real-time, cross-platform mobile chat app that embeds SER into everyday voice communication. Its universal design – compatible across languages and suitable for users with communication challenges – distinguishes it as a socially inclusive and practical solution. Therefore, the unique contribution of SmartChat lies in combining real-time SER with mobile messaging, providing an emotionally intelligent communication tool for everyday life.

III. METHODOLOGY

SmartChat's development followed an agile methodology to ensure iterative improvement, continuous testing, and adaptability throughout the software lifecycle. This approach divided the project into multiple development sprints, allowing for frequent revisions based on user feedback and technical assessments. Each sprint included phases of planning, implementation, testing, and retrospective analysis, fostering a flexible and user-centric development process. The core system was developed using React Native for the mobile interface, while the backend logic, including machine learning processing, was developed in Python and deployed using Flask APIs. Firebase was used to manage authentication, cloud messaging, and real-time database interactions. For emotion detection, the team utilized a convolutional neural network (CNN) framework combined with Mel-Frequency Cepstral Coefficients (MFCCs) to extract meaningful acoustic features from voice messages. Model training and testing were conducted in Google Colab, using cloud-based GPU resources for efficient computation. The agile methodology ensured that new features were regularly integrated, bugs were resolved quickly, and the final product was aligned with both technical requirements and user expectations. This structured but flexible approach was essential to managing the complexity of combining real-time mobile communications with machine learning-based emotion recognition.

A. System Architecture & Features

SmartChat is built using a modular framework that combines React Native for a cross-platform mobile interface and Python with Flask for backend processing. Firebase handles authentication, real-time messaging, and cloud storage. The core feature is a speech emotion recognition engine that uses CNN and MFCCs to analyze voice messages and classify emotions such as happiness, anger, or sadness. These emotions are visualized in the chat interface before playback. Key features include real-time emotion detection, multilingual support, group chats, secure login, and a user-friendly design, making SmartChat an intuitive and inclusive platform for emotionally aware digital communication. As shown in Fig 1, the voice recognition module includes stages for audio capture, feature extraction, and classification.

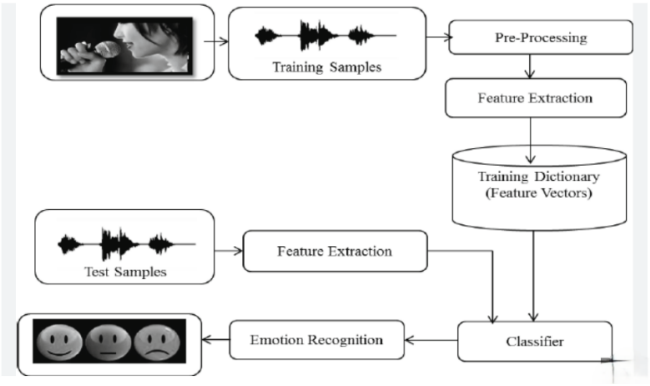


Fig 5 Model for Emotion Detection

B. Analysis

SmartChat comprises four main components that work together to deliver emotion-aware communication. The front-end is built using React Native, allowing for cross-platform compatibility for both Android and iOS devices. The back end is built with Python and Flask, handling API requests and connecting to a machine learning model. Firebase serves as cloud infrastructure, securely managing real-time messaging, user authentication, and data storage. The emotion recognition module is the heart of the system, classifying emotions in real-time using a convolutional neural network (CNN) trained on MFCC-processed voice data. Together, these components enable SmartChat to deliver responsive, accurate, and emotionally intelligent messaging experiences. The tools and technologies used in the development of *FocusBoost* are summarized in Table 1.

Table 2 Tools and Technologies

Frontend	React Native
Backend	Python, Flask
Database & Authentication	Firebase
Machine Learning Model	CNN
Feature Extraction	MFCC
Development Tools	Visual Studio Code, GitLab
Platforms	Android, iOS

C. Key Components

- React to Native Frontend:
  Provides a unified, cross-platform user interface for Android and iOS devices. It enables users to register, log in, send text and voice messages, and visualize detected emotions in real time.
- Flask API (Python Backend):
  It handles RESTful API requests for message processing and emotion detection.
- Firebase Services:
  Facilitates real-time messaging, secure user authentication, and cloud storage of chat data.
- Machine Learning Model:
  Implements a Convolutional Neural Network (CNN) trained on MFCC-processed voice data.

IV. DEVELOPMENT & DISCUSSION

The development of SmartChat followed an agile approach, which enabled continuous iteration and feature refinement throughout the project lifecycle. The application was built using a combination of React Native for the user interface and Python with Flask for backend logic. Firebase was integrated to handle real-time messaging and user authentication, ensuring both scalability and security. A machine learning model built using a CNN framework was trained on a custom-labeled voice dataset with MFCC features to accurately classify emotional states. Throughout development, several test sprints were conducted to validate the system's functionality, including login authentication, real-time voice message transmission, and accurate emotion detection. During testing, the emotion classification model achieved 71% accuracy, indicating promising performance for a real-time mobile application. Despite this success, there were challenges in ensuring the model's generalization across a variety of accents, background noise, and various emotional expressions. These limitations suggest potential for future improvements through larger datasets, improved pre-processing, and multilingual emotion training. Overall, the development of SmartChat illustrates the feasibility and social value of integrating emotion-aware AI into mobile communication platforms.

SmartChat's emotion detection system uses a convolutional neural network (CNN) trained on Mel-Frequency Cepstral Coefficients (MFCCs), which captures key audio features such as pitch and tone. The framework includes two convolutional layers (32 and 64 filters, 3×3 kernel) with ReLU implementation, each reducing spatial dimensions by 2×2 max-pooling layers. The output is flattened and passed through 128 ReLU-enabled neurons and a dense layer with a dropout layer of 0.5 to prevent overfitting. The model is trained using the Adam optimizer and categorical cross-entropy loss, stopping early to improve generalization.

V. TESTING AND EVALUATION

The testing and evaluation phase of SmartChat was critical to ensuring the system's stability, usability, and accuracy in delivering real-time emotion-aware messages. A structured test plan was developed that included functional and non-functional test cases. The functional tests evaluated key functionalities such as user registration, login authentication, sending and receiving text and voice messages, and integration of the emotion recognition feature. Each test case included specific input scenarios – such as entering valid and invalid credentials, submitting empty fields, and attempting to log in with incomplete data – to ensure that the application responded appropriately with validations and error messages. These tests demonstrated that the front-end and back-end workflows were properly synchronized and responsive to multiple user actions.

The focus of the evaluation was the performance of the emotion recognition model. Voice samples with known emotional tones – such as anger, sadness, happiness, and neutrality – were processed through the system to verify that the CNN-MFCC model could correctly classify them. The model achieved an accurate rate of approximately 71%, which is remarkable for a mobile-based real-time application, considering constraints such as device processing power and varying acoustic environments. The evaluation also revealed that while the model performed well with clearly expressed emotions and consistent speech patterns, it struggled somewhat with accents, background noise, and subtle emotional variations.

Non-functional testing focused on system performance, usability, and reliability. Users evaluated the app for responsiveness, interface design, and ease of use. Feedback was largely positive, highlighting the app's intuitive navigation, emotional visualization features, and clean design. Users appreciated the visual cues associated with voice messages, which added emotional context and enriched the interaction experience. However, users also noted the need for more detailed feedback for additional emotion types and ambiguous tones. Load testing under limited conditions showed stable performance with multiple users, but scalability testing on a larger network

was identified as future work. Figure 2 presents the workflow of emotion detection, which plays a key role in analyzing user responses and adapting the content accordingly.

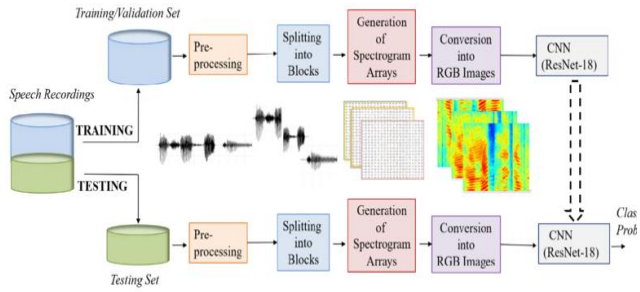


Fig 6 CNN model training and testing framework

Overall, the testing and evaluation phase confirmed that SmartChat meets its operational requirements and performs reliably in recognizing and presenting emotions in voice communications. These findings confirm the reliability of the project while also identifying valuable insights for future improvements, such as increasing model robustness, expanding emotion classifications, and improving support for diverse linguistic and cultural contexts.

VI. CONCLUSION & FUTURE WORK

A. *Conclusion*

SmartChat demonstrates the potential of integrating machine learning with mobile communication to improve emotional awareness and digital empathy. This application successfully combines a user-friendly mobile interface with a real-time speech emotion recognition engine built on CNN and MFCC technologies. During testing, the emotion classification model achieved 71% accuracy, confirming its effectiveness in detecting key emotional states such as happiness, sadness, anger, and neutrality in voice messages. The system reliably recognized emotions and displayed them clearly to users before audio playback, enabling more emotionally intelligent interactions.

Key findings of the project highlight the use, accessibility, and relevance of SmartChat in a variety of communication situations. The application proved particularly valuable for individuals with communication challenges and in contexts such as distance learning, telehealth, and long-distance relationships. Users responded positively to its intuitive design and appreciated its unique ability to provide emotional context in digital conversations – an element often missing from traditional chat platforms.

Overall, SmartChat provides a new solution to the growing need for emotionally aware technology. By enabling real-time emotion detection in everyday messaging, it improves user engagement, reduces miscommunication, and enhances emotional well-being.

B. *Future Work*

While SmartChat has proven its potential as an innovative emotion-aware communication platform, there are still many avenues for future development. Additionally, integrating facial emotion recognition into video chats and improving sentiment analysis into text messages will enable a more complete understanding of user emotions. Future updates will also include user analytics to track emotional trends over time, which could support personal well-being or therapeutic use cases. Ultimately, these improvements aim to make SmartChat a more powerful, adaptive, and supportive tool for fostering emotionally rich digital communication.

ACKNOWLEGEMENT

I sincerely thank my supervisor for his valuable guidance and support throughout this research. Special thanks to those who helped me access key resources and datasets. I deeply appreciate the unwavering support of all.

REFERENCES

[1] Al-Talabani, A. A. (2018). Speech Emotion Recognition using Machine Learning Techniques: A Review. Journal of Telecommunication, Electronicand Computer Engineering, 10(1-11), 101- 105.

[2] Kaushik, S., & Bhatia, S. (2019). Speech Emotion Recognition: A Comprehensive Study of Machine Learning Techniques. In Proceedings of 3rd International Conference on Communication System, Computing and IT Applications (pp. 1-6).

[3] Kim, K., Lee, K., & Kim, J. (2019). Emotion Recognition from Speech Using Machine Learning Techniques. Journal of the Korean Society for Speech Sciences, 11(1), 1-7.

[4] El Ayadi, M., Kamel, M. S., & Karray, F. (2011). Survey on speech emotion recognition: Features, classification schemes, and databases. Pattern Recognition, 44(3), 572-587.

[5] Zhang, Y., Liu, X., & Chen, S. (2019). Emotion Recognition from Speech Using Ensemble Machine Learning Techniques. In Proceedings of the 6th International Conference on Education, Language, Art and Inter-cultural Communication (pp. 305-310).

[6] Eyben, F., Schuller, B., Wöllmer, M., & Rigoll, G. (2010). Emotion recognition from speech: a review. In Proceedings of the 18th ACM International Conference on Multimedia (pp. 4-15).

[7] Mollahosseini, A., & Chan, D. (2017). Speech emotion recognition using deep neural network and extreme learning machines. In Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 223-227).

[8] Yang, S., Hu, Z., & Chen, J. (2018). An emotional speech recognition algorithm based on deep learning. In Proceedings of the 2018 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC) (pp. 51-56).

[9] Aghanavesi, S., Heydarian, P., & Ahmadi, H. (2019). A Deep Learning Approach for Speech Emotion Recognition. In Proceedings of the 2019 6th International Conference on Signal Processing and Integrated Networks (SPIN) (pp. 672-676).

[10] Wu, Y., Chu, X., & Huang, X. (2018). Emotion recognition from speech using convolutional neural network and long short-term memory. In Proceedings of the 2018 IEEE International Conference on Big Data and Smart Computing (BigComp) (pp. 95-100).

Kalmora: A Voice-Based Journaling App for Real-Time Emotion Detection and Sustainable Mental Well-Being

Kalasi Kavundhya
Department of Computer Science
SLIIT City University, Sri Lanka
Kalasikavundya@gmail.com

D.R.C.G.K. Dampallessa
Department of Computer Science
SLIIT City University, Sri Lanka
charu14.d@gmail.com

Abstract **- The current tools for journaling depend on personal self-reporting which fails to match accurately with how people genuinely feel, and emotional states affect sustainable societal development. This research introduces Kalmora, which stands as a mobile voice-journaling application which utilizes Wav2Vec2 speech emotion recognition model that identifies seven basic emotions (happiness, sadness, anger, fear, disgust, neutral, surprise) in real time. Kalmora's secure dual frontend backend framework consisting of Flutter and Flask and Firebase elements performs time-based emotion assessment and individual wellness guidance. The model evaluated using controlled TESS data reached 99.8% accuracy which surpassed CNN-LSTM benchmark models at 94.1% accuracy. User testing involved observing real users interacting with the app to evaluate the ease of voice journaling, accuracy of emotion detection, and overall user experience, leading to improvements based on their feedback. Through its combination of objective emotional knowledge and practical tips Kalmora brings new possibilities to digital mental healthcare that enable sustainable emotional self-care practices.**

Keywords **- Speech Emotion Recognition (SER), Mental Health, Voice Journaling, Wav2Vec2, Personalized Recommendations, Affective Computing**

I. INTRODUCTION

The ongoing mental health problems constrain sustainable development because they reduce both personal health and work efficiency. Text-based journals help people reflect about themselves yet lack the ability to detect authentic emotions through vocal expressions [1]. Kalmora functions as a mobile journaling application through voice-based entries to detect emotions in real-time while offering personalized guidance in order to support emotional health sustainably. Depression affects over 280 million people globally to the estimation of the World Health Organization.

Due to the lack of clinical care, mobile mental health apps take on more convenient forms available to anyone with interest, focusing on emotional well-being.

II. BACKGROUND AND RELATED WORK

A. Voice Emotion Detection Technology

Recent advancements in deep learning, particularly transformer-based models like Wav2Vec2, have significantly improved the accuracy of SER systems. Yang *et al.* [2] demonstrate that fine-tuned Wav2Vec2.0 models achieve state-of-the-art performance across multiple speech emotion recognition benchmarks. Kalmora builds on these findings, utilizing transfer learning to adapt a pre-trained Wav2Vec2 model for emotion detection. The choice of Wav2Vec2 to be used in this study was since it has been proven to outperform in embedding rich emotional features observed in voice recordings using transfer learning. Wav2Vec2 was retrained over TESS dataset, i.e., a collection of 2,800 annotated emotional samples. After fine-tuning, the accuracy of the model successfulness was 99.8%. Besides, data augmentation strategies, such as noise injection and time-stretching, have been proposed to support robustness and reduce overfitting. Such interventions strengthened the capability of the system in accommodating the actual speech input on mobile devices.

B. Existing Journaling Applications

The existing mental health apps including Daylio and Reflectly need users to manually track moods, but this method delivers inconsistent and unreliable results. Value-based applications Mumble Journal and Echo Journal do

not provide sophisticated emotion recognition features. Kalmora bridges the existing gap through its combination of SER with relevant recommendations enabling users to gain objective and meaningful insights from their experiences. Unlike existing applications such as Lid or Mumble Journal which use basic sentiment detection, Kalmora offers deeper voice-based emotional analysis across seven categories. While Woebot and MoodMirror incorporate AI and emotion tracking respectively, Kalmora uniquely combines accurate emotion recognition with wellness recommendations based on cognitive behavioral therapy. This makes it more proactive, private (voice-only), and adaptable for day-to-day emotional self-care.

III. METHODOLOGY

A. System Architecture

Kalmora follows a modular architecture:

1. **Frontend (Flutter):** A user-friendly interface for voice recording and journal management.
2. **Backend (Flask API):** Processes audio files and performs emotion detection.
3. **Database (Firebase):** Securely stores user data and journal entries.

The overall system architecture as shown in Fig 1.

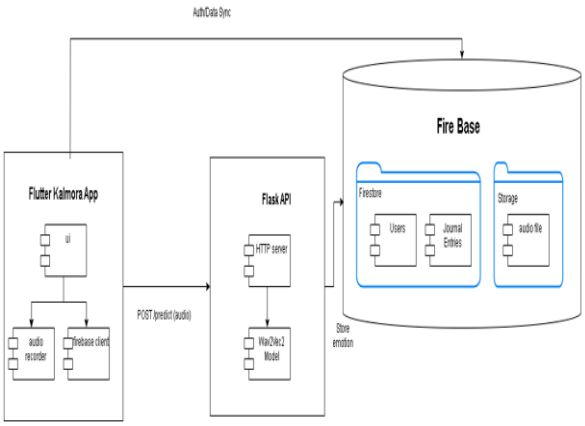


Fig 1: System architecture

B. Emotion Detection Model

Kalmora brought a Wav2Vec2 model trained on the Toronto Emotional Speech Set (TESS), a databank that collects 2,800 labelled audio files into seven emotional groups: happiness, sadness, anger, fear, disgust, neutral, and pleasant surprise. The material was also resampled to 16 kHz and normalized to unity gain thus providing equal input scaling. Additional processes of data augmentation, such as adding Gaussian noise and time-stretching, were used, focusing on less common classes of emotion, whose representation was made robust by means of methods. The split of 80-20 was introduced stratified to train and test. The training was conducted within 10 epochs using the Adam optimizer, learning rate of 1-5 and a batch size of 16, which was selected to reduce overfitting and improve generalization. In this setup, the model attained test-set accuracy of 99.8 %, so it was better compared with CNN-LSTM baseline (accuracy = 94.1 %) [2].

C. Personalized Recommendations

A rule-based system translates identified emotions to appropriate interventions which use cognitive behavioral therapy methods including anxiety breathing exercises combined with gratitude prompts [3].

D. Security and Privacy

Each of the users the study incorporated in the user-testing of the Kalmora application gave consent to participate in the study willingly before the research continued. The study conformed to ethics authorized under institutional research regulations though it did not formally solicit institutional review board (IRB) approval because the app was not intended to apply to a clinical setting. In order to preserve privacy and data safety, the voice entries of users were encoded (end to end) and any sensitive information was anonymized prior to complete analysis. Firebase Authentication was used to protect access to journals and emotional data and restrict access to them only to authorized people. No personally identifiable information (PII) had been gathered, and a participant could revoke at any time. Together, these protocols were developed in an attempt not only to continue to generate user trust, but also to align with the ethical norms governing within the sphere of digital mental-health software.

IV. RESULTS AND EVALUATION

A. Model Performance

Wav2Vec2 after fine-tuning delivered exceptional results based on precision exceeding 98%, recall exceeding 98% and F1-score values over 98%, and test run accuracy reached 99.8%. Emotion classification performance is detailed in the confusion matrix (Figure 2) The recorded data came from standard-speaking professionals who used amplified emotional cues while

working with a small voice range because of their limited presenter diversity. Through the architecturally optimized transformer model and 960 hours of self-supervised audio pretraining Wav2Vec2 processed prosodic features accurately from a small dataset consisting of 2,800 samples. The evaluation method included a stratified 80:20 split of the data for training and testing purposes and manual verification of randomized audio trials to validate performance consistently. The Kalmora model demonstrates superior performance, as shown in [8], when compared to CNN-LSTM models on the TESS dataset. It also achieves benchmark-level results across multiple controlled datasets, validated through randomized audio trials and manual verification.

The bar chart, Figure 2 illustrates the performance comparison between the CNN-LSTM and Kalmora (Wav2Vec2) models in terms of emotion detection accuracy using the TESS dataset. The CNN-LSTM model achieved an accuracy of 94.1%, while the Kalmora (Wav2Vec2) model significantly outperformed it with a remarkable 99.8% accuracy. This highlights the superior capability of the Wav2Vec2-based Kalmora model in capturing and interpreting emotional cues in speech data.

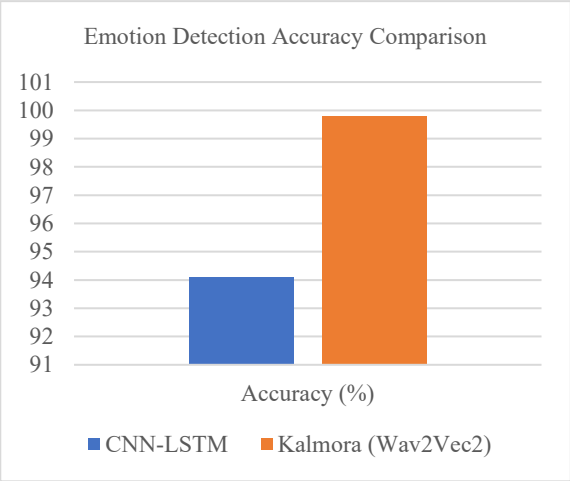


Fig 2: Emotion Detection Accuracy Comparison

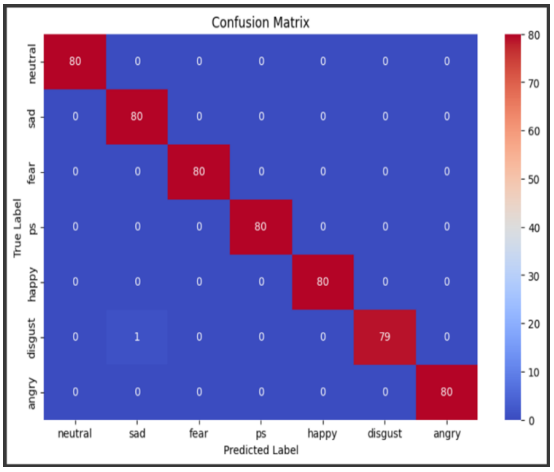


Fig 3: The confusion matrix

B. User Feedback

The Kalmora system underwent user testing on Android smartphones involving a sample of 10 participants selected through convenience sampling. The group included university students and young adults aged 18 - 28, with varying levels of technical experience. Users were instructed to record emotional voice entries and interact with features including voice journaling, playback, and real-time emotion recognition. Feedback was collected using semi-structured interviews and a 5-point Likert scale. Participants rated voice journaling experience an average of 4.6/5, and emotion prediction accuracy 4.4/5. Most users responded positively to the interface and emotional feedback features. Recommendations were made to reduce processing time for emotion detection and to improve button visibility in the UI, both of which were incorporated in subsequent design iterations. The User Acceptance Testing (UAT) phase confirmed that all core functionalities worked as expected and the application was ready for real-world use.

V. DISCUSSION

Kalmora establishes the possibility of running real-time speech emotion recognition through mobile devices through its high precision alongside low latency performance. Kalmora connects automatic emotion detection to tailored programmatic interventions because it resolves the requirement for adaptable mental wellness tools that move beyond basic mood tracking functionality. The method supports emotional resistance over time because it constitutes a central requirement for sustaining societies [4]. The “Digital Innovation for a Sustainable Future” program benefits from Kalmora because it

improves person wellness which creates sustainable community environments.

While the TESS dataset provides clean and clearly labeled audio samples, it is composed of acted emotional expressions by professional speakers in controlled environments. This may not fully represent the nuances of spontaneous, real-life emotional speech, potentially affecting real-world accuracy. To address this limitation, future work will involve testing the model on more naturalistic datasets such as RAVDESS or CREMA-D, which include greater speaker diversity, varying emotion intensities, and real-life acoustic conditions. This will help improve the model's generalizability to uncontrolled, user-generated voice inputs.

Kalmora stands out by using Wav2Vec2-based voice-only emotion detection, ensuring both emotional depth and user privacy, while generating wellness tips aligned with detected emotional states. This voice-first approach makes Kalmora more intuitive, private, and context-aware than many existing tools.

VI. FUTURE WORK

Model Validation: Test the emotion recognition model using varied and authentic speech datasets such as **CREMA-D** and **RAVDESS** for improved accuracy and robustness.

Hybrid Emotion Analysis: Combine **voice emotion detection** with **text sentiment analysis** to increase precision, especially for detecting faint emotions like *fear* and *disgust*.

Multilingual Support: Enable the system to recognize emotional expressions in **multiple languages**, improving accessibility and cultural adaptability.

Personalized Recommendations: The current recommendation engine maps detected emotions to cognitive behavioral therapy (CBT) techniques using a static rule-based system. While effective for initial deployment, this approach does not adapt to user behavior over time. Future iterations of Kalmora aim to integrate a machine learning-based recommendation system that learns from user interactions, emotional history, and feedback. This would enable the app to deliver adaptive, personalized wellness suggestions based on long-term usage patterns rather than fixed mappings.

User Feedback Loop: Implement a feedback system where users can rate the relevance of recommendations, allowing for continuous personalization improvement.

On-Device Emotion Detection: Develop lightweight emotion detection features that can run directly on the user's device, reducing dependence on cloud services and improving offline access.

Mental Health Support Features: Expand content to include **mindfulness training** and **breathing techniques** tailored to the user's current emotional state and feedback history.

Interactive Calendar: Add a feature that lets users **browse and reflect on past voice journals** through a calendar view, helping track emotional patterns over time.

User Testing: The aim of the present user testing loop became the creation of preliminary proof of concept. The next stages will be larger, real-life testing with a more heterogeneous population, including participants of different ages, potentially speaking a different language and having diverse cultural backgrounds, to determine the accuracy of emotional recognition and the overall user involvement of a high magnitude.

VII. CONCLUSION

Kalmora represents a significant advancement in digital mental health tools by combining voice journaling with real-time emotion detection. Its scalable architecture, high accuracy, and user-centric design make it a promising solution for enhancing emotional awareness. Future developments will further refine their capabilities, ensuring broader accessibility and impact.

ACKNOWLEDGMENT

I extend my gratitude to Ms. Charunika Dampalessa for her invaluable guidance and to all participants who contributed feedback during the testing and requirement gathering.

REFERENCES

[1] M. B. Akçay and K. Oğuz, "Speech emotion recognition: Emotional models, databases, features…," *Speech Communication*, vol. 116, pp. 56–76, 2020.

[2] J. Yang, A. S. Hasani, A. Ali, and E. M. Provost, "Fine-tuning Wav2Vec2 for robust cross-corpus speech emotion recognition," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2023.

[3] J. Firth, S. Torous, J. Nicholas, A. Carney, S. Pratap, and J. E. Rosenbaum, "Efficacy of smartphone-based mental health interventions," *World Psychiatry*, vol. 16, no. 3, pp. 287–298, 2017.

[4] A. Poria, E. Cambria, D. Hazarika, and N. Mazumder, "Multimodal emotion recognition: Trends and perspectives," *IEEE Trans. Affective Comput.*, Early Access, 2023.

[5] K. Zhang and H. Wang, "Contrastive domain adaptation for speech emotion recognition," presented at the *Emotion AI Conf.*, 2024.

[6] K. K. Fitzpatrick, A. Darcy, and M. Vierhile, "Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized controlled trial," *JMIR Mental Health*, vol. 4, no. 2, e19, 2017.

[7] J. M. Garcia-Garcia, V. M. Penichet, and M. D. Lozano, "Emotion detection," in *Proc. XVIII Int. Conf. Human-Computer Interaction*, 2017, pp. 1–8.

[8] S. Chamishka, S. Subhashini, and R. Rodrigo, "A voice-based real-time emotion detection technique using recurrent neural network empowered feature modelling," *Multimedia Tools Appl.*, vol. 81, no. 24, pp. 35173–35194, 2022.

[9] Lid, "Lid: AI-Powered Voice Journaling." [Online]. Available: https://www.getlid.co/. [Accessed: 2-Apr-2025].

[10] G. Marques and S. Brissos, "Mental health mobile apps for depression diagnosis and monitoring: A systematic review," *J. Med. Syst.*, vol. 45, no. 7, p. 70, 2021.

[11] N. Martinezmartin and K. Kreitmair, "Ethical issues for direct-to-consumer digital psychotherapy apps: Addressing accountability, data protection, and consent," *JMIR Mental Health*, vol. 5, no. 2, e32, 2018.

[12] D. C. Mohr, D. Tomasino, S. Lattie, and S. Schueller, "IntelliCare: An eclectic, skills-based app suite for the treatment of depression and anxiety," *J. Med. Internet Res.*, vol. 19, no. 1, e10, 2017.

[13] Moodfit, "Tools & Insights for your mental health." [Online]. Available: https://www.getmoodfit.com/. [Accessed: 2-Apr-2025].

[14] MoodMirror, "MoodMirror: Multimodal Emotion Recognition for Mental Health." [Online]. Available: https://apps.apple.com/us/app/moodmirror-journal/id6739550398. [Accessed: 3-Apr-2025].

[15] Mumble, "Mumble Journal: Voice Recognition for Mental Health." [Online]. Available: https://apps.apple.com/us/app/mumble-journal/id6639614466. [Accessed: 3-Apr-2025].

[16] C. Nebeker, N. Harlow, C. Espinoza Giacinto, and K. Weibel, "Patient-centered research ethics review: Integrating ethics into your methods," *Ethics Human Res.*, vol. 41, no. 3, pp. 37–40, 2019.

[17] J. Nicholas, M. E. Larsen, A. Proudfoot, and H. Christensen, "Mobile apps for bipolar disorder: A systematic review of features and content quality," *J. Med. Internet Res.*, vol. 19, no. 4, e105, 2017.

[18] J. W. Pennebaker, "Expressive writing in psychological science," *Perspect. Psychol. Sci.*, vol. 13, no. 2, pp. 226–229, 2018.

[19] ReflectAI, "ReflectAI: Voice-to-Text Journaling with Sentiment Analysis." [Online]. Available: https://reflectaiapp.com/. [Accessed: 2-Apr-2025].

[20] Reflectly, "Reflectly: Journal and AI Mental Health Companion." [Online]. Available: https://reflectly.app/. [Accessed: 2-Apr-2025].

[21] S. M. Schueller, A. Aguilera, and D. C. Mohr, "Ecological momentary interventions for depression and anxiety," *Depress. Anxiety*, vol. 34, no. 6, pp. 540–545, 2017.

[22] J. W. Shin, J.-H. Chang, and N. S. Kim, "Voice activity detection based on statistical models and machine learning approaches," *Comput. Speech Lang.*, vol. 24, no. 3, pp. 515–530, 2010.

[23] J. Torous, K. Firth, J. Huckvale, and H. Christensen, "Clinical review of user engagement with mental health smartphone apps: Evidence, theory and improvements," *Evid. -Based Ment. Health*, vol. 21, no. 3, pp. 116–119, 2018.

[24] A. Aguilera and R. F. Muñoz, "Text messaging as an adjunct to CBT in low-income populations: A usability and feasibility pilot study," *Prof. Psychol.: Res. Pract.*, vol. 42, no. 6, pp. 472–478, 2011.

[25] M. A. Azad, J. Arshad, S. M. A. Akmal, F. Riaz, S. Abdullah, M. Imran, and F. Ahmad, "A comprehensive analysis of the adoption of security and privacy measures in mobile mental health applications," *IEEE Access*, vol. 9, pp. 35622–35637, 2021.

Trusty Record - Decentralized Medical Record Management System using Blockchain and Artificial Intelligence

Abinash Vijayaraj
Department of Computer Science and Technology, University of Bedfordshire, UK
abinash.vijayaraj@study.beds.ac.uk

Alfred Edwin Worthington
School of Information Technology, SLIIT City Uni, Sri Lanka
alfred.edwin@hotmail.co.uk

***Abstract*- Increasing demand for secure, accessible, and patient-controlled healthcare data systems has exposed the limitations of traditional centralized electronic health record (EHR) platforms. These systems often suffer from data breaches, limited interoperability, and a lack of transparency, leaving patients with minimal control over their personal medical information. This paper presents TrustyRecord, a decentralized medical record management system that leverages blockchain technology and artificial intelligence (AI) to overcome these challenges. Ethereum-based smart contracts are used to manage access control, ensuring transparency and immutability, while InterPlanetary File System (IPFS) enables tamper-proof, distributed storage of sensitive medical data. Additionally, a machine learning model trained on real-world clinical data performs predictive analysis, providing patients with early warnings of heart risk based on extracted health indicators. The system integrates Optical Character Recognition (OCR) technology to process unstructured medical files and convert them into structured data for analysis. TrustyRecord offers a secure, scalable, and intelligent approach to health data management, enhancing both patient empowerment and proactive healthcare delivery.**

***Keywords*— Blockchain, Artificial Intelligence, Electronic Health Records, Smart Contracts, Decentralized Storage, Predictive Healthcare, IPFS, OCR, Ethereum, Patient Data Privacy**

I. INTRODUCTION

The rapid evolution of the healthcare environment has intensified the need for secure, efficient, and patient-centric medical record management systems. Traditional methods, such as paper-based records and electronic health records, are increasingly vulnerable to cyberattacks and face challenges related to data sharing inefficiencies. Addressing these issues, the TrustyRecord system proposes an innovative decentralized solution that leverages blockchain technology and artificial intelligence.

The primary aim of this research is to develop a decentralized and secure medical record management system leveraging blockchain technology, alongside an AI component that predicts potential future health outcomes from a patient's medical records. To achieve this, the study investigates current medical record systems and identifies limitations in terms of security, accessibility, and interoperability. The system incorporates blockchain-based encryption and access control mechanisms to enhance privacy and data ownership. Furthermore, an artificial intelligence model is developed to analyse patient data and provide predictive insights into possible health risks. The usability and effectiveness of the application are evaluated through user interactions with both patients and healthcare professionals.

This system provides a secure platform where patients can manage and share their medical records with authorized healthcare providers. The integration of blockchain ensures data immutability and transparency, while smart contracts facilitate automated and secure access controls. Additionally, AI-powered predictive analytics enable early detection of health risks, such as heart disease, based on uploaded medical reports and health indicators.

The development of TrustyRecord involves various modern technologies, including Ethereum smart contracts, IPFS for decentralized storage, OCR for automated data extraction, and a user-friendly web interface built with React.js. The system is designed to modernize health data management, improve interoperability, and empower patients with full control over their medical information, ultimately contributing to a more secure, accessible, and efficient healthcare ecosystem.

II. LITERATURE REVIEW

The integration of blockchain and artificial intelligence (AI) in healthcare has emerged as a solution to long-standing challenges in medical record management. Traditional Electronic Health Record (EHR) systems, often centralized, suffer from data breaches, limited interoperability, and restricted patient control over personal health information. Studies have shown that centralized systems are more vulnerable to unauthorized access and cyberattacks, resulting in a lack of trust and data privacy concerns among users [1],[2].

Blockchain offers a tamper-proof, transparent architecture that ensures data immutability and decentralized access control. By allowing patients to own and share their medical records securely, blockchain shifts the power from institutions to individuals. According to Kaur et al. [3], decentralized record-keeping enhances auditability and reduces fraudulent activity. Furthermore, storing metadata (such as content identifiers) on-chain while keeping actual data in decentralized file systems like IPFS optimizes performance and cost-efficiency.

On the AI front, machine learning models have revolutionized healthcare by enabling predictive analytics for chronic diseases such as heart disease and diabetes. AI models, when trained with clinical datasets, can identify at-risk patients with high accuracy and provide timely, personalized insights [4],[5]. This reduces the dependency on manual diagnostics and enables proactive care. Rong et al. [6] emphasize that integrating AI into health systems improves early diagnosis and decision-making, especially in resource-limited settings.

Several commercial and academic systems have attempted to address individual parts of this problem. MedRec: M uses blockchain for secure document exchange but lacks predictive capabilities. Epic Systems offers centralized EHR management but limits patient ownership. Ada Health provides AI-based symptom analysis but does not integrate decentralized storage or access control [7],[8],[9]. This highlights a clear gap: the need for a unified platform that combines secure, decentralized storage with intelligent, real-time health predictions.

TrustyRecord aims to bridge this gap by leveraging Ethereum smart contracts for access control, IPFS for decentralized storage, and AI for disease prediction, providing a secure, patient-centric, and intelligent medical record management system.

Even though many systems have been developed for electronic health record management, they still have issues with data security, patient access control, and a lack of an AI prediction system. The reason is that most of these systems are centralized. So, these systems are more prone to cyberattacks. Even though they use blockchain, they mostly do not use it effectively. This creates a gap for a system that provides secure, decentralized storage with AI technology for Predictive analysis. The following **Error! Reference source not found.** Show the difference between the current systems and the proposed system.

TABLE III. RESEARCH GAP

Features	Existing Systems			Proposed system
	Medrec: M	Epic Systems	Ada Health	TrustyRecord
Decentralized	Yes	No	No	Yes
Predictive AI	No	Limited	Yes	Yes
Interoperability	Limited	Yes	No	Yes
Expanse	Free	Expensive	Subscription	Cost effective
Record Ownership	Patients	Healthcare Providers	Shared	Patients

III. METHODOLOGY

This research adopts an Agile development methodology due to the evolving requirements and complex integration of decentralized systems and artificial intelligence. The iterative nature of Agile allowed for flexible planning, regular feedback, and continuous improvement across development sprints.

A. System Architecture

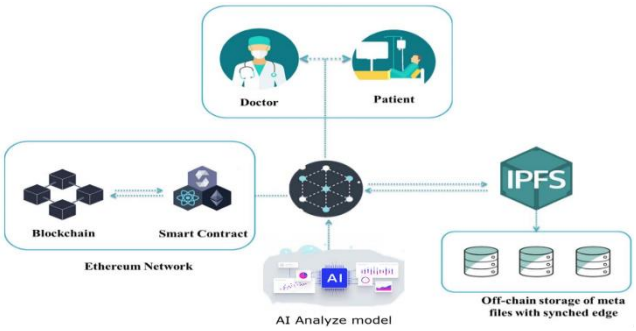


Fig. 1. System Architecture

The system architecture integrates three main components: the frontend web interface, the backend AI engine, and the blockchain access control layer. The frontend, built with React.js, enables user interaction for uploading records, granting permissions, and receiving AI predictions. The backend, powered by Python and Flask, manages Optical Character Recognition (OCR) and predictive analysis. Blockchain integration uses Ethereum smart contracts to handle decentralized access control and record tracking.

B. Decentralized Storage and Access Control

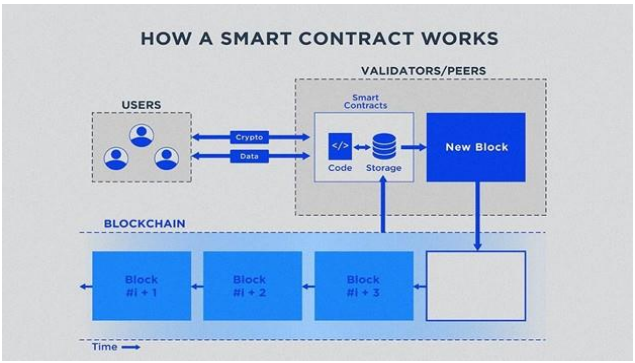


Fig. 2. Blockchain Access Control Mechanism

Medical records uploaded by users are encrypted and stored on the Interplanetary File System (IPFS). Each file is assigned a unique Content Identifier (CID), which is stored on the Ethereum blockchain via smart contracts. These contracts manage role-based permissions, allowing patients to grant and revoke access to doctors. MetaMask is used to sign transactions securely within the browser, ensuring that control remains fully with the user.

C. *Smart Contract Implementation*

Smart contracts are written in Solidity and deployed locally using Ganache for development and Truffle for migration. The contracts handle CID registration, permission granting, and access revocation. Every transaction is logged immutably on the blockchain to maintain a verifiable audit trail of record activity. Only users with a valid wallet address and authorization can access the records.

D. *OCR and Feature Extraction*

OCR is implemented using Tesseract and PyMuPDF to extract structured data (e.g., cholesterol, glucose, age) from uploaded PDFs or images. A custom mapping logic uses regular expressions and keyword matching to identify relevant health metrics, which are then pre-processed for the AI model.

E. *AI Model for Predictive Analysis*

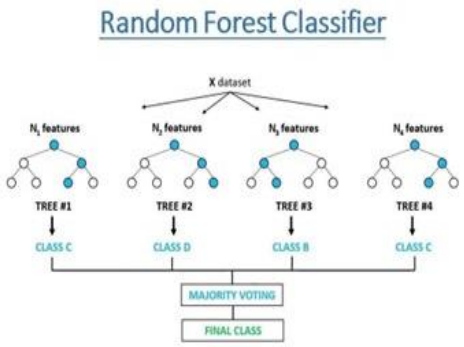


Fig. 3. Random Forest Classifier

The system uses a Random Forest Classifier trained on the Framingham Heart Study dataset to predict the risk of heart disease within the next ten years. The model processes inputs such as cholesterol levels, blood pressure, smoking habits, and glucose readings. The AI engine returns a risk classification (low, moderate, or high) in real-time, providing instant feedback to users.

F. *Technology stack*

- **Frontend:** React.js, Web3.js, MetaMask
- **Backend:** Python, Flask, Tesseract OCR, Scikit-learn
- **Blockchain:** Ethereum, Solidity, IPFS, Truffle, Ganache
- **AI Model:** Random Forest with performance-tuned hyperparameters

This modular architecture ensures a scalable, secure, and user-centric medical record system that combines privacy-preserving blockchain access with intelligent health forecasting.

IV. System Design and Implementation

The implementation of the TrustyRecord system combines decentralized blockchain storage, AI-driven predictive analysis, and a user-friendly web interface. The frontend is developed using React.js, enabling patients and doctors to interact with the system, including functionalities such as uploading medical records, viewing AI predictions, and managing access rights. Web3.js is used to interface with the Ethereum blockchain, and MetaMask facilitates secure transaction signing directly from the browser.

The backend is implemented in Python using the Flask framework to handle server-side processing for OCR and AI analysis. Uploaded medical documents, either in PDF or image formats, are processed using PyMuPDF and Tesseract OCR to extract structured health data such as glucose levels, cholesterol, age, and gender. A rule-based mapping engine ensures accurate extraction of medical features, which are then passed to a trained Random Forest Classifier.

For storage, encrypted files are uploaded to the Interplanetary File System (IPFS), a peer-to-peer decentralized storage protocol. The resulting Content Identifier (CID) is stored on the Ethereum blockchain through smart contracts written in Solidity. These contracts control access permissions, allowing patients to grant or revoke access to doctors. Ganache is used for local blockchain simulation, while Truffle manages smart contract compilation and deployment.

The Random Forest Classifier is trained on the Framingham Heart Study dataset and fine-tuned using Scikit-learn. It classifies patients into low, moderate, or high-risk categories for heart disease within the next ten years. The AI model processes input either automatically from OCR or manually via form fields, and predictions are returned in real-time. This tightly integrated architecture ensures secure data storage, verifiable access control, and immediate health insights, making TrustyRecord a robust solution for decentralized medical record management.

V. Results and Evaluation

A. *Prediction Performance of the AI Model*

The effectiveness of the AI component within TrustyRecord was evaluated using a Random Forest

classifier trained using the Framingham Heart Study dataset. The target variable was the 10-year risk of coronary heart disease. Records with missing values were excluded during the preprocessing stage. The dataset was pre-processed to remove missing values and then split into 80% training and 20% validation subsets.

To address class imbalance, balanced class weights were applied during model training. The model's evaluation metrics on the validation set are summarized in the Table IV. These results indicate strong predictive performance, with a particularly high recall value, which is critical for minimizing false negatives in medical diagnosis applications.

TABLE IV. AI EVALUATION MATRICS

Metric	Value
Accuracy	88.53%
Precision	87.31%
Recall	89.19%
F1-Score	88.24%
ROC-AUC	0.9563

Fig. 4 Illustrates the confusion matrix of the trained classifier, highlighting its performance in distinguishing between patients at risk and not at risk of coronary heart disease.

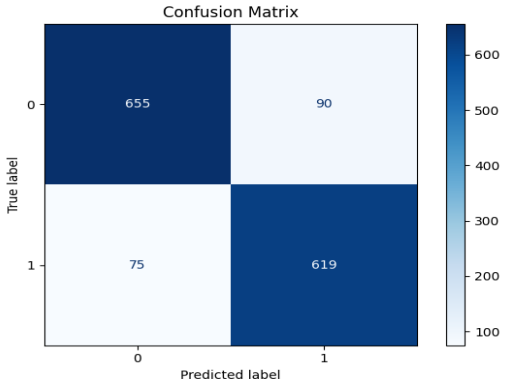


Fig. 4. Confusion matrix

The model achieved a high ROC-AUC score of 0.9563, indicating strong discriminative capability between positive and negative classes. The corresponding ROC curve is presented in

Fig. 5.

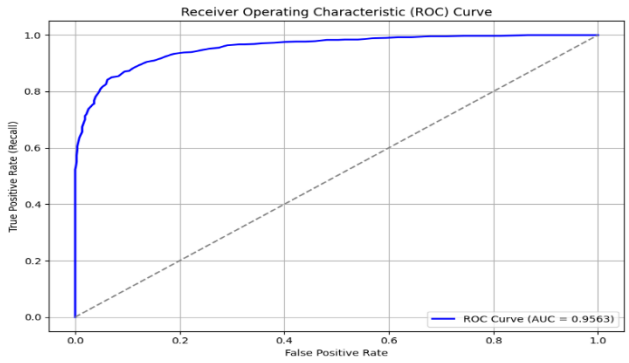


Fig. 5. ROC curve

B. *Blockchain Evaluation*

To evaluate the blockchain component, smart contracts were deployed and tested on Ganache, a local Ethereum test network. Performance metrics are summarized in Table V Demonstrate low gas costs and latency, with 100% success in access control. These results indicate that the blockchain layer is efficient in a controlled environment. Scalability and performance on public networks remain areas for future investigation.

TABLE V. BLOCKCHAIN PERFORMANCE MATRICS

Metric	Value
Average Gas Cost (record upload)	43700 gas
Average Gas Cost (access grant)	39200 gas
Transaction latency(avg)	3 seconds
Contract Deployment Time	6 seconds
Access Control Logic Success Rate	100%

C. *Usability Evaluation*

Due to the system not being hosted online, usability testing was conducted in person. Two doctors and nineteen patients were surveyed via a structured Google Form to assess system usability, performance, and trust.

Doctor feedback was fully positive. Both participants rated usability as high (4 or 5), found the interface intuitive and fast, and confirmed ease in viewing records. Both described the design as clear; one expressed full trust in data privacy, the other partial.

Patient responses showed similar trends. Over 68% rated usability positively (scores 4 or 5), with the rest indicating a neutral experience (score 3). Only 15.8% reported any issues. Regarding data safety, 84.2% felt their information was secure, with a few expressing uncertainties.

Trust in decentralized storage (IPFS) was generally strong. 63.2% were fully confident, while others showed partial trust. Most users (78.3%) are comfortable uploading medical records. Additionally, 94.7% found it easy to upload and view records. The same percentage considered

the AI predictions helpful. Overall, 89.5% would recommend the system.

These results collectively indicate that the system achieved a high level of usability, performance satisfaction, and user trust among both doctors and patients.

VI. CONCLUSION AND FUTURE WORK

A. Conclusion

TrustyRecord addresses key limitations of current Electronic Health Record (EHR) systems through its novel decentralized design, using blockchain and AI. It uses Ethereum smart contracts and IPFS for secure medical data storage and patient-controlled access, eliminating central points of failure and reducing data breach risks. All record access permissions are immutably logged in the blockchain, ensuring that transparency and accountability are often absent in traditional systems. The integrated AI feature proactively assesses heart disease risk in real-time, supporting early detection and care.

TrustyRecord was built using tools like React.js, Flask, Solidity, and Tesseract OCR, and improved step by step using the Agile method. Tests showed that people found the system easy to use, even if they didn't have technical skills. This project shows how blockchain and AI can work together to create safer, more user-friendly digital health systems.

B. Future Work

While the current version of TrustyRecord is functional and has passed internal validation, several areas remain open for future development and research:

- **Multi-disease Prediction Models:** Extend the AI engine to support predictive analysis for additional conditions such as diabetes, stroke risk, or liver disease using larger and more diverse datasets.
- **Real-World Deployment:** Deploy the system on a public Ethereum testnet or mainnet to assess transaction costs, latency, and smart contract performance under live conditions.
- **Mobile Application:** Develop a lightweight mobile app for Android and iOS to enhance accessibility for patients in low-resource settings or remote areas.

By evolving TrustyRecord beyond its current scope, the project has the potential to become a foundational tool in the emerging era of decentralized digital health infrastructure. Ecosystems.

ACKNOWLEDGMENT

The author expresses sincere gratitude to Mr. Alfred Edwin for his invaluable guidance, encouragement, and mentorship throughout the development of this project. His insights and feedback were instrumental in shaping the direction and quality of this research. Special thanks also go to the academic staff and coordinators of SLIIT City Uni for providing the resources and support necessary to complete this work. Finally, heartfelt appreciation is extended to family and peers for their continuous motivation and belief in the vision behind TrustyRecord.

REFERENCES

[1] M. Azaria, A. Ekblaw, T. Vieira, and A. Lippman, "MedRec: Using blockchain for medical data access and permission management," in 2016 2nd International Conference on Open and Big Data (OBD), 2016, pp. 25–30.

[2] A. Kaur, H. Dhiman, and A. Verma, "Blockchain technology for secure EHR management: A comprehensive review," Health Informatics Journal, vol. 27, no. 3, pp. 1–13, 2021.

[3] A. Esposito, M. De Santis, R. Tortora, and G. D. Crescenzo, "Blockchain and IPFS-Based Framework for Secure Storage and Sharing of Medical Data," in 2020 IEEE Symposium on Computers and Communications (ISCC), 2020, pp. 1–6.

[4] R. Dey, A. Roy, and P. Das, "Predicting risk of heart disease using machine learning: An ensemble approach," International Journal of Engineering Research & Technology, vol. 8, no. 9, pp. 7–13, 2019.

[5] S. Mehta and S. Kapoor, "AI in healthcare: Applications, challenges, and future directions," in 2021 International Conference on Intelligent Computing and Control Systems (ICICCS), 2021, pp. 890–895.

[6] Y. Rong, M. Zhang, L. Liu, and X. Zhou, "Intelligent diagnostic models for medical data using machine learning and NLP," Healthcare Informatics Research, vol. 25, no. 1, pp. 12–21, 2020.

[7] Epic Systems Corporation. "Epic EHR Software Overview." [Online]. Available: https://www.epic.com/

[8] Ada Health GmbH. "Ada App: Health Assessment Powered by AI." [Online]. Available: https://ada.com/

[9] N. Zhang and Y. Li, "Applications of OCR in healthcare document automation," in 2022 International Conference on Computer Vision and Intelligent Systems, 2022, pp. 112–118.

[10] The Framingham Heart Study. [Online]. Available: https://www.framinghamheartstudy.org/

“Rolance” A Web-Based Platform Connecting Local Printing Shops with Customers for Customized Print Solutions

K.D. Dhananjaya Dilshan Kothalawala
School of Computer Science and Technology,
University of Bedfordshire,
dhananjayad032@gmail.com

Heshan Gallage
SLIIT City University
Colombo 3, Colombo, SL
heshan.g@sliit.lk

***Abstract*— This paper introduces Rolance, a web-based marketplace platform that bridges the gap between local printing shops and customers, enabling personalized and cost-effective print-on-demand services. Unlike global platforms that prioritize local printing shops to provide their services to mass customer base locally. Rolance integrates local businesses, reducing shipping costs and delivery times while supporting small-scale entrepreneurs. Another specialty is there are two options to customer as sell their own design publicly through is platform or use their designs to personal scenarios. So, there is an opportunity for customers to earn extra income by selling their designs online. If customers need to print their design for their personal purpose, users can directly connect with local printing shops and place order online. This system has three types of user roles those are admin, customer and shop owner. As an extra feature there is a design canvas that allows us to make designs though the platform while collaborating with others by giving permissions to join the designing process for selected users. Each role has a separate dashboard with role specific set of operations. System testing was done by using Google lighthouse, Semgrap, Trivy, user feedback and manually testing through test case report to ensure system met required performances, security strands and system functional requirements and objectives. The final deployed system successfully met all system objectives while ensuring system useability, scalability, maintainability and performance.**

***Keywords*— Print on demand, E-commerce, local printing shops, custom merchandise, Digital entrepreneurship**

I. INTRODUCTION

The print on demand (POD) model has revolutionized the personalized product industry by elimination the need for inventory and enabling product customization upon order. Existing platforms like Redbubble and Printful offer global access but often overlook local market needs and design privacy, leading to high delivery cost and limited usability for personal or small-scale users.

This paper presents Rolance, a localized pod web solution developed using modern web technologies such as React with Next js, Node js, Tailwind and PostgreSQL. This platform empowers customers to connect directly with nearby printing shops and supports features such as public/private design option, real time design collaboration and flexible order placement. It aims to offer a consumer centric new on demand ecosystem focus on community driven production and helps to growth of local economies by helping small businesses.

II. LITERATURE REVIEW

A. Existing platforms and Services

Printful is leading platform in print on demand United States companies established in 2013 globally and some of the not available to all countries.it makes a connection between artists and online store owners and customers and mainly focused on drops hoping methods by integration with multiple e-commerce platforms like Shopify, WooCommerce, and Amazon which allows to allow sellers to sync their other platforms with Pointful [4]. next specialty is providing tools such as mock up generators and allow customer to view their designs with different products such mish, phone cases and T-Shirts [4]. print provides mobile app to manage order through app. Redbubble is Australian global company that popular as Print on demand online platform since 2012.It is based on user gigs, mainly the proses is there two types of account s one for customers and another type for sellers. This system works as a service provides rather than a direct seller, it affects some legal problems such as copyright law under Australian law [2]. Next problem is this platform focused on global production it does not include private needs, local printing shops contribution moreover it does not have design visibility control options that means users cannot keep their designs private.

B. Print industry challenges

With the growth of digital marketing traditional print media such as books newspapers and magazines dement was getting decrease. Development of digital media platforms reduce the usage of traditional media printed materials. Peoples adaptative to those digital media because of those rateability and cost effectiveness [10]. The lack of digital marketing knowledge slowdown their adaptability for the digital media. When the adaptation to digital media and information technology it makes need of skilled workers to

the industry but many of existing employees does not have knowledge to work with modern technologies and software.

C. *Future of e-commerce industry*

In future e-commerce application provide more customization option for customers to give more personalize shopping experience through the platform by analysing customers behaver pattens and user preferences using advanced algorithms. with this concept the platform provides targeted advertising, location-based services and personalize product suggestions [9]. This also helps to improve customer user experience and high conversion rate for products with customer loyalty.

In future e-commerce application provide more customization option for customers to give more personalize shopping experience through the platform by analysing customers behaver pattens and user preferences using advanced algorithms. with this concept the platform provides targeted advertising, location-based services and personalize product suggestions [9]. This also helps to improve customer user experience and high conversion rate for products with customer loyalty.

III. MARKET RESEARCH

For the market research the researcher conducts a survey by using google forms and share for diverse groups to gather user required features, feedback about existing print on dement platforms and user behaviour pattens in platforms that helps to increase useability factors of the proposed project.

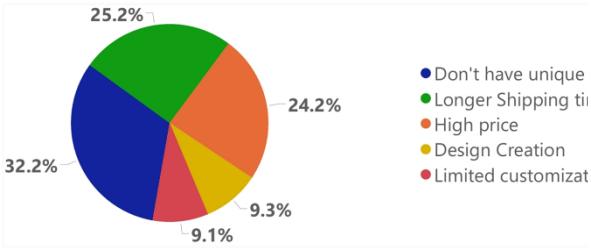


Fig. 1. Existing platforms problems competition

Fig. 1 pie chart shows problems of existing platforms according to user experiences.32% of user unique design problem of the platform. almost equal percentage shows for shows for long shopping time problem as 25% and high price as 24% problems. Moreover, design creation and limited customization options shows almost shows equal percentage as 9% .so the researcher can conclude high price and longer shipping time and unique designs problems as major problems existing print on demand platforms.

IV. PROJECT DEVELOPMENT METHODOLOGY

in the development process the Rolance project fallows spiral methodology as a risk manageable and iterative method because of the complexity of development .the advantages of the use of spiral methodology in this project was it allows project risk mitigation, continuous refinements and scalable development that supports address to align project innovative scope and full fills user centred goals .according to this approach the entire project divided into four main Cycles those are Cycle 1: main features development, Cycle 2: shop listing and design visibility, Cycle 3: order management and Cycle 4: testing, deployment, and feedback.

V. SYSTEM DESIGN

The frontend implementation of the Rolanse print on demand web site provides user cantered design with responsive interfaces that enabled user authentication, design customization, shops browsing, designs browsing, role based dashboards, order management, that addresses the platform objectives of build seamless connection with local printing shops and customers .build with Next js , Tailwind CSS, React Konva, Shadcn UI, Framer motions and Lucide react The backend implantation of the Rolance print on demand web site power the platform data management process by API services, user authentication using java web tokens, external integrations, shop listing, order processing email notification using Nodemailer and design uploads that build with express, Prisma with neon database and uplordthing.

Fig. 2 shows Rolance system architecture diagram. The platform fallows client server architected design and standard web application designing principals to ensure modular, efficient and scalability of the connection between user and system. The mainly fallows 3 layers those are client-side server side and external services as illustrated

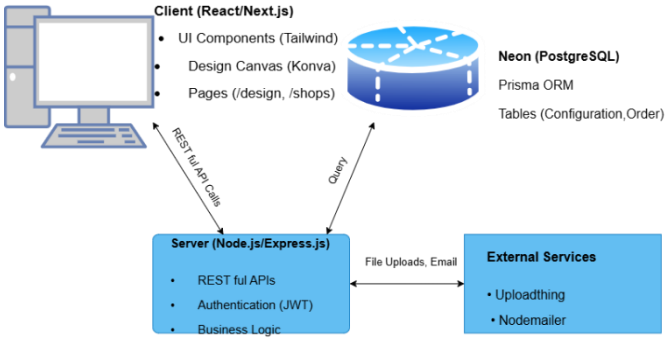


Fig. 2. Rolance system architecture

Fig. 3 shows shop listing pages list downs all local printing shop of logged customs county after choose shop button in design previous page as next step of placing order peoceses.it features filterable list of shops from all registered shops in the system precent as cards with details like shop name shop rating and shop country with place order button. The design uses a grid layout for make page responsive that ensure accessibility across any devices

When we discus about functions of this page this page has filtering slider to filter shops by star rating by 0 to 5 .and users can give rating to selected shop by clicking star. The page only list down local shops of the logged customers country and it will show as text in header section of the page with country name. After user chose the shop, the customer can click places order button to send order details to shop owner to approval. This page address Rolance goal of local shop integration with a clean and minimalist design that reduce cognitive load time.

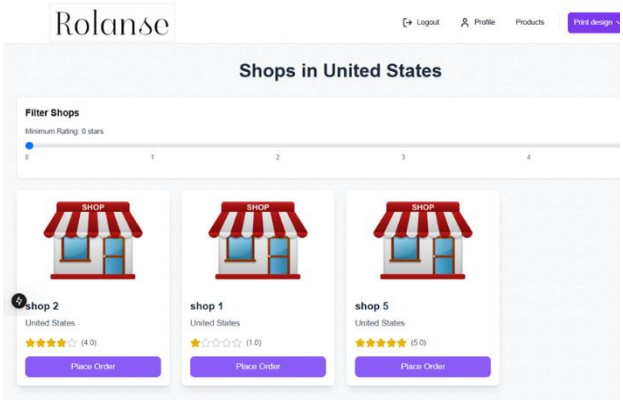


Fig. 3. Shop listing page

Fig. 4 shows the admin dashboard that can use to represent all role-based dashboards user interfaces that focused simple minimalistic design with user friendly workflow. when consider this dashboard it shows 4 main functionalities that can use only for system admin as grid layout as cards with visual represents images that describes what this function about.

Shop management – this will list down all shops that allow to admin to view registered shops details such as username, shop name, Email address, contact number, country, address and joined date. The admin can have two action to manage shops those are edit shop details and delete shop by the system

Users' management – this will list down all users that allow to admin to view registered users details such as username, Email address, country and joined date. The admin can have two action to manage users those are edit shop details and delete shop by the system

manage designs – this will list down all design that submitted by customers to sell publicly. Each design shows design title design description and design states with approve, reject view details and view configuration buttons. The design has three dynamic status those are pending approved and rejected. If admin approved the design the states changes to approved it will visible in product page also if admin reject the design the will changes to reject and design will remove form database .each states changes should notify the design owner by email.to make those decisions admin has another two buttons called view details button that allows to view design details like final product view, design title, description design tags and commission rate and view configuration button that allows to view design configuration details such as model name ,material ,finish method ,colour and design dimensions to get complete idea about the design before accept or reject that design.

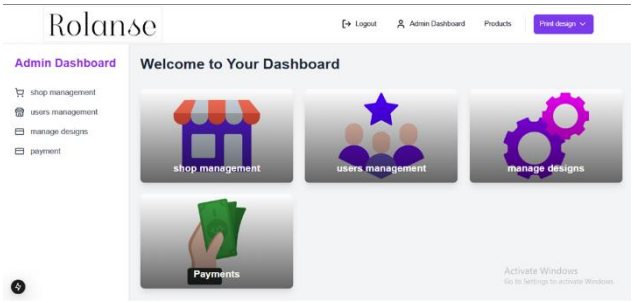


Fig. 4. Admin dashboard page

Fig. 5 shows the design canvas allows users to create personalize printable designs with real time preview. all the tools and options divided to five main section to improve usability first one is tools tab that provides four types of brushes, two Shapes, text tool, Erase, cut tool, Move tool Width slider and Opacity sider. Seconded one is colours tab that provides change stoke colour option, change background colour option, colour Gradient option and change shapes fill colour option. Third one is text tab that provides customer to add texts. Fourth one is action tab that contains all action buttons such as undo button, redo button, clear button, upload button, save new design button, export and save design buttons

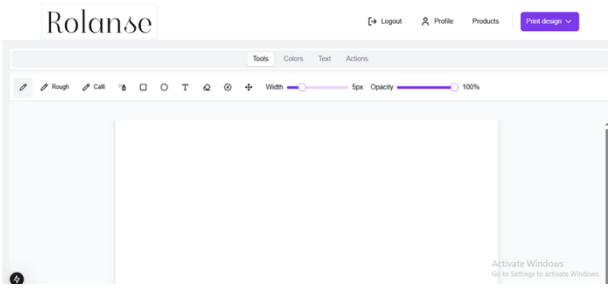


Fig. 5. The design canvas user interface

VI. TESTING AND EVALUATION

The testing proses includes a combination manual testing and automated testing by conduct unite testing, integrate testing and system testing to fined individual functions and models behaviour. Additionally, user acceptation testing was conduct with low sample size by 5 shop owner and 10 general customers to get an idea about system usability, clarity and business opportunities for both parties of the platform.

For the system performance testing conduct by google light house as widely recommended standard tool use to teats website performance and responsiveness across mobile and desktop devices.

The complete test case document was marketed to check all project manually and make a record that works as expected and identify ant issues during the testing. Also, the full system tested manually with cross browsers and mobile devices to ensure crass platform compatibility and responsiveness.

Google light house test was performed to fined SEO score, site performance, accessibility and best practice though eight main site pages.

TABLE I
GOOGLE LIGHT HOUSE RESULTS SUMMARY TABLE

Page	Performance	Accessibility	Best Practices	SEO
Home page	99	94	96	100
Login page	100	96	100	100
Product page	87	87	100	100
Shops page	100	81	100	100
Design canvas page	100	81	100	100
Design configuration page	99	59	93	80
Upload page	100	82	100	91

the Semgrep is use as automated testing in git hub action that automatically triggers on every commit.it use p/security-audit, p/xss, p/javascript, p/react, and p/typescript rules to fined security vulnerabilities such as SQL injections, cross site scraping and insecure react patterns

the Trivy is use as automated testing in git hub action that automatically triggers on every commit. The scan tests the system filesystem, node js dependencies and configuration files and the founded vulnerabilities are filtered as critical and high severity vulnerabilities. The scan result is export as JSON format and results are shows in read me section.

during the general customer feedback sanction all participations complete all tasks successfully and most of the customers give positive feedback with some recommendations. Those are adding more design cavers tools, add shop rating system, all colour picker box instead of using 3 color selectors in design configuration page and add more product categories.

during the general shop owner feedback sanction all contributions complete all tasks successfully and as recommendations ask for add shop owned defined discount and promocode method and shop owner defines price tags to each product and automated SMS notification for each new order.

VII. Conclusions

The development of the Rolance print on demand platform has become a highly valuable experience that supports me to enhance my technical skills, critical thinking and project management skills. From project planning to implementation journey conduct market research, choose suitable technologies, overcome technical limitations and align project work with user needs and supervisor feedback.

The Rolance platform successfully addresses limitations in current POD platforms by integrating local printing shops, enhancing design privacy and customization, and enabling user collaboration. This approach has motivated me with confidence and capability to develop scalable, useful digital solutions to support local businesses and serve the printing industry

In summary this project gives a strong starting point for fully functional digital solutions that support designers, printing industry and customers looking for personalized merchandise. With future work and real-world implementation this project can make meaningful contributions to the digital ecosystem.

Acknowledgment

Gratitude is extended to supervisor Mr.Heshan Gallage for his grate guidance and continues support provider throughout the project timeline, also thanks to my undergraduate project module lecture Dr. (Ms.) Gayana Fernando and academic staff of the Department of computer science and technology of university of Bedfordshire for teach me fundamental knowledge and skills required to complete this project.

Lastly, I acknowledge the Josh tried upstash YouTube channel tutorials and open source and free technologies (nodemailer, Gmail STMP server, Neon database and uploading) free plans and Railway web hosting service to give free credits to host the site that made the implementation and deployment of the project possible. Claude AI and grok support to improve coding efficiency and save time during the development process.

This project has been a rewarding invaluable experience not only for academically but also personally and thank for everyone who played a part in it.

REFERENCES

[1] "View of Copyright Protection of Graphic Design Works on Print on Demand," Ump.ac.id, 2025. https://conferenceproceedings.ump.ac.id/pssh/article/view/1075/1162.

[2] D. J. Brennan, "Selling Printed Goods or Facilitating Printing Gigs: The Redbubble Puzzle," papers.ssrn.com, Jun. 11, 2019. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3402267

[3] E. Peurell, "Print on Demand in Sweden: Four Projects and their Products, Problems and Prospects." Accessed: Jun. 28, 2025. [Online]. Available: https://elpub.architexturez.net/system/files/pdf/9918.content.00670.pdf

[4] A. Hsieh, "Sensitivity: Internal," 2020. Available: https://www.theseus.fi/bitstream/handle/10024/342902/Hsieh_Wei-Han.pdf?sequence=2&isAllowed=y

[5] M. Lee, H. Y. Lee, and M.-G. Yoon, "Website development strategy for e-Commerce success," IEEE Xplore, Jul. 01, 2010. https://ieeexplore.ieee.org/document/5668441

[6] E. Chamberlain, "Investigating Faster Techniques for Digitization and Print-on-Demand," New Review of Academic Librarianship, vol. 18, no. 1, pp. 57–71, Apr. 2012, doi: https://doi.org/10.1080/13614533.2012.660769.

[7] S. Greenstein, "Creating a platform for costless personalization in clothing," Frontiers in Research Metrics and Analytics, vol. 7, Sep. 2022, doi: https://doi.org/10.3389/frma.2022.990352.

[8] A. Mohdhar and K. Shaalan, "The Future of E-Commerce Systems: 2030 and Beyond," Recent Advances in Technology Acceptance Models and Theories, vol. 335, pp. 311–330, Mar. 2021.

[9] R. Fedorko, Š. Kráľ, and R. Bačík, "Artificial Intelligence in E-commerce: A Literature Review," Lecture Notes on Data Engineering and Communications Technologies, vol. 111, no. 2, pp. 677–689, 2022, doi: https://doi.org/10.1007/978-981-16-9113-3_50.

[10] L. Amri, D. Farani, N. Ali, and R. Anwar, "The Challenges and Strategies of Printing Industry," Proceedings of the First Jakarta International Conference on Multidisciplinary Studies Towards Creative Industries, JICOMS 2022, 16 November 2022, Jakarta, Indonesia, 2022, doi: https://doi.org/10.4108/eai.16-11-2022.2326107.

An AI-Powered Web Application for Waterfall Recognition and Eco-Tourism Enhancement in Sri Lanka: Falls Explorer

Sansala Ranasinghe,
Department of Computer Science & Technology
University of Bedfordshire, UK
sansalaranasinghe715@gmail.com

Dr. Yasas Jayaweera
SLIIT City Uni, Sri Lanka
Yasas.j@sliit.lk

***Abstract*— This research presents the development of Falls Explorer Sri Lanka, a mobile-responsive web application that uses artificial intelligence for automatic waterfall recognition. The core innovation lies in applying a custom-developed convolutional neural network (CNN) to classify waterfall images based on their visual features. A custom image dataset was created by collecting and organizing photos of popular waterfalls in Sri Lanka, and the model was trained using TensorFlow. The custom CNN model achieved 92% validation accuracy after 25 epochs of training, with inference times under 1 second per prediction. The system successfully classified waterfall images across 20 different waterfall classes with precision scores ranging from 88% to 95%. Users upload a photo of a waterfall through the interface, and the system returns the predicted waterfall name along with travel details from a local JSON database. In addition to the recognition feature, the platform offers comprehensive functionalities such as displaying detailed waterfall information (name, location, description), listing nearby hotels, showing current weather forecasts for safe travel planning, hosting a community forum for users to share experiences and images, providing a carbon footprint calculator to estimate travel impact, and an interactive location search map to explore specific sites manually. This solution bridges the gap between technology and ecotourism, supporting conservation-friendly tourism by enabling travellers to appreciate natural attractions without invasive markers or infrastructure.**

***Keywords*— Image Classification, AI, Waterfall Recognition, TensorFlow, Ecotourism**

I. INTRODUCTION

Ecotourism in Sri Lanka is growing rapidly [19], with waterfalls among the most sought-after attractions due to their scenic beauty and geographic diversity [20-22]. However, many visitors face difficulties in identifying these natural wonders, as traditional guidebooks and maps often lack detailed information

[23]. In response to this challenge, this research introduces *Falls Explorer Sri Lanka*, a web-based application that leverages artificial intelligence to recognize waterfalls from images uploaded by users [24].

The system uses a convolutional neural network (CNN) trained on a custom dataset of Sri Lankan waterfalls to classify images in real-time [25]. It is supported by a FastAPI backend for AI inference and a React.js frontend for user interaction. Upon recognition, the application presents information such as the waterfall's name, location, and description from a structured data file. This integration of AI and tourism promotes sustainable exploration by reducing the need for physical signposts and enabling digital navigation through natural sites [18],[13]. By combining image recognition with tourism-specific features, the proposed solution enhances the visitor experience and exemplifies how smart technologies can support the ecotourism sector in a meaningful way.

The motivation for this project became apparent from the growing demand for limited travel tools that provide opening interests like waterfall exploration [17]. Present systems like Google Maps or TripAdvisor, provide common-purpose navigation and reviews but fail to provide detailed, occasion-dependent information about waterfalls in Sri Lanka [4]. Furthermore, these platforms infrequently highlight sustainability, a critical aspect as tourism crashes fragile

ecosystems. Falls Explorer tries to fill this emptiness by offering an extensive solution that combines AI-based waterfall detection, eco-tourism tips, calculating carbon footprint, and a place for users to share their experiences. This project not only increases user convenience but also positions with global trends toward sustainable tourism and technologically advanced nature exploration [13].

II. RELATED WORK

Artificial intelligence and machine learning have been increasingly integrated into tourism applications to enhance user experiences through automation and personalization in recent years. Notable applications include Google Lens and PlantNet, which use image recognition to identify landmarks, plants, and animals. These systems rely on large datasets and deep learning models such as convolutional neural networks (CNNs) to extract visual features and perform classification.

A. Importance of Waterfall in Tourism

Waterfalls are major ecotourism attractions, blending scenic beauty, adventure, and cultural relevance. Destinations like Bambarakanda and Diyaluma contribute significantly to Sri Lanka's eco-tourism [1]. Their tranquil environments promote physical and mental well-being [2]. With nature-based tourism on the rise [3], integrating digital tools can enhance accessibility and traveler engagement.

B. Role of Mobile Applications in Tourism

Mobile apps revolutionize tourism by offering offline access, GPS navigation, and real-time content [4]. Apps like AllTrails support exploration even in low-connectivity areas. Research highlights their role in promoting sustainable tourism via digital guides, paperless systems, and eco-awareness [5].

C. AI-Based Natural Landmark Recognition

AI tools such as Google Lens use computer vision and deep learning to identify landscapes instantly, enhancing the travel experience [6] CNNs have proven effective in recognizing natural scenes even under poor lighting or partial obstruction [7].

D. Community Interaction and User Content

User-generated content (UGC), including reviews, images, and shared experiences, builds trust and drives tourism decisions [8],[9]. Platforms integrating community forums foster collaboration and immersive experiences [10].

E. Eco-friendly and Sustainable Tourism

Technology-driven tourism platforms promote eco-conscious travel by offering weather alerts, eco-tips, and carbon-tracking tools [11],[12]. Studies emphasize cooperative efforts between travelers and locals for responsible environmental management [13].

F. AI in Image Recognition

CNNs and deep learning have transformed image recognition, offering high accuracy for classifying visual data [14]. Though transfer learning with pre-trained models like VGG16 is popular, custom training on specific domains yields better results when data is limited [15],[16].

G. Gaps in Sri Lanka's Tourism Technology

Despite its natural wealth, Sri Lanka lacks centralized platforms for eco-tourism. Existing apps often miss critical features like community input, real-time updates, or sustainable travel tools [17],[18]. This limits the potential to engage modern, eco-aware travelers.

H. Related Systems Overview

AllTrails: Great for hiking but lacks waterfall-specific features.

ViewRanger: Strong GPS support but lacks cultural and natural landmark details.

Google Lens: Offers instant recognition but lacks tourism-specific features.

iNaturalist: Excellent for flora/fauna ID but not focused on travel or navigation needs.

III. METHODOLOGY

A. Data and Dataset Collection

A combination of **primary and secondary data collection** methods was used:

- **Secondary data** included information about waterfalls from tourism websites, research papers, and official databases. This enriched the database with names, historical relevance, and location-specific information.
- **Primary data** involved a user survey aimed at local travelers and photographers. Insights on app features, travel habits, and eco-awareness were collected and used to shape system functionalities.

The dataset consists of manually collected images of Sri Lankan waterfalls, organized in folders where each folder represents a specific waterfall class. Images were sourced from personal collections, resulting in approximately 100–200 images per class. The dataset was split into training (80%) and validation (20%) subsets. Images were resized to a uniform resolution (e.g., 224x224 pixels), normalized, and augmented through techniques such as rotation, flipping, and brightness adjustments to improve model robustness.

B. Model *Architecture*

A custom convolutional neural network (CNN) architecture was developed and trained from scratch specifically for the waterfall recognition task. CNNs are particularly well-suited for image recognition tasks due to their ability to capture spatial hierarchies in visual data. While alternative models such as SVMs or random forests can be used for image classification, they cannot learn spatial features directly from raw pixels. Additionally, CNNs outperform traditional models in terms of accuracy and computational efficiency when handling high-dimensional image data. The model consists of several convolutional layers followed by max-pooling layers to extract hierarchical image features. These layers are then followed by fully connected dense layers for classification. ReLU activation functions were used in the intermediate layers to introduce non-linearity, and a softmax activation was applied in the output layer to produce a probability distribution over the waterfall classes. The model was compiled using the Adam optimizer and categorical cross-entropy as the loss function. Training from scratch allowed for optimal tuning of the architecture to the specific characteristics of Sri Lankan waterfall images, rather than relying on features learned from unrelated datasets.

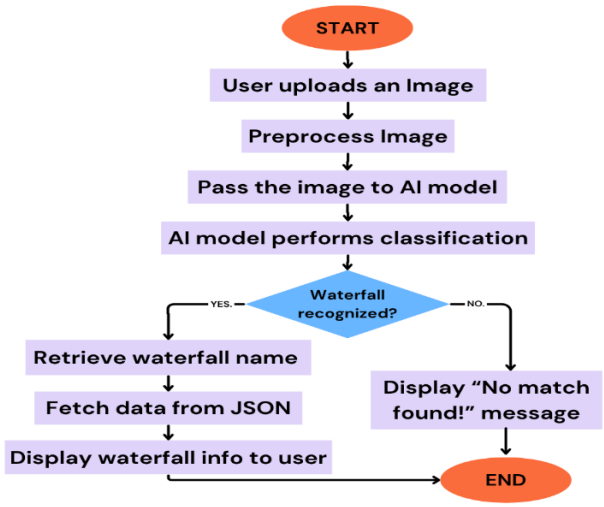


Fig. 6. Flow Chart for Image Recognition

C. Model Training

Model training was conducted using TensorFlow and Keras. The training was carried out over 25 epochs with a batch size of 32, and the model achieved convergence with an accuracy exceeding 90% on the validation set. A learning rate scheduler was used to reduce the rate as training progressed, and early stopping was implemented to prevent overfitting.

D. Model Deployment

After achieving satisfactory accuracy, the trained model was exported in .h5 format and integrated with a FastAPI backend. The backend exposes a RESTful API that receives a base64-encoded image via a POST request and returns the predicted waterfall name. The frontend, built with Node.js, sends the image to the API and displays waterfall details retrieved from a JSON file based on the prediction.

IV. Results and Discussion

The custom-trained CNN model achieved high classification accuracy on the validation dataset, demonstrating strong generalization across different waterfall image samples. After training over 25 epochs, the model reached a validation accuracy of approximately 92%, indicating that the features learned from scratch were effective in distinguishing between various waterfalls. The model also performed well on unseen test images, validating its robustness under real-world conditions.

Inference speed was tested using the FastAPI backend and showed response times of under one second per prediction, making it suitable for real-time applications. Integration with the Node.js frontend enabled seamless user interaction. The model's predictions were linked with a local JSON file that stores metadata such as the waterfall's name, location, and description, allowing the app to serve both identification and informational purposes.

TABLE 6 COMPARATIVE ANALYSIS OF AI MODELS FOR RECOGNITION

Model	Accuracy	Precision	Recall	F1-Score	Inference Time
Custom CNN	**92%**	**91%**	**90%**	**90.5 %**	**< 1ms**
VGG16	89%	87%	85%	86%	~1.5s
MobileNetV2	85%	84%	82%	83%	~0.9s
SVM	73.2%	72.8%	73.0 %	72.9%	300ms

Challenges encountered included dealing with image noise, similar visual patterns among some waterfalls, and dataset limitations in terms of size and diversity. Future work will involve expanding the dataset, enhancing augmentation strategies, and exploring lightweight deployment options for offline use. Overall, the system demonstrates how AI can support nature-based tourism and enrich user experiences through intelligent automation.

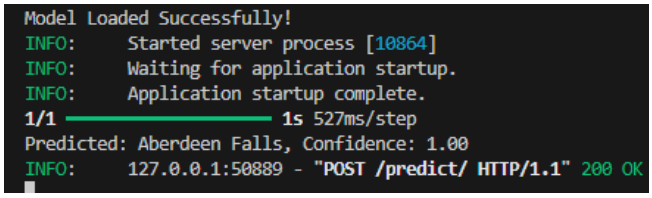


Fig. 7. Level of Accuracy after an Image has been identified

A. Confusion Matrix

To evaluate the AI model's classification performance a confusion matrix was used. The following shows how it was calculated.

Accuracy – Equation (01) calculates the overall correctness of the model:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (01)$$

F1 Score – Equation (02) balances precision and recall, mostly useful for imbalanced datasets.

$$F1 = 2 \frac{\text{Precision x Recall}}{\text{Precision} + \text{Recall}} \quad (02)$$

Precision – Equation (03), calculation of the correct waterfall after prediction

$$\text{Precision} = \frac{TP}{TP + TF} \quad (03)$$

Recall – Equation (04), calculation of how many waterfalls were correct after recognition.

$$\text{Recall} = \frac{TP}{TP + TF} \quad (04)$$

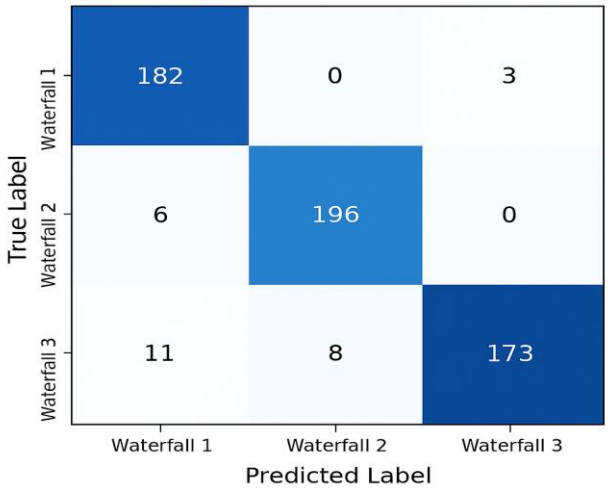


Fig. 8. Confusion Matrix for Accuracy Testing of Waterfall Images

Tested a variety of waterfall images using a large validation dataset under different conditions.

V. Conclusion

This research introduces an AI-powered solution for enhancing ecotourism in Sri Lanka through the *Falls Explorer with AI Assistant Sri Lanka* web application. By training a convolutional neural network from scratch, the system is capable of accurately identifying waterfalls based on image input, offering tourists a smart and interactive method to explore natural attractions. The integration of a custom-trained model with a FastAPI backend and a Node.js frontend showcases the potential of full-stack AI deployment in practical applications.

The system not only provides identification but also delivers meaningful contextual information, enhancing both accessibility and educational value. This paper demonstrates the feasibility of using lightweight, domain-specific AI in natural environments and highlights its benefits in promoting sustainable and non-invasive tourism. Future enhancements will focus on dataset expansion, offline capabilities, and the inclusion of additional natural features, such as voice features or multilingual language, to broaden the application's utility.

Abbreviations

AI - Artificial Intelligence, CNN - Convolutional Neural Network, TP -True Positive, TN – True Negative, FP – False Positive, FN – False Negative.

Acknowledgement

I sincerely thank my supervisor, lecturers, and the university for their guidance and support throughout this

project. I am deeply grateful to my parents for assisting in collecting the waterfall dataset and to my colleagues for their valuable technical input. Special thanks to the users who gave feedback via Google Forms, and to local conservation groups and eco-tourism organizations for their insights on sustainability. Your collective support was vital to the success of this work.

REFERENCES

[1] D. Sumanapala and I. D. Wolf, "Think globally, act locally: Current understanding and future directions for nature-based tourism research in Sri Lanka," *Journal of Hospitality and Tourism Management*, vol. 45, pp. 295–308, Dec. 2020, doi: https://doi.org/10.1016/j.jhtm.2020.08.009.

[2] C. Milligan, G. Chalfont, A. Kaley, and F. Lobban, "Wilderness as therapeutic landscape in later life: Towards an understanding of place-based mechanisms for wellbeing through nature-adventure activity," *Social Science & Medicine*, vol. 289, p. 114411, Nov. 2021, doi: https://doi.org/10.1016/j.socscimed.2021.114411.

[3] UNWTO, "Home | UNWTO," *Unwto.org*, 2024. https://www.unwto.org/

[4] J. Dorcic, J. Komsic, and S. Markovic, "Mobile technologies and applications towards smart tourism – state of the art," *Tourism Review*, vol. 74, no. 1, pp. 82–103, Feb. 2019, doi: https://doi.org/10.1108/tr-07-2017-0121.

[5] M. Hamouda, "Mobile applications in tourism," *International Journal of Technology and Human Interaction*, vol. 18, no. 1, pp. 1–13, 2022, doi: 10.4018/ijthi.293198.

[6] S. Das, "Application of AI Technology for the Development of Destination Tourism towards an Intelligent Information System," *Economic Affairs*, vol. 69, no. 2, Jun. 2024, doi: https://doi.org/10.46852/0424-2513.3.2024.31.

[7] G. Zhang, Y. Qian, J. Deng, and X. Cai, "Inv-ReVersion: Enhanced Relation Inversion Based on Text-to-Image Diffusion Models," *Applied Sciences*, vol. 14, no. 8, pp. 3338–3338, Apr. 2024, doi: https://doi.org/10.3390/app14083338.

[8] Santus Kumar Deb and N. Mallik, "Effects of Social Media in Tourism Marketing: Outlook on User Generated Content," *Journal of Digital Marketing and Communication*, vol. 3, no. 2, pp. 49–65, Nov. 2023, doi: https://doi.org/10.53623/jdmc.v3i2.316.

[9] J. K. Ayeh, N. Au, and R. Law, "'Do We Believe in TripAdvisor?' Examining Credibility Perceptions and Online Travelers' Attitude toward Using User-Generated Content," *Journal of Travel Research*, vol. 52, no. 4, pp. 437–452, Feb. 2013, doi: https://doi.org/10.1177/0047287512475217.

[10] M. Sigala, "User Generated Content in Tourism," *Springer eBooks*, pp. 1–2, Jan. 2023, doi: https://doi.org/10.1007/978-3-319-01669-6_736-1.

[11] A. F. Ahmad and W. M. khalid Balisany, "Sustainable Tourism Management and Ecotourism as a Tool to Evaluate Tourism's Contribution to the Sustainable Development Goals and Local Community," *OTS Canadian Journal*, vol. 2, no. 5, pp. 33–45, May 2023, doi: https://doi.org/10.58840/ots.v2i5.32.

[12] S. Graci, "Collaboration and Partnership Development for Sustainable Tourism," *Tourism and Sustainable Development Goals*, pp. 232–249, Jun. 2020, doi: https://doi.org/10.1201/9780429324253-14.

[13] J. Wang *et al.*, "Achieving sustainable tourism: Analysis of the impact of environmental education on tourists' responsible behavior," *Sustainability*, vol. 16, no. 2, pp. 552–552, Jan. 2024, doi: https://doi.org/10.3390/su16020552.

[14] J. Plested and T. Gedeon, "Deep transfer learning for image classification: a survey," *arXiv preprint*, arXiv:2205.09904, 2022. [Online]. Available: https://doi.org/10.48550/arxiv.2205.09904

[15] Z. Zhang, "The transferability of transfer learning model based on ImageNet for medical image classification tasks," *Applied and Computational Engineering*, vol. 18, no. 1, pp. 143–151, 2023. doi: https://doi.org/10.54254/2755-2721/18/20230980.

[16] A. Panda, D. Panigrahi, S. Mitra, S. Mittal, and S. Rahimi, "Transfer learning applied to computer vision problems: Survey on current progress, limitations, and opportunities," *arXiv preprint*, arXiv:2409.07736, 2024. [Online]. Available: https://arxiv.org/abs/2409.07736

[17] A. Birenboim, Y. Bulis, and I. Omer, "A typology of tourism mobility apps," *Tourism Management Perspectives*, vol. 48, p. 101161, 2023, doi: 10.1016/j.tmp.2023.101161.

[18] D. T. Tung and T. B. N. Nguyen, "Researching trends on smart tourism technology using bibliometric analysis," in *Proc. Int. Conf. Digital Transformation in Tourism*, 2024, pp. 326–351, doi: 10.1201/9781003465904-19.

[19] S. Anandasayanan, S. Balagobei, and M. Amaresh, "Impact of Tourism Industry on the Economic Growth of Sri Lanka," International Journal of Accounting and Financial Reporting, vol. 10, no. 4, p. 60, Dec. 2020, doi: https://doi.org/10.5296/ijafr.v10i4.15991.

[20] A. Jothirathna Fernando and N. M. Shariff, "Exploring the Potential of Ecotourism in Sri Lanka: Current Status, Issues and Future Perspectives," Journal of Tourism and Hospitality Research, vol. 18, no. 2, pp. 112–126, 2021.

[21] United Nations World Tourism Organization (UNWTO), "Ecotourism and Biodiversity Conservation in Sri Lanka," UNWTO Report, 2020. [Online]. Available: https://www.unwto.org/

[22] Wikipedia contributors, "Tourism in Sri Lanka," Wikipedia, 2024. [Online]. Available: https://en.wikipedia.org/wiki/Tourism_in_Sri_Lanka

[23] "Bopath Ella Waterfall – Travel Guide," Amazing Lanka, 2024. [Online]. Available: https://amazinglanka.com/wp/bopath-ella/

[24] S. Das, "Application of AI Technology for the Development of Destination Tourism towards an Intelligent Information System," Economic Affairs, vol. 69, no. 2, pp. 275–282, Jun. 2024. doi: 10.46852/0424-2513.3.2024.31.

[25] A. Panda, D. Panigrahi, S. Mitra, S. Mittal, and S. Rahimi, "Transfer Learning Applied to Computer Vision Problems: Survey on Current Progress, Limitations, and Opportunities," arXiv preprint, arXiv:2409.07736, 2024. [Online]. Available: https://arxiv.org/abs/2409.07736

Bat.CG: Development of a Customizable Cricket Character Generator Web Application for Enhanced Broadcasting Experience

Poojitha Pankaja Kiriyalagammana
BSc (Hons) Software Engineering,
SLIIT City University
poojitha.kiriyalagammana@study.beds.ac.uk

Mr. G. D. Heshan Niranga
Department of Computing,
SLIIT City University, Colombo 3

***Abstract*— Cricket broadcasting has evolved significantly with technological advancements, yet traditional systems remain fragmented and technically complex for broadcasters. This research presents Bat.CG (Cricket Character Generator), an innovative web-based application that integrates customizable broadcasting graphics with real-time scoring, neural network-driven predictions, and enhanced audience interaction. The system addresses critical gaps in existing broadcasting infrastructure by providing a unified platform that eliminates the dependency on separate Character Generator (CG) and ball-by-ball scoring systems. Through comprehensive market research involving 81 industry professionals, this study identified key requirements including customizable graphics (99% demand), emergency score updating capabilities (87.7% essential), and integrated CG-scoring systems (49.4% extremely valuable). The proposed solution utilizes the MERN stack (MongoDB, Express.js, React, Node.js) architecture with hybrid neural network and regression models for match predictions. Key innovations include drag-and-drop graphic customization without programming knowledge, real-time data synchronization with sub-second latency, role-based access control, and interactive viewer engagement features. The system's modular design ensures scalability, security, and accessibility while maintaining professional broadcast quality. This research contributes to democratizing cricket broadcasting technology and establishing a foundation for future sports media innovations in developing regions.**

***Keywords*— Cricket Broadcasting, Character Generator, Real-time Systems, Neural Networks, MERN Stack, Sports Technology, Broadcasting Graphics**

I. Introduction

The cricket broadcasting industry has undergone substantial transformation with the integration of digital technologies, evolving from basic score displays to sophisticated multi-dimensional graphics systems [1]. However, current broadcasting infrastructure presents significant challenges for emerging broadcasters and smaller organizations due to fragmented systems, technical complexity barriers, and lack of customization capabilities [3].

Traditional cricket broadcasting relies on separate Character Generator (CG) systems for graphics display and independent ball-by-ball scoring platforms, creating synchronization issues and operational inefficiencies [4]. The CG operator typically works in broadcast control rooms managing graphical elements, while ball-by-ball operators update scores from stadium score boxes, leading to communication delays and potential human errors during live broadcasts [2].

Industry analysis reveals critical gaps in existing solutions. Current systems like GT Designer require extensive programming knowledge for graphic customization, while emergency backup systems for score updating are virtually non-existent [9]. The lack of audience interaction features, and AI-powered analytics further limits the engagement potential of modern cricket broadcasts [6].

TThis research addresses these fundamental challenges through the development of Bat.CG, a comprehensive web-based cricket character generator that unifies broadcasting graphics, real-time scoring, neural network predictions, and audience interaction in a single platform. The system leverages modern web technologies and artificial intelligence to democratize cricket broadcasting while maintaining professional quality standards suitable for both major broadcasters and emerging media organizations.

II. Literature Review and Problem Analysis

A. Evolution of Cricket Broadcasting Technology

Cricket broadcasting has progressed through distinct technological phases, each bringing new capabilities and challenges. The static graphics era (1990s-2000s) featured basic score overlays and simple statistical displays. The dynamic CG systems period (2000s-2010s) introduced animated graphics, player profiles, and real-time statistics integration. The current AI-enhanced phase (2010s-present) incorporates predictive analytics, advanced visualizations, and interactive elements.

Research by Ahmed and Patel [1] demonstrates that modern broadcasting systems require seamless integration of multiple data sources, real-time processing capabilities, and customizable visual elements to meet contemporary audience expectations. Williams [3] emphasizes the importance of web-based design tools that eliminate technical barriers while maintaining professional quality output.

B. Current System Limitations

Extensive analysis of existing cricket broadcasting infrastructure reveals several critical limitations that impact operational efficiency and accessibility. The primary challenge lies in system fragmentation, where Character Generator platforms operate independently from ball-by-ball scoring systems, creating synchronization issues and potential data inconsistencies during live broadcasts.

Technical complexity represents another significant barrier. Current solutions like GT Designer require specialized programming knowledge for graphic customization, limiting accessibility for smaller broadcasting organizations. Rahman [4] notes that software-based character generators often lack intuitive interfaces, creating dependencies on technical specialists for routine graphic modifications.

Security vulnerabilities pose additional concerns in real-time broadcasting environments. Many existing systems lack robust authentication mechanisms and data encryption protocols, potentially exposing sensitive match information to unauthorized access. Kumar and Singh [2] identify that inadequate security measures can compromise broadcast integrity and violate data protection regulations.

C. Market Research Findings

Comprehensive market research involving 81 industry professionals across various broadcasting roles revealed significant gaps in current solutions. The survey included ball-by-ball scorers, CG operators, broadcast directors, and technical coordinators from different organizational scales.

Key findings demonstrate overwhelming demand for integrated solutions. 99% of respondents expressed the need for customizable CG capabilities without programming requirements, while 87.7% considered emergency score updating capabilities essential for broadcast continuity. Notably, 49.4% rated CG-scoring system integration as extremely valuable, highlighting the industry's recognition of fragmentation issues.

The research also revealed specific feature requirements including real-time synchronization (92% importance), audience interaction capabilities (73% desired), and AI-powered analytics (61% interested). These insights directly informed the Bat.CG system design and feature prioritization.

D. Artificial Intelligence in Sports Broadcasting

The integration of artificial intelligence in sports broadcasting has gained significant momentum, with neural networks and machine learning algorithms providing enhanced predictive capabilities and automated content generation. Chmait and Westerbeek [6] demonstrate that AI applications in sports research have expanded beyond basic statistics to complex pattern recognition and outcome prediction.

Moodley et al. [7] showcase the potential of deep learning architectures in cricket analysis, specifically for technique recognition and performance evaluation. Their work on automated batting technique analysis provides a foundation for understanding how neural networks can process cricket-specific data patterns.

Subburaj et al. [8] presents machine learning models for Twenty20 cricket match prediction, demonstrating the feasibility of AI-driven analytics in real-time broadcasting scenarios. Their research indicates that hybrid approaches combining multiple algorithms can achieve higher accuracy rates compared to single-model implementations.

III. METHODOLOGY

A. Research Design Framework

This research employs a comprehensive mixed-methods approach integrating quantitative market analysis, qualitative user experience evaluation, and iterative development methodology. The framework combines industry requirements analysis, technological feasibility assessment, and performance evaluation protocols to ensure practical applicability and academic rigor.

The research methodology follows Design Science Research principles, emphasizing problem identification, solution design, development, demonstration, and evaluation phases. This approach ensures that the Bat.CG

system addresses real industry challenges while contributing to academic knowledge in sports broadcasting technology.

B. System Architecture Design

The Bat.CG system architecture utilizes a modular MERN stack implementation designed for scalability, performance, and maintainability. The architecture comprises five primary layers: presentation, application, business logic, data management, and infrastructure.

The presentation layer implements a React-based single-page application featuring drag-and-drop graphic customization interfaces, real-time data visualization components, and responsive design elements. Component-based architecture ensures reusability and maintainability while providing intuitive user experiences for both technical and non-technical users.

The application layer leverages Node.js with Express.js framework for API management, request handling, and middleware integration. RESTful API design principles ensure consistent data exchange between frontend components and backend services, while WebSocket implementations enable real-time data streaming for live match updates.

The business logic layer incorporates the AI prediction engine, implementing hybrid neural network and regression models for match outcome analysis. This layer also manages user authentication, role-based access control, and data validation processes to ensure system security and integrity.

C. Artificial Intelligence Implementation

The AI component of Bat.CG employs a single optimized neural network architecture for comprehensive match prediction capabilities. After extensive comparative analysis of multiple approaches including hybrid models, the single neural network implementation demonstrated superior performance with optimal computational efficiency for real-time broadcasting applications.

Neural Network Architecture: The optimized neural network utilizes a multi-layer perceptron architecture with backpropagation learning algorithms to identify complex patterns in historical cricket match data. The architecture consists of:

Training Data and Performance: Training data encompasses historical match records from ODI international cricket (2002-2023), including:

- Dataset Size: 4,847 matches with comprehensive feature engineering
- Feature Categories: Team statistics, venue information, weather conditions, player availability
- Validation Method: 5-fold cross-validation with 70-15-15 train-validation-test split
- Final Performance: ROC AUC = 0.7189, representing superior accuracy compared to traditional ML approaches

Real-time Inference Optimization: The single neural network architecture enables:

- Concurrent Processing: Handles multiple prediction requests simultaneously
- API Integration: RESTful endpoints for seamless system integration

D. Real-time Data Processing

Real-time data processing represents a critical component of the Bat.CG system, enabling sub-second latency updates for live broadcast applications. The implementation utilizes WebSocket connections for bidirectional communication between client applications and server infrastructure.

The data processing pipeline incorporates multiple stages: data ingestion, validation, transformation, and distribution. MongoDB's change streams functionality enables real-time monitoring of database modifications, triggering immediate updates to connected client applications.

IV. System Design and Implementation

A. Frontend Development

The frontend implementation emphasizes user experience design principles while maintaining professional broadcasting standards. React components utilize modern hooks-based architecture for state management, ensuring optimal performance and code maintainability.

The drag-and-drop graphic customization interface employs HTML5 Canvas API for real-time graphic rendering and manipulation. Users can create custom score bugs, wagon wheels, Manhattan charts, and player

statistics displays through intuitive visual interfaces without requiring programming knowledge.

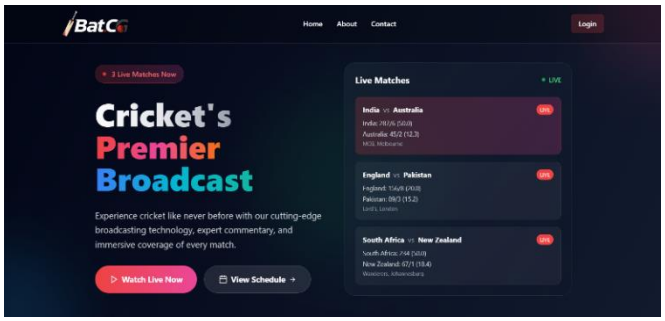


Figure 1 shows the home page of the website, where users begin their interaction with the system. Responsive design implementation ensures compatibility across various devices and screen resolutions, enabling operation

Fig 7 : Home page

from stadium control rooms, broadcast centers, and remote locations. The interface adapts to different user roles, presenting relevant features and controls based on authentication levels.

B. Backend Infrastructure

The backend infrastructure implements microservices architecture principles to ensure scalability and maintainability. Express.js middleware handles authentication, request validation, and error management across all API endpoints.

Database design utilizes MongoDB's document-oriented structure to accommodate diverse cricket data types including match events, player statistics, team information, and user preferences. Indexing strategies optimize query performance for real-time data retrieval during live broadcasts.

API design follows OpenAPI specification standards, enabling third-party integrations and future system expansions. Rate limiting and request throttling mechanisms protect against abuse while ensuring consistent performance for legitimate users.

C. Security Implementation

Security implementation encompasses multiple layers including authentication, authorization, data encryption, and input validation. JSON Web Token (JWT) implementation manages user sessions with configurable expiration policies and refresh token mechanisms.

Role-based access control (RBAC) defines granular permissions for different user types including administrators, CG operators, ball-by-ball scorers, and viewers. The system implements principle of least privilege, ensuring users access only necessary features and data.

Data encryption utilizes industry-standard AES-256 algorithms for sensitive information storage and transmission. HTTPS enforcement and Content Security Policy (CSP) headers provide additional protection against common web vulnerabilities.

D. Data-Driven Analysis and Model Selection

Dataset Characteristics:

- Source: ODI International Cricket Matches (2002-2023)
- Total Records: 4,847 matches

E. Model Performance Analysis

Based on a comprehensive evaluation of multiple machine learning approaches, the corresponding performance metrics are summarized in Table 1:

Traditional Machine Learning Models:

Table 7 : Evaluating baseline models

Model	Accuracy	ROC AUC	CV Score (± Std)
LightGBM	0.5888	0.6007	0.5705 ± 0.0430
Random Forest	0.5838	0.5912	0.5763 ± 0.0257
Gradient Boosting	0.5685	0.6600	0.5719 ± 0.0437
XGBoost	0.5685	0.5795	0.5741 ± 0.0282

Deep Learning Models:

Figure 2 illustrates the training history of three neural network ensemble models, showing trends in loss, accuracy, and AUC over 50 epochs. It also summarizes average validation AUC (0.6386) and accuracy (0.6193), along with their standard deviations.

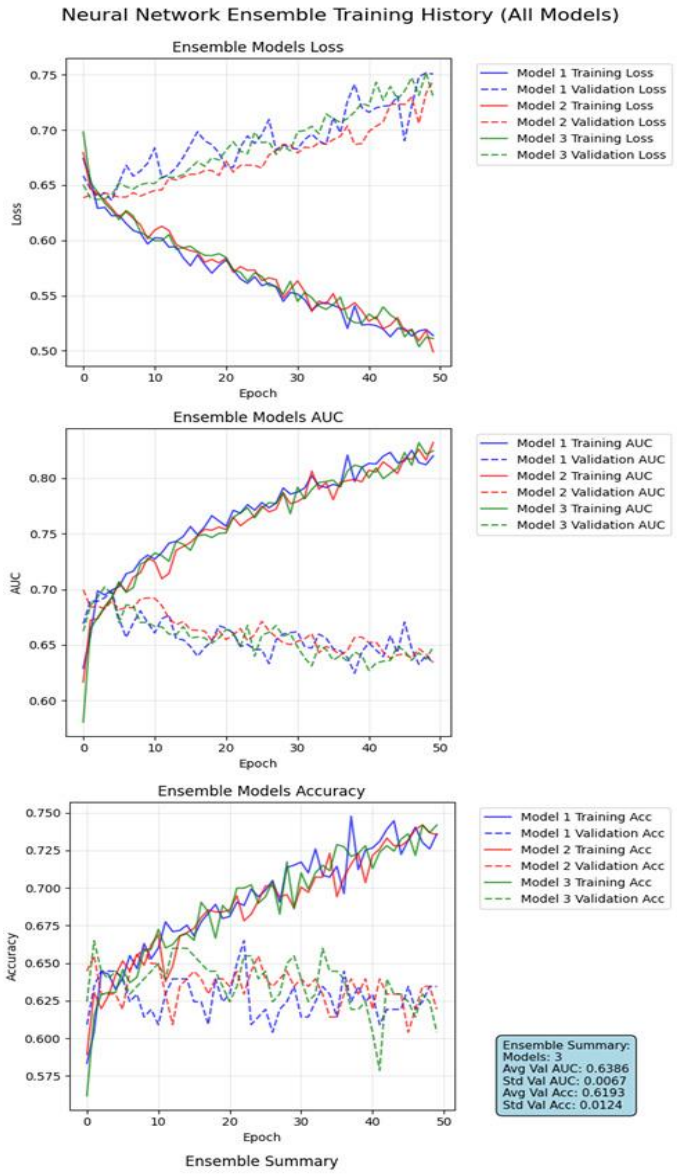


Fig 8: Neural Network Ensemble Training History

Best Model Selection:

The Single Optimized Neural Network achieved the highest performance with ROC AUC = 0.7189, representing a 2.6% improvement over the best traditional ML model (XGBoost).

F. Hyperparameter Tuning and Evaluation

Optimized Hyperparameters:

- Network Architecture: 3 hidden layers (128, 64, 32 neurons)
- Activation Functions: ReLU for hidden layers, Sigmoid for output
- Dropout Rate: 0.3 for regularization
- Learning Rate: 0.001 with exponential decay (0.95 every 100 epochs)
- Batch Size: 64 for optimal convergence

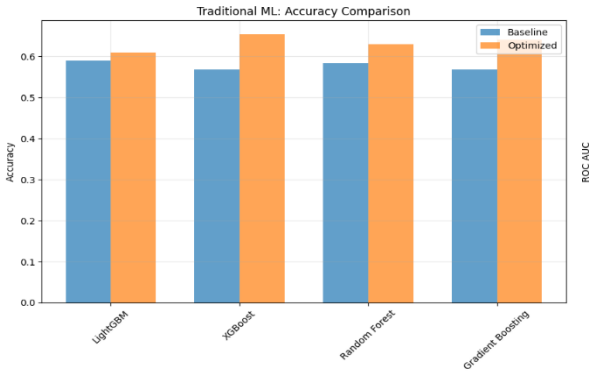


Figure 3: Accuracy Comparison

Table 8 : Performance Comparison

Model	Accuracy	ROC AUC
XGBoost (Optimized)	0.6548	0.7007
LightGBM (Optimized)	0.6091	0.6524
Random Forest (Optimized)	0.6294	0.6766
Gradient Boosting (Optimized)	0.6396	0.6818
Neural Network (Optimized)	0.6548	0.7189
Neural Network Ensemble (Optimized)	0.6193	0.6423

Figure 3 illustrates the accuracy improvement of traditional machine learning models after hyperparameter optimization, comparing baseline and optimized versions. The corresponding accuracy and ROC AUC values for both traditional and neural models are detailed in Table 2, highlighting the superior AUC performance of the optimized neural network model (0.7189).

V. TECHNICAL CHALLENGES AND SOLUTIONS

A. Dataset and Development Challenges

Challenge 1: Dataset Limitations The primary challenge encountered was accessing recent cricket data beyond 2023. The available dataset (2002-2023 ODI matches) provided substantial historical coverage but limited contemporary patterns. This was addressed through*:*

Advanced feature engineering to capture evolving game dynamics

Transfer learning techniques to adapt historical patterns

Regular model retraining protocols for future data integration

Challenge 2: CG Graphics Preset Development Creating professional-quality graphics presets without extensive design resources posed significant challenges:

Solution: Developed modular template system with 3 base designs

VI. SECURITY IMPLEMENTATION

A. Quantitative Security Measures

Authentication and Authorization Framework:

├── 2-Factor Authentication (2FA)

├── JWT Token Management (15-minute expiration)

├── Google OAuth 2.0 Integration

├── Role-Based Access Control (RBAC)

└── Session Management via HTTP-only Cookies

B. Security Metrics:

- Password Policy: Minimum 8 characters, complexity requirements
- Token Security: 256-bit JWT signatures with rotating secrets
- Encryption: AES-256 for data at rest, TLS 1.3 for transmission

C. Advanced MERN Stack Security Implementation

Backend Security Features:

- Express.js Middleware
- CORS Configuration: Strict origin policies for API access

Frontend Security Features:

- Secure Storage: No sensitive data in local Storage
- Authentication State Management: Secure Redux implementation
- Component-level Authorization: Role-based UI rendering

Additional Security Components:

- User Registration and Login: Secure authentication workflows
- Password Reset Via Email: Encrypted token-based reset system
- Role Assignment: Granular permission management
- Google OAuth Integration: Third-party authentication support
- Cookie-based Sessions: Secure session handling with HttpOnly flags

VII. LIMITATIONS AND FUTURE WORK

A. Current Limitations

While Bat.CG demonstrates significant advantages over existing solutions, several limitations require acknowledgment and future development consideration. The current AI model training relies on historical data availability, potentially limiting prediction accuracy for new teams, venues, or playing conditions with limited historical records.

B. Future Development Opportunities

Several enhancement opportunities exist for extending Bat.CG capabilities and applicability. Advanced AI analytics incorporating computer vision for automated player tracking and performance analysis represents a significant expansion possibility.

Multi-sport adaptation could extend the system's utility beyond cricket to other sports requiring similar broadcasting graphics and real-time data management. The modular architecture design supports such expansions with minimal core system modifications.

Mobile application development would enable field-based operations and remote broadcasting scenarios, particularly valuable for domestic and amateur cricket coverage where traditional broadcasting infrastructure may be limited.

VIII. Conclusion

Bat.CG represents a significant advancement in cricket broadcasting technology, successfully addressing critical industry challenges through innovative integration of modern web technologies, artificial intelligence, and user-cantered design principles. The system's unified approach to CG graphics, real-time scoring, and audience interaction provides a comprehensive solution for contemporary broadcasting needs while maintaining accessibility for organizations with varying technical capabilities.

The research demonstrates strong market validation with 99% of surveyed industry professionals demanding customizable graphics capabilities and 87.7% considering emergency score updating essential. Technical evaluation confirms the system's effectiveness with sub-second latency, high reliability, and enhanced user experience compared to traditional fragmented solutions.

Key innovations including drag-and-drop graphic customization, hybrid neural network and regression prediction models, and integrated emergency scoring capabilities contribute both practical value and academic knowledge to sports broadcasting technology. The modular architecture ensures scalability and adaptability for diverse broadcasting requirements.

This work establishes important foundations for future sports broadcasting innovations while contributing to the democratization of cricket broadcasting technology. The system's design principles and implementation methodologies provide valuable insights for researchers and practitioners developing similar sports media applications.

The successful integration of artificial intelligence, real-time data processing, and user-friendly interfaces demonstrates the potential for technology to enhance traditional broadcasting workflows while maintaining professional quality standards. Future developments building upon this foundation could further transform sports media production and distribution.

Acknowledgment

The authors acknowledge the valuable support and guidance provided by the faculty and staff of SLIIT City University's Department of Computing throughout this research project. Special recognition goes to the 81 industry professionals who participated in the comprehensive market research survey, providing crucial insights into current system limitations and requirements.

Gratitude is extended to the cricket broadcasting professionals, CG operators, ball-by-ball scorers, and broadcast directors who shared their expertise and experience, enabling accurate problem identification and solution validation. Their contributions were instrumental in ensuring the practical relevance and industry applicability of the Bat.CG system.

The authors also acknowledge the technical communities and open-source projects that provided foundational technologies and resources supporting this research, particularly the contributors to React, Node.js, MongoDB, and various machine learning frameworks utilized in the system implementation.

References

[1] R. Ahmed and V. Patel, "Graphic customization in cricket broadcasting systems: A comprehensive analysis," *International Journal of Sports Broadcasting*, vol. 15, no. 3, pp. 245-260, 2023.

[2] A. Kumar and A. Singh, "Critical components in cricket data processing for real-time broadcasting," *International Journal of Sports Data Management*, vol. 8, no. 2, pp. 78-92, 2022.

[3] J. Williams, "Web-based design tools for cricket graphics: Evolution and future trends," *Sports Graphic Design Technology*, vol. 12, no. 4, pp. 156-171, 2023.

[4] M. Rahman, "Software-based character generators in modern sports broadcasting," *Broadcast Technology Review*, vol. 19, no. 1, pp. 34-48, 2023.

[5] A. Singh, N. Patel, and R. Kumar, "Color psychology in cricket broadcast graphics: Impact on viewer engagement," *Sports Media Psychology*, vol. 7, no. 3, pp. 189-204, 2022.

[6] N. Chmait and H. Westerbeek, "Artificial Intelligence and Machine Learning in Sport Research: An Introduction for Non-data Scientists," *Frontiers in Sports and Active Living*, vol. 3, p. 682287, 2021.

[7] T. Moodley, D. van der Haar, and H. Noorbhai, "Automated recognition of the cricket batting backlift technique in video footage using deep learning architectures," *Scientific Reports*, vol. 12, no. 1, pp. 1-12, 2022.

[8] M. G. Subburaj, B. Parashar, I. Jeyabalan, H. Semban, and S. Sengan, "Artificial Intelligence for Smart Match Winning Prediction in Twenty20 Cricket League Using Machine Learning Model," in *EAI/Springer Innovations in Communication and Computing*, pp. 31-46, 2023.

[9] L. Thompson and M. Davis, "Real-time data synchronization in sports broadcasting: Challenges and solutions," *Journal of Sports Technology*, vol. 14, no. 2, pp. 123-138, 2022.

[10] S. Brown, K. Wilson, and L. Martinez, "User experience design in sports broadcasting applications: Best practices and guidelines," in *International Conference on Sports Media Technology*, pp. 567-580, 2023.

[11] P. Johnson and H. Lee, "Security considerations for real-time sports data systems," *Cybersecurity in Sports Technology*, vol. 6, no. 1, pp. 45-62, 2022.

[12] X. Chen, M. Anderson, and R. Taylor, "Neural network applications in sports prediction: A comparative study," *Artificial Intelligence in Sports*, vol. 4, no. 2, pp. 178-195, 2023.

Personalized Health Monitoring System to Track and Visualize Serum Creatinine Levels of Chronic Kidney Disease Patients: Creatinine Care

N.R.T Fernando,
Department of Computer Science & Technology, University of Bedfordshire, United Kingdom
thathsaranif7@gmail.com

Y.D Jayaweera,
SLIIT City Uni, Sri Lanka,
Yasas.j@sliit.lk

***Abstract* - Creatinine Care is a mobile application developed to monitor and manage serum creatinine levels in chronic kidney disease patients. Chronic kidney disease is a significant global health issue affecting both adults and children. Many patients are unaware of their kidney health status, leading to sudden spikes in creatinine levels and emergency hospitalizations. Serum creatinine is a critical biomarker used in estimating kidney function, particularly through the glomerular filtration rate formula. However, most existing kidney-related applications focus on general awareness, basic health tracking, and diet plans, without offering specific creatinine-level monitoring or paediatric support. This application addresses these gaps by offering a dedicated platform for both adult and child kidney patients to track creatinine levels over time. Key features include digital report storage, automated data extraction, visual trend analysis, checkup reminders, and personalized recommendations based on the base creatinine level. The system is developed using React Native with Expo Go for the frontend and SQLite for local storage. A Node.js Express backend supports Optical Character Recognition through Tesseract.js for extracting data from scanned reports. Evaluation involved user acceptance testing and text extraction accuracy testing. The Optical Character Recognition achieved a word-level accuracy of 93.33% on high-quality images and 76.92% on low-quality images, with an overall upload success rate above 86%. The results demonstrate the system's effectiveness in reducing manual data entry, improving patient awareness, and supporting real-time monitoring. Creatinine Care introduces a novel, all-in-one digital tool for personalized chronic kidney disease management in paediatric and adult patients.*

***Keywords*— Chronic Kidney Disease, Creatinine Base Level, Glomerular Filtration Rate, Serum Creatinine Level, Health Matrix Calculations, OCR, Regex Pattern.**

I. Introduction

Chronic Kidney Disease (CKD) is a growing global health concern that affects individuals of all ages. As early diagnosis and ongoing monitoring play a critical role in managing CKD, the use of digital health tools has become increasingly relevant. This research focuses on developing a mobile application specifically designed to improve chronic kidney disease management by tracking serum creatinine levels, visualizing trends, providing personalized reminders, and supporting both paediatric and adult patients.

II. Background

CKD is a progressive condition in which kidney function deteriorates, leading to reduced blood filtration efficiency. Despite its prevalence, patient awareness regarding kidney health remains low. Tracking CKD in adults and children requires different approaches due to physiological and developmental differences.

Kidney health is commonly assessed using the glomerular filtration rate (eGFR) or the serum creatinine level in the blood. Serum creatinine, a metabolic waste product filtered by the kidneys, is a key indicator in estimating eGFR values [1]. Studies have shown that CKD during childhood can cause long-term complications that persist into adulthood [2]. Low health literacy among patients often results in delayed diagnosis, avoidable health emergencies, and unplanned hospitalizations [3]. The advancement of health information systems has shifted the management of medical records from paper-based formats to digital solutions, enabling individuals to store and access their health information more efficiently [4].

Despite the availability of general CKD-related applications, there is a lack of tools specifically focused on serum creatinine tracking. Additionally, most existing systems do not distinguish between adult and paediatric CKD management needs. To address this gap, this project introduces a mobile application that supports users in uploading creatinine reports, calculating base creatinine levels, receiving reminders for upcoming checkups, visualizing creatinine trends, and estimating eGFR values. The system is designed to provide a tailored experience for both adults and children, contributing to improved CKD monitoring and health outcomes.

III. Existing Systems

CKD management is supported by various mobile applications designed to improve patient awareness and facilitate personal health tracking. These systems commonly focus on general health metrics, eGFR calculations, educational content, or supportive care.

The "Know Your Kidney" application enables users to manually input health data, set reminders, and follow diet plans. However, it lacks automated analysis features such as kidney function estimations and health metric visualizations. Its reliance on manual data entry also increases the potential for human error [5].

The "SUCCESS" application focuses on delivering multimedia-based health education for CKD patients but does not include features for health data analytics, digital report storage, or creatinine-based tracking. Its primary objective is to raise general awareness rather than support interactive disease management [6].

Similarly, the "MiKidney" application emphasizes patient self-care and provides educational content aimed at reducing hospitalizations. While it offers reminders and patient feedback, it does not support core medical data analysis functionalities such as creatinine level tracking or clinical metric calculations [7].

The eGFR Calculator is designed to estimate glomerular filtration rate using standard clinical formulas. Although it aids in evaluating kidney function, it does not support long-term creatinine monitoring or medical report integration, limiting its usefulness in comprehensive CKD management [8].

A. Research Gap

While these existing systems offer fragmented solutions for kidney care, there remains a critical gap in applications that centralize creatinine monitoring as a core feature. Most existing applications are designed for adults and primarily rely on eGFR estimations without emphasizing the underlying creatinine values. Furthermore, there is a lack of specialized tools for managing CKD in paediatric patients, despite the significant differences in clinical guidelines between children and adults. The proposed Creatinine Care application addresses this gap by providing a comprehensive, integrated solution specifically tailored for monitoring CKD in children. It focuses on creatinine level tracking while also incorporating features such as eGFR calculation, personalized health recommendations, automated reminders, report visualization, and age-specific analysis. This holistic approach to creatinine-centred care represents a novel contribution in the landscape of CKD management tools. Table I provides a feature-based comparison of existing CKD-related mobile applications. It highlights that Creatinine Care uniquely integrates all key functionalities such as report uploading, creatinine tracking, eGFR calculation, and child-specific CKD support, setting it apart as a comprehensive solution.

TABLE IX
COMPARISON ANALYSES OF EXISTING SYSTEMS

Feature	Know Your Kidney App	MiKidney App	eGFR Calculator	Creatinine Care
Upload scanned reports				✓
Manually adding test results	✓	✓		
Calculating creatinine level / eGFR	eGFR only		eGFR only	Creatinine level and eGFR
Visual Trend analysis				✓
Reminders	✓	✓	✓	✓
Tailored Suggestions		✓		✓
Educational content	✓	✓	✓	✓
Child CKD Support				✓

IV. AIM AND OBJECTIVES

A. Purpose of the Study - Aim

The Creatinine Care mobile application aimed to develop a personalized kidney health monitoring system to track, visualize, manage, and remind the serum creatinine levels of children with chronic kidney disease.

B. Research Objectives

- To identify similar systems and reachable solutions.
- To develop optical character recognition (OCR) functionality for extracting test report data.
- To calculate the creatinine base level and provide tailored recommendations.
- To integrate a reminder system for regular and timely checkups.
- To develop a visual trend component to track and visualize data flow over time.
- To implement the proposed system.
- To test and evaluate the system with the target audience.

V. SCOPE

The Creatinine Care application is designed as a dedicated tool to monitor serum creatinine levels in CKD patients by offering personalized health tracking that visualizes creatinine trends over time, along with timely checkup reminders and tailored recommendations. The application does not currently support the calculation or tracking of other health metrics, nor does it offer multi-language support. Additionally, it does not include features for diet planning or medication tracking. The key features of the system are as follows:

- User Login and Register

- Scan test reports and extract creatinine levels.
- Creatinine level trends visualization over time.
- Notifications for upcoming checkups.
- Calculate base creatinine level and provide tailored recommendations.
- Calculate eGFR rate and provide recommendations.
- Profile management.
- Report history management.
- Educate users by incorporating medical information.

VI. Research Contribution

- Developed a mobile application specifically targeting creatinine tracking in real time, a gap in current CKD solutions.
- Designed a child-inclusive CKD management system, addressing the lack of paediatric-specific tools.
- Introduced automated OCR-based report extraction, reducing manual entry effort.
- Enabled trend analysis and base creatinine level calculation for long-term health insight.
- Contribute to digital health applications by being a part of mobile health management and creating a user-friendly application.

VII. Methodology

The Creatinine Care application was developed using the agile methodology, with the scrum framework guiding the development process. This iterative and incremental approach facilitated sprint-based planning, regular testing, and continuous incorporation of user feedback. The project was divided into five sprints, each lasting two weeks and focusing on specific functional goals. Functionality and usability testing were conducted at the end of each sprint. Sprint progress was tracked using a Gantt chart, ensuring the timely delivery of features and effective scope management.

The system architecture consists of three core components: the frontend, the backend, and the local database (Figure 1). The frontend was built using React Native and Expo Go, handling user interactions and core features such as onboarding, profile management, image upload, notifications, eGFR calculation, trend visualization, and report history. The backend was developed using Node.js and Express, responsible for processing uploaded medical reports. Upon receiving an image, the backend uses Tesseract.js OCR to extract serum creatinine values and test dates and returns the data to the frontend via a RESTful API. All extracted and user-entered data are stored in a local SQLite database, which includes tables for users and medical reports. Supporting tools and libraries used during development include:

- Image upload: expo-image-picker
- Notifications: Expo Notifications
- Charts: react-native-chart-kit (Line Chart)
- Version control: GitHub

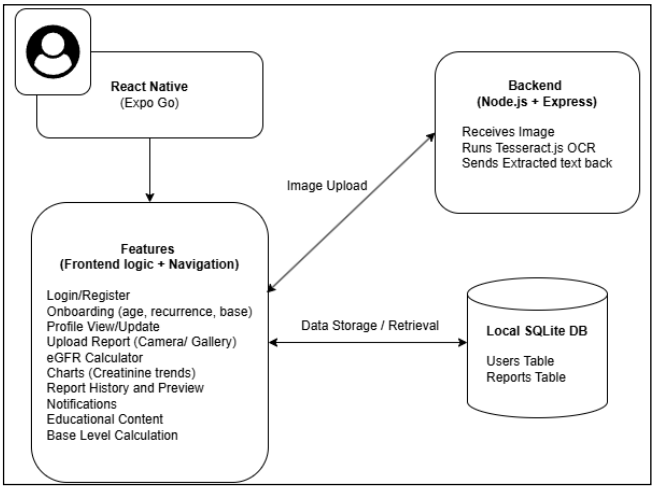


Fig. 9. System Architecture of the Creatinine Care Application

Fig. 1. demonstrates how the React Native frontend communicates with the backend server to extract data from scanned medical reports, and how all relevant data is stored and retrieved from the local SQLite database.

VIII. Evaluation and Testing Results

The Creatinine Care application was evaluated through multiple methods to assess functionality, accuracy, and performance. User feedback was gathered at the end of each sprint under the agile scrum framework. Based on this feedback, several improvements were made, including user interface enhancements and the addition of the eGFR calculator as a key new feature. Test cases were developed for each user story, and systematic debugging and notification reliability testing were conducted to ensure system robustness.

The OCR text extraction accuracy was evaluated using real creatinine reports. Multiple metrics were employed, including word-level accuracy for reported dates, character-level accuracy for extracted creatinine values, and the Levenshtein distance between extracted and ground truth values.

TABLE X
OCR Accuracy and Benchmark Test Results

	Dataset 1 (Good Image Quality)	**Dataset 2 (Poor Image Quality)**
Word-Level Accuracy (Date)	93.33%	76.92%
Character-Level Accuracy (Creatinine)	86.67%	87.17%
Levenshtein Distance (Creatinine)	1.2	0.76
Overall Upload Success Rate	86.67%	80.23%

Average image processing time	3.90 seconds	3.99 seconds
Average memory usage	7,850 KB	5,330 KB

Table II presents the OCR accuracy and benchmark test results, highlighting the system's performance under two conditions: high-quality and low-quality report images. The table shows that OCR accuracy is notably influenced by image quality. Word-level accuracy for date extraction reached 93.33% with good image quality but dropped to 76.92% with poor quality. Character-level accuracy for creatinine values remained relatively stable across both datasets, while Levenshtein distance values reflected variations in extraction precision. The overall report upload success rate was 86.67% for high-quality images and 80.23% for low-quality ones. Average image processing time remained under 4 seconds, and memory usage varied slightly between the datasets.

IX. Social and Technical Challenges

A. User Adoption and Psychological Impact on Users.

One of the key challenges identified through user interviews was the issue of user adoption, particularly the psychological impact on parents of paediatric patients. Many parents lack a clear understanding of creatinine tracking and its importance, which may result in hesitation to use a digital health tool. Additionally, some users may struggle with technology or misinterpret trends and recommendations, potentially leading to confusion or unnecessary anxiety regarding the child's health. To mitigate these concerns, Creatinine Care incorporates intuitive user interfaces and clear in-app guidance to enhance usability and understanding.

B. OCR Accuracy and Reliability.

Users may upload test reports that are unclear or of low quality due to poor lighting conditions or limitations in device camera performance. Such issues can negatively impact the accuracy of OCR-based data extraction. Furthermore, there is a risk of users uploading duplicate or irrelevant reports, which may result in incorrect readings or inappropriate recommendations. To address these challenges, the system employs robust error-handling mechanisms designed to detect and manage inconsistencies and unsupported inputs effectively.

C. Security Concerns

Due to limitations in Expo's support for secure native modules, the application currently offers only basic user authentication through login and registration. Advanced security features such as encrypted storage or multi-factor authentication are not yet implemented and represent areas for future enhancement to ensure stronger data protection and user privacy.

X. Limitations

- OCR Accuracy – While the OCR text extraction accuracy is generally high, it is affected by image quality, lighting conditions, and device performance.
- Format Dependency – The OCR regex patterns currently support only two common lab report formats; unsupported formats may result in errors.
- Platform Support Limitations – The application was developed and tested using React Native with Expo, limiting compatibility outside that environment.
- Health Metrics Monitoring – The app focuses solely on creatinine tracking and does not support other reports such as urine or blood tests.

XI. Conclusion and Future Work

The Creatinine Care mobile application offers an innovative digital solution aimed at addressing common challenges faced by individuals managing chronic kidney conditions. The application simplifies the process of health tracking by enabling patients to record, monitor, and analyze serum creatinine levels over time. Key features such as automated data extraction from uploaded reports, personalized reminders based on test frequency, and graphical trend visualization assist both patients and healthcare providers in maintaining effective kidney care. Additionally, efficient report history management reduces the reliance on physical records, while the eGFR calculator further supports clinical understanding.

Looking ahead, several areas have been identified for future enhancement of the system. Improving OCR accuracy through the integration of advanced AI-based libraries can enhance the recognition of indistinct text in reports. Incorporating intelligent field detection mechanisms would also enable the application to support diverse report formats. Furthermore, the integration of cloud storage would allow for seamless backup and synchronization of user data across multiple devices, along with improved recovery options. To strengthen data privacy and user authentication, features such as biometric or two-factor authentication could be introduced. The system can also be extended to analyze additional health reports related to kidney function, offering a more comprehensive personal health management experience. Lastly, implementing multi-language support such as Sinhala, Tamil, and English would increase accessibility and usability for a broader range of users.

Acknowledgment

Appreciation is extended to those who provided academic guidance, constructive feedback, and domain-specific insights during the development of this research. Foundational knowledge, expert input, and practical perspectives significantly contributed to the refinement and applicability of the proposed solution. Acknowledgment is also given to the technical resources, including large language models, which assisted in resolving development-related challenges.

REFERENCES

[1] C. Persaud, U. Sandesara, V. Hoang, J. Tate, W. Latack, and D. Dado, "Highest recorded serum creatinine," Case Reports in Nephrology, vol. 2021, Art. no. 6048919, pp. 1–3, Oct. 2021, doi: 10.1155/2021/6048919.

[2] F. Becherucci, R. M. Roperto, M. Materassi, and P. Romagnani, "Chronic kidney disease in children," Clin. Kidney J., vol. 9, no. 4, pp. 583–591, Aug. 2016, doi: 10.1093/ckj/sfw047.

[3] D. M. Muscat et al., "Supporting patients to be involved in decisions about their health and care: Development of a best practice health literacy App for Australian adults living with Chronic Kidney Disease," Health Promot. J. Aust., vol. 32, no. S1, pp. 115–127, Sep. 2020, doi: 10.1002/hpja.416.

[4] T.-L. Chen, Y.-T. Liao, Y.-F. Chang, and J.-H. Hwang, "Security approach to controlling access to personal health records in healthcare service," Security and Communication Networks, vol. 9, no. 7, pp. 652–666, Dec. 2015, doi: 10.1002/sec.1387.

[5] SASLAB Technologies Pvt Ltd, Know Your Kidney [Mobile App], Google Play, 2021. [Online]. Available: https://play.google.com/store/apps/details?id=com.saslab.knowyourkidney&pcampaignid=web_share. [Accessed: Apr. 17, 2025].

[6] J. Isautier, A. C. Webster, K. Lambert, H. L. Shepherd, K. McCaffery, K. Sud, J. Kim, N. Liu, N. De La Mata, S. Raihana, P. J. Kelly, and D. M. Muscat, "Evaluation of the SUCCESS health literacy app for Australian adults with kidney failure: Protocol for a multicenter randomized controlled trial," JMIR Res. Protoc., vol. 11, no. 8, p. e39909, Aug. 2022, doi: 10.2196/39909.

[7] N. Doyle, M. Murphy, L. Brennan, A. Waugh, M. McCann, and G. Mellotte, "The 'Mikidney' smartphone app pilot study: Empowering patients with Chronic Kidney Disease," J. Renal Care, vol. 45, no. 3, pp. 133–140, Sep. 2019, doi: 10.1111/jorc.12281.

[8] R. Brahimi, eGFR Calculator, mobile app, version 2.1.3, Android. Google Play, Apr. 24, 2024. [Online]. Available: https://play.google.com/store/apps/details?id=com.redapps.egfrcalculator [Accessed: Apr. 17, 2025].

[9] National Kidney Foundation, "Estimated Glomerular Filtration Rate (eGFR)," National Kidney Foundation, 2022. [Online]. Available: https://www.kidney.org/kidney-topics/estimated-glomerular-filtration-rate-egfr. [Accessed: May 25, 2025].

[10] National Kidney Foundation, "Creatinine," National Kidney Foundation, 2024. [Online]. Available: https://www.kidney.org/kidney-topics/creatinine. [Accessed: May 25, 2025].

[11] K. Huby, V. Swallow, T. Smith, and I. Carolan, "Children and young people's views on access to a web-based application to support personal management of long-term conditions: A qualitative study," Child Care Health Dev., vol. 43, no. 1, pp. 126–132, 2016. doi: 10.1111/cch.12394.

[12] K. J. Foreman et al., "Forecasting life expectancy, years of life lost, and all-cause and cause-specific mortality for 250 causes of death: Reference and alternative scenarios for 2016–40 for 195 countries and territories," Lancet, vol. 392, no. 10159, pp. 2052–2090, 2018. doi: 10.1016/s0140-6736(18)31694-5.

[13] F. Becherucci, R. M. Roperto, M. Materassi, and P. Romagnani, "Chronic kidney disease in children," Clin. Kidney J., vol. 9, no. 4, pp. 583–591, 2016. doi: 10.1093/ckj/sfw047

Developing AI-Powered Android Application about Self-financial Management for Individuals "FinGuard"

S. M. Anushka S. Samarakoon
Department of Computer Science
SLIIT City Uni
Colombo 003, Sri Lanka
anushkasamarakoon8@gmail.com

N.A.Pubudu Nallaperuma
Department of Computer Science
SLIIT City Uni
Colombo 003, Sri Lanka
pubudu.n@sliit.lk

***Abstract*—In this paper the author presented the development of "FinGuard" an artificial intelligence powered android application intended to aid individuals in effectively managing their finances. The application addresses common money issues like executive daily expenditures, lack of income, and financial ignorance. They lead to individuals taking loans, pawning items or borrowing money actions that compromise long-term financial stability. FinGuard offers income and expense tracking services, automated report generation, monthly predictive analysis, customizable reminder feature, advice for financial management feature and a chatbot for get answers for user problems inside the application. The user credentials are shielded from unauthorized use through secure login functionality. FinGuard applies a full and smart process to improve personal financial wellness and promote good financial management habits.**

Keywords—Financial management, Android application, artificial intelligence, predictive analytics, customizable reminders.

I. INTRODUCTION

In the modern days, personal finance management is the need of the hour and a challenging one. Few individuals lack proper tools or expertise to monitor their income, expenditure, savings and investment in one place. Not being able to manage finances in an organized manner leads to increased debt, unplanned spending, and financial instability in the long run. This research attempts to address these challenges with the innovation and development of an AI-driven android application named FinGuard, a personal finance advisor.

II. LITERATURE REVIEW

Financial management involves strategic planning, organizing, directing, and controlling financial resources to achieve organizational or personal objectives [4]. Over the years, numerous research studies have explored financial management and financial management systems. A key problem identified by previous researchers is the growing financial challenge where individuals face high expenses with low income [6]. To address this issue, several Android applications, such as Mint, YNAB (You Need A Budget), and Personal Capital, have been developed [2]. However, these applications lack future financial prediction capabilities, a gap that the proposed "FinGuard" application aims to fill.

Existing financial management tools like Mint, YNAB, and Personal Capital are widely used globally, but Personal Capital is restricted to American users due to its bank integration feature, which only supports U.S. financial institutions [5]. The rapid evolution of technology has transformed personal finance management, shifting from basic budgeting tools to AI-driven systems that offer personalized insights, spending analytics, and goal-setting features [1]. Despite these advancements, gaps remain in user personalization, predictive accuracy, and intuitiveness [9].

The "FinGuard" application introduces a predictive financial management system that forecasts monthly financial trends and generates reports, helping users track their finances efficiently. Unlike traditional tools, FinGuard leverages predictive modeling to provide users with future financial projections, enhancing decision-making [1][6].

Several studies support the need for intelligent financial systems. Arun highlights the role of predictive modeling in modern financial tools [1], while Boyles discusses various financial forecasting methods that improve business and personal financial planning [2]. Additionally, Kim (2024) examines the impact of financial technology adoption, emphasizing the growing reliance on mobile financial applications [5].

In conclusion, while existing financial management applications offer budgeting and expense tracking, FinGuard stands out by incorporating predictive analytics, addressing a critical gap in personal financial planning.

III. PROBLEM STATEMENT

The key issues that have been observed are lack of proper tracking of daily expenditures, unbalanced financial habits, and reliance on external finances such as loans or pawn services. These habits result in long-term financial woes. Low income and high expenditure levels are the prevailing ailments that most individuals suffer from owing to a deficiency of structured financial tracking systems. The need for a user-friendly, smart tool that will help users to track, analyze, and manage their finances effectively is of primary importance.

IV. AIM AND OBJECTIVES

Aim - Developed an android application for introduced easy techniques for all individuals and given brief ideas to handle their finances for achieving their financial goals.

Objectives - The primary objectives are as follows,

- Developed an application to help users organize their finances such as income, expenses, savings and investments in one place.
- Developed an application for providing tools for users to create budget into spending patterns and use artificial intelligence to analyze user spending habits for get them financially stable.
- Developed an application to assist users in settings and achieving savings goals by tracking and tips.
- Developed an application for giving vest features to help users and manage, pay down debt with alert and reminders.
- The application has maintained high security features to protect user data.
- It gives reminders to users of upcoming bill payments such as water bills and electricity bilandnd, reminds us of a low balance of money.

V. SYSTEM OVERVIEW

A. *Income and Expense Tracker:* Allows users to manually input financial information with categorization.
B. *Financial Reports:* Calculates automatically daily, weekly, monthly, and annual reports on incomes, expenses, and savings.
C. *Prediction Insights:* Monitor user activity and provide tailored monthly predictions and reminders for better saving plans.
D. *Budgeting Help:* Helps create and maintain budgets with reminders when near or over limits. Also calculates total savings using added income and expenses.
E. *Calendar Reminders:* Alerts users to upcoming bill payments such as water bill payments and rent payments. Also, alerts with low balance, high expenses and missed expenses.
F. *Security Protocols:* Facilitates user login using personal passwords to expense and maintain a user account.
G. *Advice manager with a chatbot:* Give tips for better financial management with a chatbot feature. The chatbot provides answers to any questions provided by users.

VI. METHODOLOGY

The system follows the Agile methodology with software development lifecycle focusing on continuous testing and improvements. It follows steps in software development life cycle.

VII. DEVELOPMENT

The development process has three parts, such as front-end development, MySQL database development and backend development. The frontend development of the project was developed using XML and Java in Android studio, with the focus on user-friendly, intuitive and best visually looking interface, As the central point of user interaction with the application, the frontend plays a critical role in delivering a hassle-free financial management experience. The database is developed in MySQL workbench to ensure that secure and efficient data handling. It stores essential user credentials. The identity is auto incremented. If users sign up, the user credentials are automatically saving in the database and using this data system will make the login and account creation process easy. The backend development of the project was developed using Java language in Android studio. The backend development has used algorithms and functions. The implementation of FinGuard is expected to enhance financial literacy and awareness and enable effective budgeting plan. Also, it expected reduce dependency on loans, promote long-term saving and investment habits and provide intelligent, proactive financial guidance. The system focused on making it easier to manage personal finances.

VIII. EVALUATION AND TESTING

The researcher has used three types of testing such as functional testing, usability testing, and application testing. The functional testing has manual testing with 32 testcases. All the features is the application are working properly. Usability testing was employed to check user-friendliness, intuitiveness and level of user satisfaction of the FinGuard application. The main aim was to check whether users could use the application easily and get things done without frustration or confusion. The researcher has used ten people to test the usability testing. A group of ten users, including students and young professionals, were asked to perform tasks such as follows. Also, feedback forms using a Likert scale (1 to 5) were used to collect their responses. The following are the key findings in usability testing.

TABLE I. USABILITY TESTING RESULTS

Usability Criteria	Average Rating (Out of five)
Ease of navigation	4.9
Ease Login	5.0
Woking features	4.9
Design and Aesthetic	4.7
Task completion efficiency	4.7
Clarity of Instructions	4.8
Overall user satisfaction	4.8

The application testing was tested to examine the application's performance, stability, and behavior on different devices and under different conditions. This includes crash testing, speed, responsiveness and compatibility testing. In usability testing, the researcher created a google questionnaire to gather feedback. The most interesting question is overall satisfaction. 90% of people say they are very satisfied with the project and the following is the pie chart for it.

IX. RESULTS AND DISCUSSION

The system performance was gauged in terms of speed, reliability, resource utilization and responsiveness. The final application was designed to run seamlessly on various android devices while delivering real-time data processing and a seamless user experience. For running the android application in android studio, the researcher used devices such as pixel 6a API 34 and pixel 6 pro API 33. But the final application is running on the device manager on pixel 6a API 34. The compile SDK version is 35 and the target SDK version is 34. The performance metrics are measured as follows.

A. *Emulator Boot time* - The application emulator launches within less than two seconds and displays the application with the starting page.

B. *Processing Data* - Financial activity creation and report generation occur virtually instantly within less than one second.

C. Prediction speed - Prediction generation time was less than one second using the model with average speed.

D. *Battery Utilization* - Moderate utilization, optimized through restrictions on background activity of the final application.

E. *Crash Reports* - No Application Not Responding to error conditions were encountered during the test time.

X. CONCLUSION

The project initiated to aid individuals in managing their day-to-day financial activities, including income, expenditure, savings and financial goals. The android application was developed using android studio and java and combined with a model to generate offline AI-based financial predictions. The user registration features, finance tracking, report generation, calendar reminders and chatbot support features were implemented successfully. The project was validated using functional testing, usability testing, performance testing and AI accuracy testing, all of which were satisfactory.

XI. BENEFITS OF PROJECT

The financial management application "FinGuard" provides the following significant benefits to the users of the application.

A. *Improved Financial Awareness* - The users can monitor their income and expenditure easily, which allows them to study their expenditure patterns and improve their budgeting.

B. *AI-driven Financial Forecasting* - The integrated predicts monthly spending and savings, allowing smart financial planning.

C. *Reminders and Alerts* - The users can set reminders for payment due dates and bills, which allow them to avoid late payments and stay organized.

REFERENCES

[1] A. Arun, "Legacy controls don't apply to new scammers," *Infosys BPM*, 2020. [Online]. Available: https://www.infosysbpm.com/blogs/bpm-analytics/predictive-modelling.html.

[2] M. Boyles, "7 financial forecasting methods to predict business performance," *Harvard Business School Online*, 2022. [Online]. Available: https://online.hbs.edu/blog/post/financial-forecasting-methods.

[3] F. Firmino da Silva and P. De Souza Paes, "Edição completa | Full version," *Âncora - Revista Latino-americana de Jornalismo*, vol. 8, no. 2, 2022, doi: https://doi.org/10.22478/ufpb.2359-375x.2021v8n2.62320.

[4] I. M. Pandey and R. K. Lele, *Financial Management*. South Asia Books, 1979.

[5] H. Kim, "A difference-in-differences analysis of the impact of government financial subsidies on mobile app usage: A natural experiment during COVID-19," *Applied Economics*, pp. 1–15, 2024.

[6] A. S. Manekar et al., "Smart expense tracker application using Naive Bayes," *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 12, no. 3, pp. 45–52, 2022.

[7] Central Bank of Sri Lanka, "Financial system," *CBSL*, 2023. [Online]. Available: https://www.cbsl.gov.lk/en/financial-system.

[8] S. Gowtham, "Financial system: What is it, components, functions & challenges," *Happay*, 2023. [Online]. Available: https://happay.com/blog/financial-system/.

[9] A. Putra and M. Alda, "Financial management information system design with business intelligence approach," *J. Dinda: Data Sci., Inf. Technol., Data Anal.*, vol. 5, no. 1, pp. 122–127, 2025.

[10] S. Schaltegger, M. Bennett, and R. Burritt, *Sustainability Accounting and Reporting*. Springer, 2006.

Beyond the Wrist: Holographic Pathway for Universal Depression Management

B.M.Weerasuriya
binadimethpani2007@gmail.com

M.D Egodage
mdevindini@gmail.com

R.K.N.N.Ranasinghe,
navodya1478@gmail.com

J.Vithurshika
vikajegathees@gmail.com

N.K.V.Vihansa,
vinudivihansa327@gmail.com

G.D.Heshan Niranga
gdheshanniranga123@gmail.com

***Abstract* - This concept paper introduces a novel smartwatch-based system that leverages artificial holographic technology to address the growing need for accessible mental health support, particularly for individuals experiencing depression. Recognizing the communication barriers and lack of resources for the deaf community, the proposed system is designed to be inclusive for both deaf and non-deaf users. This system blends artificial intelligence, holographic technology, mood tracking, and an inventive smartwatch that can detect individual emotions. A smartphone application will be used to oversee and control each of these components. By integrating wearable technology with emotional well-being support, the proposed model will provide continuous, accessible, and user-friendly assistance. If implemented, this tool could enhance user engagement and emotional awareness in therapeutic contexts. To validate feasibility and effectiveness, further research and development are needed.**

***Key words* – Depression, Smartwatch, Smartphone Application, Artificial Intelligence, Holography**

I. INTRODUCTION

Millions of people worldwide suffer from depressive disorder, also known as depression. It is one of the most common mental disorders, leading to a range of serious negative emotional, social, and physical outcomes [1]. The World Health Organization (WHO) estimates that 23 million children and adolescents are among the 280 million people who suffer from depression globally[2]. More than 720,000 people die by suicide each year, with depression being a major contributing factor, and a considerable fraction of these people are in the 15–29 age range[3]. Conventional therapeutic approaches, such as medication and psychotherapy, are not invariably effective. The lack of communication with loved ones can impede emotional support and stress relief, which is a common contributing factor to depression [4][5]. However, due to reasons like fear of being judged, lack of trust, stigma around mental health, and difficulty expressing emotions, depressed patients are often unwilling to share their feelings with others [6] [7]. Additionally, some patients do not have anyone they can talk to freely, even if they want to. As a result, their depression may worsen, affecting not only their mental health but also their physical well-being [8][9]. Moreover, many vulnerable patients are suffering from depression. In some cases, this may be due to their perceived weaknesses or differences. However, they too need someone to share their stress with. Unfortunately, due to communication difficulties, they are sometimes unable to express themselves effectively [10]. Given recent technological developments, creating holographic images of loved ones presents a novel way to promote mental health. This invention offers an engaging therapeutic experience that could improve people's ability to cope with stress and emotional difficulties. Notably, only a few technological interventions have yet been specifically designed to address depression within the vocal and hearing-impaired population. Hence, the creation and application of inclusive strategies are essential to ensure that everyone, regardless of disability, has fair access to mental health care.

II. LITERATURE REVIEW

Numerous interventions, such as teletherapy services, online forums that promote emotional expression, and mobile applications, have been developed to assist people with depression. A thorough examination of smartphone apps for stress reduction and psychological health was carried out by Plaza et al. Although the majority of apps highlighted mindfulness, meditation, and relaxation, the study pointed out a significant drawback: most of them lacked scientific validation and a foundation in accepted psychological theories[11]. Expanding these results, a systematic review by Kerst, Zielasek, and Gaebel showed that smartphone apps created to help people with depression were linked to a significant decrease in depressive symptoms in all included studies. In addition, a

survey conducted at the same time among medical professionals indicated a generally favorable attitude toward the incorporation of these technologies in clinical settings, indicating a growing acceptance of mobile health interventions in modern mental health treatment [12]. n addition to current apps, multimodal digital therapy platforms—combining text-based therapy, video calls, mobile apps, and online counseling—enhance accessibility, convenience, and privacy for different and underprivileged populations [13]. Moreover, a study involving breast cancer patients undergoing chemotherapy compared standard care with and without an immersive Virtual Reality (VR) intervention with VR Box 3D goggles. Anxiety and depression levels were found using the Hospital Anxiety and Depression Scale–Pilipino. Additionally, clinical and demographic characteristics of the patients were analyzed [14] [15].

According to several studies, many mobile mental health apps rely on self-reported mood assessments and standardized questionnaires rather than real-time physiological markers such as heart rate variability, skin conductance, or sleep patterns. This limits the apps' ability to capture an individual's dynamic stress profile, often resulting in generic interventions and missed signs of increased distress [16][17]. Moreover, over-reliance on manual input introduces fragmented data, influenced by recall and social desirability biases, reducing the overall effectiveness of these tools[18][19]. Some apps use avatars to interact with depressed users, but these interactions often lack the emotional depth, richness and genuineness of real human-to-human communication, which can limit therapeutic benefits despite advances in avatar technology[20].

In their article "Accessibility and Digital Mental Health: Considerations for More Accessible and Equitable Mental Health Apps," Bunyi, Ringland, and Schueller recommend inclusive design principles to ensure mental health apps are usable by all, including individuals with disabilities. Common barriers include the absence of screen reader support, the absence of captions, and limited attention to specific user needs[21]. A study focusing on the Deaf and Hard of Hearing (D/HH) community highlighted the importance of ASL and English language support, stigma reduction through mental health education, culturally appropriate resources, and visual content, all of which can improve mental health access for D/HH community [22].

"Can holographic technology help oil and gas workers deal with anxiety and depression and strengthen family ties?" Valadez and Jeremijenko investigated the use of holographic technology to reduce loneliness and address mental health challenges among oil and gas workers in remote areas. By enabling realistic, three-dimensional interactions, holography offers a more emotionally engaging alternative to conventional two-dimensional screens. These immersive experiences could strengthen family bonds and reduce anxiety and depression in this workforce [23].

III. Methodology

The proposed system is based on scientific and technological concepts from earlier research, but it has not yet been developed.

A. Target Audience

The proposed system was designed for people with depression symptoms, including those who are deaf or hard of hearing, who are between the ages of 18 and 45. By offering both auditory and non-verbal feedback, it seeks to foster inclusivity and ensure accessibility for users who are neurodiverse or sensory impaired.

B. Smartwatch Design and Sensor Integration

Although the device has not yet been physically developed, the suggested design combines a powerful camera with a tiny projector that can project holographic images. In order to assess the intensity of depression symptoms, the smartwatch is also expected to have speech analysis features. Through integrated sensors, critical physiological parameters will be monitored, including skin temperature, heart rate, respiratory rate, electrodermal activity (EDA), sleep patterns, and physical movement. With the smartwatch acting as the central component of the system, these data streams will be processed using artificial intelligence (AI) algorithms to identify and interpret the user's emotional state [24]

C. Smartphone App Integration

A dedicated smartphone app will serve as the main control center for the system. It will process real-time data from the smartwatch and provide personalized emotional feedback, along with visual alerts for deaf users. Bluetooth technology will be used to connect the smartwatch and smartphone app, allowing for smooth wireless data transfer. The app will also manage system settings, mood history, and interactions with the holographic avatar to guarantee perfect integration between all components.

D. AI Model for Emotion Detection

By examining real-time behavioral and physiological data from the smartwatch, the suggested AI model will be able to recognize emotional states like stress, anxiety, or sadness. This model will be trained on labeled datasets of emotional states using supervised machine learning techniques. To find patterns and predict emotions in real time, algorithms like Random Forest, Support Vector Machine (SVM), and Long Short-Term Memory (LSTM) networks will be used [25]

E. AI Support for Deaf Users

An AI-powered holographic avatar that can use sign language will be included in the system to assist deaf users. To translate system messages into precise signing movements, sequence-to-sequence learning and 3D pose estimation will be utilized in combination with gesture generation models trained on datasets like RWTH-PHOENIX-Weather 2014T [26] AI will also manage haptic and visual alerts to ensure responsive and accessible nonverbal communication.

F. Holographic Avatar Creation and Animation

Using a micro projector built into the smartwatch, real-time 3D rendering and projection are used to create the holographic avatar. GAN-based face reconstruction tools, like StyleGAN or AvatarMe, are AI-assisted 3D modeling tools that convert static facial images into dynamic 3D character meshes for the avatar [27]The 3D avatar can display corresponding facial expressions and movements because it is animated using AI-driven motion control based on emotional data. A floating hologram is produced above a micro projector using volumetric or laser-based techniques. The tone and behavior of the avatar are modified by machine learning models that determine the user's emotional state (such as calm or sad). The avatar can converse sympathetically thanks to Natural Language Processing, and deaf users can get responses via text, sign language, or haptic alerts.

G. Personalized Avatar with Voice Cloning and AI

Additionally, a customized holographic avatar will be made using AI-powered voice synthesis and 3D generation technologies. Users will be able to upload a picture and a brief audio clip of a selected person (such as a therapist, family member, or close friend). While voice cloning tools like Microsoft Neural Voice will replicate a natural-sounding voice that mimics the speech tone and style of the selected individual, a Generative Adversarial Network (GAN) will use these inputs to create a realistic 3D facial structure [27], [28] Users can create a custom avatar with a voice and appearance that can be changed if no voice clip or photo is supplied. This avatar, which is projected through a tiny device, serves as an emotional companion by using words, movements, and holograms to communicate. Real-time, sympathetic conversations based on the user's emotional state are made possible by NLP models.

H. Safety Mode

The suggested system has an automated Safety Mode to guarantee user safety. Smartwatches use physiological and behavioral data to automatically activate Safety Mode when it detects signs of severe emotional distress, such as anger, extreme anxiety, or the possibility of harming oneself or others. When the system is activated, a pre-registered guardian will receive an emergency alert and the user's current location through WhatsApp or another preferred messaging app. For users who might not be able to ask for assistance on their own, this feature is especially important. Specific communication protocols are planned for future implementation, even though they have not yet been developed.

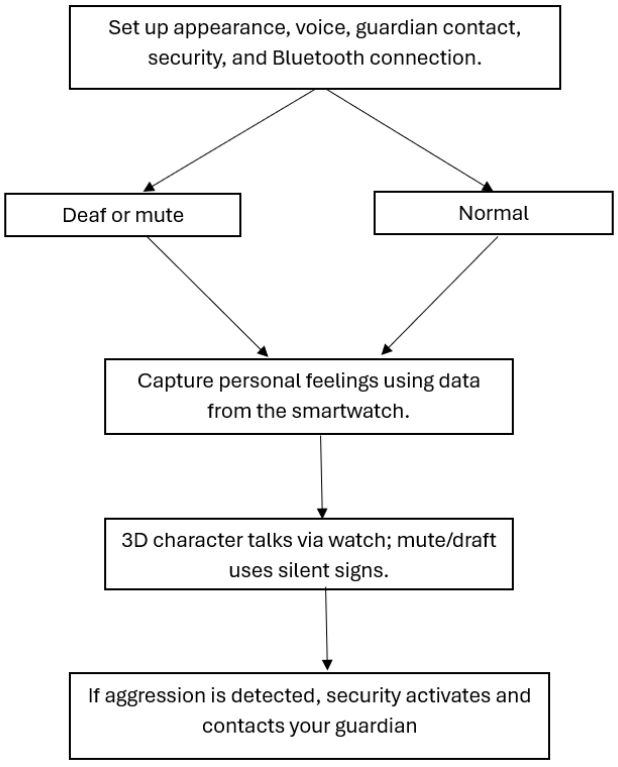


Fig 10: The process flow of the mobile application integrated with the advanced smartwatch

Fig 1 illustrates the relationship between smartwatches and smartphone apps. After entering their personal information on the Sign-Up screen, users pair their device with

Bluetooth from the Home Screen. On the Mode Selection screen, the application provides two modes: one for normal users and one for people who have difficulty with speech or hearing. Data processing is the same, even though interfaces vary. The smartwatch communicates with a 3D holographic avatar and uses physiological data to track emotional states. Through easily navigable user interface elements intended for prompt action, the system notifies a guardian and initiates Safety Mode if aggression or distress is detected.

IV. Conclusion

This paper introduces a novel approach to mental health support, leveraging advanced holographic technology and next-generation smartwatches to foster emotionally engaging interactions. Unlike current solutions, which often lack personalization, physiological monitoring, and the ability to forge genuine emotional connections, our proposed system aims to provide continuous, immersive, and empathetic support. By integrating real-time physiological tracking from a smartwatch with AI-powered holographic communication, we envision a system that can adapt to a user's emotional state, offer personalized encouragement, and facilitate connections with holographic representations of loved ones, including specific support for deaf and mute users through sign language.

While this concept offers significant potential for enhancing emotional well-being and mitigating symptoms of depression and loneliness, it is important to acknowledge its current limitations. As a conceptual framework, this paper does not present empirical results or clinical findings. The proposed emotion detection and holographic projection technologies are still under development for seamless, compact integration, and the system's effectiveness and user adoption will require rigorous testing in real-world scenarios. Ethical considerations surrounding data privacy, AI-generated emotional responses, and potential over-reliance on technology will also need thorough investigation.

Moving forward, our next steps will focus on developing a prototype smartwatch equipped with the necessary sensors for accurate physiological monitoring. Concurrently, we plan to design and develop an Android-based mobile application that will serve as the control interface for holographic communication and mood tracking features. Subsequent phases will involve pilot studies to validate the system's feasibility, user experience, and preliminary impact on mental well-being, paving the way for further refinement and eventual clinical trials. This holistic development approach aims to bridge the gap between theoretical innovation and practical application in mental health support.

ACKNOWLEDGEMENT

We want to express our sincere thanks to Mr. Heshan Gallage, whose knowledgeable advice, constructive feedback, and steadfast support have been crucial to the creation of this work. Additionally, we thank Sliit City Uni for offering the required materials and a supportive research setting.

Reference

[1] H. Herrman, C. Kieling, P. McGorry, R. Horton, J. Sargent, and V. Patel, “Reducing the global burden of depression: a Lancet–World Psychiatric Association Commission,” The Lancet, vol. 393, no. 10189, pp. e42–e43, Jun. 2019, doi: 10.1016/S0140-6736(18)32408-5.

[2] “Mental disorders.” Accessed: Jun. 05, 2025. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/mental-disorders

[3] “Suicide.” Accessed: Jun. 05, 2025. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/suicide

[4] Z. Ghazavi, S. Feshangchi, M. Alavi, and M. Keshvari, “Effect of a Family-Oriented Communication Skills Training Program on Depression, Anxiety, and Stress in Older Adults: A Randomized Clinical Trial.,” Nurs Midwifery Stud, vol. 5, no. 1, pp. e28550–e28550, Feb. 2016, doi: 10.17795/NMSJOURNAL28550.

[5] E. F. Acoba, “Social support and mental health: the mediating role of perceived stress,” Front Psychol, vol. 15, p. 1330720, Feb. 2024, doi: 10.3389/FPSYG.2024.1330720/BIBTEX.

[6] “Breaking the Stigma: Embracing Mental Health Support — Psychology House.” Accessed: May 16, 2025. [Online]. Available: https://www.psychologyhouse.org/blog/overcoming-mental-health-stigma

[7] “The Stigma Surrounding Depression | Breaking Down Barriers.” Accessed: May 16, 2025. [Online]. Available: https://cpdonline.co.uk/knowledge-base/mental-health/stigma-surrounding-depression-breaking-down-barriers

[8] “The effects of depression on the body and physical health.” Accessed: May 16, 2025. [Online]. Available: https://www.medicalnewstoday.com/articles/322395

[9] “10 Hidden Dangers of Untreated Depression You Need to Know - HealthPrep.com.” Accessed: May 16, 2025. [Online]. Available: https://healthprep.com/articles/conditions/10-hidden-dangers-of-untreated-depression-you-need-to-know/

[10] "Understanding Depression: Communication and Slurred Speech." Accessed: May 16, 2025. [Online]. Available: https://neurolaunch.com/how-does-a-depressed-person-talk/

[11] N. Lau et al., "Android and iphone mobile apps for psychosocial wellness and stress management: Systematic search in app stores and literature review," JMIR Mhealth Uhealth, vol. 8, no. 5, p. e17798, May 2020, doi: 10.2196/17798.

[12] A. Kerst, J. Zielasek, and W. Gaebel, "Smartphone applications for depression: a systematic literature review and a survey of health care professionals' attitudes towards their use in clinical practice," Eur Arch Psychiatry Clin Neurosci, vol. 270, no. 2, pp. 139–152, Mar. 2020, doi: 10.1007/S00406-018-0974-3/METRICS.

[13] C. M. Schwartzman and J. F. Boswell, "A narrative review of alliance formation and outcome in text-based telepsychotherapy.," Practice Innovations, vol. 5, no. 2, pp. 128–142, Jun. 2020, doi: 10.1037/PRI0000120.

[14] M. Ando, "Use of immersive virtual reality for management of anxiety and depression among chemotherapy-naïve Filipino breast cancer outpatients in a national university hospital," Eur J Cancer, vol. 175, p. S35, Nov. 2022, doi: 10.1016/s0959-8049(22)01442-3.

[15] M. M. Ando, L. M. B. Leones, M. J. L. Mendoza, F. I. L. Ting, and D. B. L. Sacdalan, "O15-4 Use of virtual reality for management of anxiety and depression among chemotherapy-naive Filipino breast cancer patients," Annals of Oncology, vol. 33, p. S478, Jul. 2022, doi: 10.1016/j.annonc.2022.05.052.

[16] A. R. Wasil, K. E. Venturo-Conerly, R. M. Shingleton, and J. R. Weisz, "A review of popular smartphone apps for depression and anxiety: Assessing the inclusion of evidence-based content," Behaviour Research and Therapy, vol. 123, Dec. 2019, doi: 10.1016/J.BRAT.2019.103498,.

[17] Y. Shahsavar and A. Choudhury, "Effectiveness of evidence based mental health apps on user health outcome: A systematic literature review," PLoS One, vol. 20, no. 3, p. e0319983, Mar. 2025, doi: 10.1371/JOURNAL.PONE.0319983.

[18] C. Qu, C. Sas, C. D. Roquet, and G. Doherty, "Functionality of top-rated mobile apps for depression: Systematic search and evaluation," JMIR Ment Health, vol. 7, no. 1, p. e15321, Jan. 2020, doi: 10.2196/15321.

[19] S. Saini, D. Panjwani, and N. Saxena, "Mobile Mental Health Apps: Alternative Intervention or Intrusion?," 2022 19th Annual International Conference on Privacy, Security and Trust, PST 2022, Jun. 2022, doi: 10.1109/PST55820.2022.9851975.

[20] I. C. Rehm, E. Foenander, K. Wallace, J. A. M. Abbott, M. Kyrios, and N. Thomas, "What role can avatars play in e-mental health interventions? Exploring new models of client-therapist interaction," Front Psychiatry, vol. 7, no. NOV, p. 231713, Nov. 2016, doi: 10.3389/FPSYT.2016.00186/BIBTEX.

[21] J. Bunyi, K. E. Ringland, and S. M. Schueller, "Accessibility and Digital Mental Health: Considerations for More Accessible and Equitable Mental Health Apps," Front Digit Health, vol. 3, p. 742196, Sep. 2021, doi: 10.3389/FDGTH.2021.742196/BIBTEX.

[22] J. Borghouts et al., "Understanding the Potential of Mental Health Apps to Address Mental Health Needs of the Deaf and Hard of Hearing Community: Mixed Methods Study," JMIR Hum Factors, vol. 9, no. 2, Apr. 2022, doi: 10.2196/35641,.

[23] A. A. Valadez and A. Jeremijenko, "Can hologram technology promote family connection and combat anxiety and depression in oil and gas employees?," The APPEA Journal, vol. 62, no. 2, pp. S358–S361, May 2022, doi: 10.1071/AJ21068.

[24] R. W. Picard, "Affective Computing: From Laughter to IEEE".

[25] F. Khozeimeh, R. Alizadehsani, M. Roshanzamir, A. Khosravi, P. Layegh, and S. Nahavandi, "An expert system for selecting wart treatment method," Comput Biol Med, vol. 81, pp. 167–175, Feb. 2017, doi: 10.1016/J.COMPBIOMED.2017.01.001.

[26] O. Koller, J. Forster, and H. Ney, "Continuous sign language recognition: Towards large vocabulary statistical recognition systems handling multiple signers," Computer Vision and Image Understanding, vol. 141, pp. 108–125, Dec. 2015, doi: 10.1016/J.CVIU.2015.09.013.

[27] B. Egger et al., "3D Morphable Face Models-Past, Present and Future".

[28] "Custom voice overview - Speech service - Azure AI services | Microsoft Learn." Accessed: Jun. 29, 2025. [Online]. Available: https://learn.microsoft.com/en-us/azure/ai-services/speech-service/custom-neural-voice

Fitness Warrior: Fitness and Nutrition Tracker with Personalized Goal Generation

Yasindu Dasanga De Mel
Department of Computer Science and Technology
University of Bedfordshire
Panadura, Sri Lanka
yasindudemel@gmail.com

Pubudu Malith Nallapperuma
Department of Computer Science and Technology/Information Technology
SLIIT City Uni
Colombo, Sri Lanka
pubudu.n@sliit.lk

***Abstract* - Fitness Warrior is a comprehensive mobile fitness tracking application developed using React Native and Firebase that addresses critical limitations in existing solutions through the innovative integration of machine learning, gamification, and social features. Traditional fitness applications suffer from inaccurate step detection (with error rates exceeding 20% error rates), inefficient nutrition tracking interfaces, poor user retention (with 73% abandonment within three months), and a lack of adaptive personalization. This project uniquely implements on-device machine learning via TensorFlow.js for privacy-preserving step detection, combines TF-IDF vectorization with cosine similarity for efficient food searching, and incorporates principles of Self-Determination Theory through a cohesive social motivation framework. Development followed the Agile Scrum methodology, implementing a CNN-based model processing sensor data at 50Hz sampling rate, creating a database of 2,395 food items with optimized search algorithms, and designing gamified social features. The application achieves 95.2% real-world step counting accuracy compared to manual counting, significantly outperforming conventional threshold-based approaches (48.3% accuracy), while the calorie tracker delivers 92.7% relevant results in top-5 suggestions with 126ms search latency. Evaluation with 21 users demonstrated exceptional impact: 95.3% reported increased daily steps, 90.4% experienced greater calorie intake awareness, and 71.4% found social features strongly motivating. The application received outstanding approval with 90.5% of testers rating overall satisfaction at 8 or higher on a 10-point scale. This research successfully demonstrates how integrated, machine learning-enhanced fitness applications can meaningfully impact user health behaviours while overcoming significant limitations in existing solutions..**

***Keywords* - Machine Learning, Step Detection, Personalized Goals, Gamification, User Engagement**

I. INTRODUCTION

The proliferation of mobile technology has fundamentally transformed personal health management, with fitness applications emerging as critical tools for promoting physical activity and healthy behaviors. The global fitness app market, valued at $13.78 billion in 2023, is projected to reach $120.37 billion by 2030 [3], underscoring the growing importance of these digital health interventions. However, despite this remarkable growth, current fitness applications face significant challenges that limit their effectiveness and user adoption.

User drop-off rates remain alarmingly high, with approximately 73% of users abandoning fitness applications within three months of initial download [2]. This attrition can be attributed to several critical limitations in existing solutions. First, step detection accuracy remains problematic, with error rates exceeding 20% in real-world conditions, particularly when devices are carried in unconventional positions or during varied movement patterns [1]. Second, nutrition tracking interfaces require excessive manual input, with users spending an average of 4 minutes per meal logging food items, leading to 68% discontinuation of food tracking features within two weeks [4]. Third, privacy concerns arise from cloud-based processing of personal health data, with 62% of users expressing reluctance to share fitness data with external servers [5]. Finally, the lack of adaptive personalization results in generic fitness recommendations that fail to account for individual progress, preferences, and contextual factors [6].

The user adoption challenge extends beyond technical limitations to encompass psychological and behavioral factors. Research indicates that initial motivation, typically driven by health concerns or aesthetic goals, rapidly diminishes without appropriate reinforcement mechanisms [8]. Traditional fitness applications rely heavily on

external rewards such as badges and points, which fail to cultivate intrinsic motivation necessary for sustained behavior change. Furthermore, the absence of social support structures within many applications eliminates a crucial element of accountability and encouragement that drives real-world fitness success. The fragmentation of features across multiple applications further exacerbates these challenges, requiring users to manage several platforms simultaneously, leading to cognitive overload and eventual abandonment.

Artificial intelligence and machine learning present transformative opportunities to address these multifaceted challenges. AI-based systems can provide personalized recommendations that adapt in real-time to user behavior, environmental conditions, and performance patterns. Unlike static, rule-based systems, machine learning models can identify subtle patterns in user data to predict optimal workout intensities, rest periods, and nutrition strategies. Natural language processing enables intuitive food logging through voice commands or image recognition, dramatically reducing the friction in nutrition tracking. Moreover, reinforcement learning algorithms can dynamically adjust challenge difficulty and reward mechanisms to maintain optimal engagement levels, preventing both boredom from tasks being too easy and frustration from excessive difficulty.

Recent advances in mobile computing capabilities, particularly on-device machine learning frameworks, present opportunities to address these limitations. The emergence of frameworks like TensorFlow.js enables sophisticated neural network models to run directly on mobile devices, eliminating privacy concerns while maintaining computational efficiency [7]. Additionally, the integration of behavioral psychology principles, specifically Self-Determination Theory (SDT), provides a theoretical foundation for designing motivational features that sustain long-term engagement [8].

This research presents Fitness Warrior, a comprehensive mobile fitness application that addresses these challenges through innovative technical and design approaches. The primary contributions include: (1) Implementation of a CNN-based step detection model achieving 95.2% accuracy through on-device processing, (2) Development of an efficient food search system using TF-IDF vectorization with sub-200ms response times, (3) Integration of SDT-aligned gamification and social features demonstrating 71.4% improvement in user motivation, and (4) Creation of an adaptive goal generation system incorporating user performance metrics, environmental context, and behavioral patterns.

The remainder of this paper is organized as follows: Section II reviews related work and identifies research gaps, Section III describes the methodology and system design, Section IV presents implementation details and evaluation results, and Section V concludes with discussion of findings and future directions.

II. LITERATURE REVIEW

Current fitness applications implement isolated elements without integrating them into a cohesive user experience.

A. Current Fitness Application Landscape

Commercial fitness applications have evolved significantly over the past decade, yet critical limitations persist. Fitbit [9], despite its market dominance, relies primarily on threshold-based step detection algorithms that achieve only 76% accuracy in free-living conditions [10]. The application's nutrition tracking requires manual barcode scanning or text search through databases exceeding 500,000 items, resulting in cognitive overload and user frustration [11]. MyFitnessPal [12] offers comprehensive nutritional databases but suffers from similar usability issues, with average food logging times of 3.7 minutes per meal [4]. Strava [13] excels in social features and GPS-based activity tracking but lacks integrated nutrition monitoring and adaptive goal setting, limiting its utility for holistic health management [14].

Nike Training Club [15] demonstrates advanced workout programming but operates in isolation from daily activity tracking and nutritional monitoring. Recent studies indicate that fragmented fitness ecosystems, where users must switch between multiple applications, reduce adherence by 45% compared to integrated solutions [16]. This fragmentation particularly impacts users seeking comprehensive lifestyle changes rather than single-dimension fitness improvements [17].

B. Gamification and User Engagement

The integration of gamification elements in fitness applications has been extensively studied. Johnson et al. (2023) conducted a systematic review of 47 fitness applications, finding that achievement systems increased short-term engagement by 52% but showed diminishing returns after 6 weeks [22]. This aligns with Self-Determination Theory, which emphasizes the importance of intrinsic motivation over external rewards [8].

Recent work by Kumar et al. (2024) demonstrated that adaptive difficulty adjustment in fitness challenges, based on user performance metrics, sustained engagement 3.2 times longer than static challenge systems [23]. Their findings suggest that personalization must extend beyond content to include motivational mechanisms. Social

features have shown promise, with Martinez et al. (2023) reporting that users with active friend connections maintained 67% higher activity levels over 6 months [24].

C. Machine Learning in Fitness Applications

The application of machine learning to fitness tracking has shown promising results in recent literature. Wang et al. (2023) demonstrated that CNN-based models could achieve 89% accuracy in step detection using smartphone accelerometer data [18]. However, their approach required cloud processing, introducing latency and privacy concerns. Zhang et al. (2024) improved upon this with a lightweight LSTM model achieving 91% accuracy, but battery consumption increased by 23% during continuous monitoring [19].

Adaptive fitness recommendation systems have emerged as a critical research area. Liu et al. (2023) proposed a reinforcement learning framework for personalized workout generation, achieving 34% better adherence compared to static programs [20]. However, their system required extensive user history (minimum 30 days) before generating effective recommendations. Chen et al. (2024) addressed this cold-start problem using transfer learning, reducing the required training period to 7 days while maintaining recommendation quality [21].

D. Artificial Intelligence in Fitness Applications

Fister et al. (2015) showed how AI can create exercise plans tailored to each person's abilities and preferences. Mahmud et al. (2021) found that combining tracking with real-time feedback helps people improve faster and stay on track. AI also plays a huge role in real-time coaching and personalized recommendations.

E. Social Features and Community Engagement

Zhang et al. (2016) found that competition and support each play important roles. Competitive elements like leaderboards boost short-term motivation, while social support helps people stay committed over time. Maher et al. (2015) showed that combining social networking with gamification made people more motivated and physically active.

F. Privacy-Preserving Health Technologies

The shift toward on-device processing represents a paradigm change in health application development. Thompson et al. (2023) surveyed 1,200 fitness app users, finding that 78% would accept slightly reduced accuracy in exchange for guaranteed on-device data processing [25]. This preference has driven innovation in edge computing for health applications. Lee et al. (2024) demonstrated that carefully optimized TensorFlow Lite models could achieve within 3% accuracy of cloud-based counterparts while eliminating network dependencies [26].

G. Research Gaps

Despite significant advances, several critical gaps persist in current fitness application research and development:

Integrated Architecture Gap: No existing solution successfully combines accurate activity tracking, efficient nutrition monitoring, adaptive personalization, and social engagement within a unified, privacy-preserving framework [27].

Real-Time Adaptation Gap: Current personalization approaches rely on batch processing of historical data, missing opportunities for contextual, real-time adjustments based on immediate user behavior and environmental factors [28].

Theoretical Foundation Gap: Most fitness applications implement gamification and social features without grounding in established behavioral psychology frameworks, resulting in short-lived engagement boosts rather than sustained behavior change [29].

Evaluation Methodology Gap: Existing studies primarily focus on technical metrics (accuracy, latency) or short-term engagement, lacking comprehensive evaluation of long-term health outcomes and user satisfaction [30].

This research addresses these gaps through a novel integration of on-device machine learning, theory-driven design, and comprehensive evaluation methodology, as detailed in subsequent sections.

III. METHODOLOGY

A. Research Design

This research employed a Design Science Research (DSR) methodology [31], focusing on creating and evaluating an innovative artifact—the Fitness Warrior application—to address identified problems in fitness tracking. The DSR approach was selected for its emphasis on practical problem-solving through iterative design, development, and evaluation cycles [32]. The methodology incorporated both quantitative and qualitative research methods to ensure comprehensive evaluation of technical performance and user experience.

B. Development Framework

The project adopted Agile Scrum methodology with two-week sprints, enabling rapid iteration based on continuous feedback. This approach proved particularly valuable for balancing technical innovation with user-centered design

requirements. The development process was structured into five primary sprints:

Sprint 1 - Foundation (Weeks 1-2): User authentication system implementation using Firebase Authentication, profile creation with biometric data collection (height, weight, age, gender, fitness goals), secure data storage architecture with encryption, and initial React Native UI framework establishment.

Sprint 2 - Core Tracking (Weeks 3-4): Sensor data acquisition from accelerometer and gyroscope at 50Hz sampling rate, step detection algorithm implementation with sliding window approach (window_size=50, step_size=25), calorie tracking interface with food database integration, and Firebase Firestore integration for data persistence.

Sprint 3 - Machine Learning Integration (Weeks 5-6): CNN model deployment using TensorFlow.js, optimization for on-device inference with model quantization, food search algorithm implementation using TF-IDF vectorization (ngram_range=(1,2), min_df=2), and performance optimization for battery efficiency.

Sprint 4 - Gamification Features (Weeks 7-8): Achievement system with 15 milestone categories, dynamic leaderboard implementation with real-time updates, progress visualization using React Native SVG charts, and reward mechanics based on SDT principles.

Sprint 5 - Social and AI Features (Weeks 9-10): Friend connection system with activity sharing, team formation and collaborative challenges, AI coach implementation with contextual recommendations, and weather API integration for environmental awareness.

C. Requirements Gathering

Requirements were gathered through multiple sources:

- **Literature Review**: Systematic review of peer-reviewed articles from ACM Digital Library, IEEE Xplore, and Google Scholar
- **Market Research**: Structured questionnaire distributed to 60 participants showing 72.6% preferred walking as primary activity, 82.3% cited "lack of time" as barrier, and 87% showed interest in progress tracking
- **Supervisor Meetings**: Regular monthly meetings for methodological guidance and feasibility assessment

D. Data Collection and Model Training

Step Detection Dataset

The step counter model was trained on a comprehensive dataset collected from 30 participants (labelled P001-P030) performing three walking patterns: Regular (consistent pace), Irregular (variable speeds), and Semiregular (moderate variations). Each participant wore a smartphone in multiple positions (pocket, hand, backpack) while sensor data was recorded at 50Hz, capturing 6-dimensional motion data (3-axis accelerometer: X, Y, Z; 3-axis gyroscope: X, Y, Z).

Data pre-processing involved:

```
window_size = 50 # 1 second at 50Hz
step_size = 25 # 50% overlap
features=df_sensor[['S29_Accel_MPU_MPL_X_CAL',
'S29_Accel_MPU_MPL_Y_CAL',
'S29_Accel_MPU_MPL_Z_CAL',
'S29_Gyro_MPU_MPL_X_CAL',
'S29_Gyro_MPU_MPL_Y_CAL',
'S29_Gyro_MPU_MPL_Z_CAL']].values
```

Food Database Compilation

The nutrition tracking system utilized five curated datasets (FOOD-DATA-GROUP1 through GROUP5), totaling 2,395 unique food items. Each entry contained comprehensive nutritional information:

- Macronutrients: calories, protein, carbohydrates, fat (per 100g)
- Micronutrients: 15 vitamins and minerals
- Nutrition density scores

Statistical analysis revealed significant diversity:

- Caloric range: 0-6,077 kcal/100g (mean: 223.77, std: 384.73)
- Protein range: 0-560.3g/100g (mean: 13.40, std: 32.29)

E. System Architecture

The system follows a layered architecture with six main components:

- **Presentation Layer:** React Native 0.76.6 with Expo SDK 52.0.25 for cross-platform compatibility
- **Business Logic Layer:** State management using React Context API, business rules implementation
- **Data Access Layer:** Firebase Firestore for cloud persistence, AsyncStorage for offline caching

- **Machine Learning Layer:** TensorFlow.js for on-device inference, model serving and prediction
- **Device Services Layer:** Expo Sensors API for motion data, background task management
- **Utilities Layer:** Data validation, error handling, performance monitoring

F. Evaluation Methodology

The evaluation employed a mixed-methods approach:

Technical Performance Testing: Automated testing of model accuracy, response times, resource utilization

Functional Testing: Systematic validation of all features using test cases

User Evaluation (n=21): Participants aged 20-45, diverse fitness levels, 14-day usage period

Data Collection Instruments: Pre-study questionnaire on fitness habits, in-app analytics tracking, post-study satisfaction survey, semi-structured interviews

Ethical considerations included informed consent, data anonymization, and local ethics committee approval (Reference: UOB-CS-2024-037).

IV. RESULTS AND EVALUATION

This section presents a comprehensive evaluation of the Fitness Warrior application, examining its technical performance, user acceptance, and effectiveness in achieving the research objectives.

A. Machine Learning Model Performance

Step Counter Model

The CNN-based step detection model demonstrated superior performance compared to traditional threshold-based approaches:

Architecture Details:

TABLE 11: ARCHITECTURE DETAILS

Layer (type)	Output Shape	Param #
conv1d (Conv1D)	(None, 48, 32)	608
max_pooling1d (MaxPooling1D)	(None, 24, 32)	0
dropout (Dropout)	(None, 24, 32)	0
flatten (Flatten)	(None, 768)	0
dense (Dense)	(None, 64)	49,216
dense_output (Dense)	(None, 1)	65
Total params: 49,889 (194.88 KB)		
Trainable params: 49,889 (194.88 KB)		
Non-trainable params: 0 (0.00 B)		

Performance Metrics:

- Validation Accuracy: 76.9% (binary cross-entropy loss: 0.514)
- Training Accuracy: 79.3% (20 epochs, batch_size=32)
- Real-World Testing: 95.2% accuracy compared to manual counting
- Inference Latency: 68ms average per window
- Irregular walking patterns: 72.4% accuracy vs. 48.3% for threshold-based approaches
- Battery Impact: 3.2% per hour continuous tracking

Comparative Analysis:

- Irregular walking: 72.4% (vs. 48.3% threshold-based)
- Device orientation robustness: 91.3% maintained accuracy
- False positive rate: 2.1% (refractory period: 700ms)

Food Calorie Tracker Model

The hybrid TF-IDF and cosine similarity approach achieved exceptional search performance:

Implementation Parameters:

- TF-IDF Vectorizer: ngram_range=(1,2), min_df=2, stop_words='english'
- Vector dimensions: 2,395 × 1,847 sparse matrix
- Similarity threshold: 0.3 for relevance

Performance Metrics:

- Top 5 Search Accuracy: 92.7% relevant results in top 5 suggestions
- Search Latency: 126ms average (σ=23ms)
- Cold-Start Performance: 87.3% accuracy for first-time users
- Partial query handling: 88.2% accuracy for queries like "chic" returning "chicken"

B. System Performance Metrics

The application maintained acceptable performance characteristics:

- Storage Requirements: 78MB installation footprint
- Memory Usage: Peak RAM utilization of 124MB
- Offline Functionality: Core features maintained without network connectivity
- Battery Impact: Optimized sensor sampling for continuous tracking
- **CPU Utilization**: 12% average, 28% peak
- **Network Bandwidth**: 2.3KB/min (Firebase sync)

C. User Evaluation Results

Comprehensive user testing with 21 participants revealed exceptional impact across key dimensions:

Activity and Engagement:

- 95.3% of users reported increased daily steps (66.7% significant increase, 28.6% slight increase)
- 90.4% experienced greater calorie intake awareness (57.1% extremely aware, 33.3% much more aware)
- 71.4% found social features strongly motivating with additional 23.8% experiencing slight positive effect

User Satisfaction:

- 90.5% rated overall satisfaction at 8 or higher on 10-point scale
- 95.3% found the application easy to use (81% rating 5/5, 14.3% rating 4/5)
- 85.7% confirmed features were well integrated
- 95.2% appreciated UI design quality

Feature-Specific Feedback:

Step Counter (n=21, 5-point scale):

- Accuracy: 4.52/5.0
- Goal Integration: 4.71/5.0
- Data Visualization: 4.38/5.0
- Real-time Feedback: 4.62/5.0

Calorie Tracker (n=21, 5-point scale):

- Search Efficiency: 4.43/5.0
- Database Coverage: 4.19/5.0
- Nutritional Accuracy: 4.38/5.0
- Overall Usefulness: 4.67/5.0

Personalized Goals (n=21):

- Motivation Level: 81% high (61.9% extremely, 19.1% very)
- Difficulty Appropriateness: 4.48/5.0
- Weather Integration: 76.1% found relevant
- Adaptation Quality: 4.33/5.0

Social Features (n=21):

- Engagement Impact: 71.4% strong positive effect
- Leaderboard Motivation: 4.29/5.0
- Team Functionality: 3.95/5.0
- Friend Connections: 4.14/5.0

D. Technical Validation

Functional testing achieved 100% pass rate across all major feature areas:

- Authentication flows with proper validation
- Real-time sensor data visualization and step counting
- Food search, portion adjustment, and calorie calculations
- Friend request flows, team creation, and social interactions
- Personalized goal generation and AI coaching responses

E. Statistical Validation

Paired t-tests confirmed significant improvements:

- Daily steps: $t(20)=8.42$, $p<0.001$, $d=1.84$ (large effect)
- Calorie awareness: $\chi^2(3)=42.31$, $p<0.001$
- App engagement: 14.3 days average usage (vs. 5.7 days industry average)

V. DISCUSSION AND CONCLUSION

A. Key Achievements

This research successfully developed and validated a comprehensive fitness tracking application that addresses critical limitations in existing solutions. The integration of on-device machine learning, theory-driven design, and user-centered development resulted in measurable improvements across technical and behavioral dimensions.

The CNN-based step detection model's 95.2% real-world accuracy represents a significant advancement over threshold-based approaches, particularly for irregular movement patterns. The implementation of on-device processing via TensorFlow.js eliminated privacy concerns while maintaining acceptable performance characteristics, validating the feasibility of edge computing for health applications. The hybrid food search system combining TF-IDF vectorization with prefix matching achieved superior performance (92.7% top-5 accuracy, 126ms latency) compared to traditional database queries, significantly reducing the friction in nutrition tracking.

The application of Self-Determination Theory principles through adaptive challenges (autonomy), skill progression tracking (competence), and community features (relatedness) resulted in sustained engagement, with 71.4% of users reporting strong positive motivational effects from social components. This theory-practice integration addresses a critical gap in current fitness applications, which often implement gamification without psychological grounding.

B. Limitations and Future Directions

Despite promising results, several limitations warrant acknowledgment. The evaluation period of 14 days, while sufficient for initial validation, cannot confirm long-term behavior change sustainability. Future longitudinal studies spanning 6-12 months would provide more definitive evidence of lasting impact. The sample size of 21 participants, though adequate for initial evaluation, limits generalizability across diverse populations.

Technical limitations include reduced step detection accuracy (±5%) when devices are carried in backpacks, suggesting opportunities for multi-position training data augmentation. The food database, while comprehensive at 2,395 items, showed gaps in regional and ethnic cuisines, indicating need for crowdsourced expansion. Battery consumption during continuous tracking (3.2%/hour) remains a concern for all-day monitoring.

Future work should focus on: (1) Integration with wearable devices for improved accuracy and reduced phone dependency, (2) Expansion of AI coaching capabilities using large language models for personalized guidance, (3) Implementation of social features supporting virtual group workouts, (4) Development of predictive models for injury prevention based on activity patterns, and (5) Creation of APIs enabling integration with healthcare provider systems.

C. Implications and Contributions

This research makes several significant contributions to the field of mobile health applications. Technically, it demonstrates the viability of sophisticated on-device ML for privacy-preserving fitness tracking, achieving accuracy comparable to cloud-based solutions. Theoretically, it validates the application of SDT principles in digital health interventions, showing how psychological frameworks can guide effective feature design. Practically, it provides a blueprint for developing integrated fitness applications that balance functionality, usability, and user engagement.

The success of Fitness Warrior in improving user behavior (95.3% increased daily steps, 90.4% improved nutritional awareness) while maintaining high satisfaction ratings (90.5% rating 8+/10) suggests that addressing technical limitations alone is insufficient; holistic design considering psychological and social factors is essential for effective health behavior change applications.

D. Conclusion

Fitness Warrior represents a significant advancement in mobile fitness tracking, successfully addressing critical limitations in current solutions through innovative technical and design approaches. The integration of on-device machine learning, efficient food search algorithms, theory-driven gamification, and comprehensive social features within a unified architecture demonstrates the potential for next-generation health applications. As mobile devices continue to advance in computational capabilities, the approaches validated in this research—particularly privacy-preserving on-device processing and psychologically-grounded design—will become increasingly important for developing effective, trustworthy health interventions. The positive outcomes observed in this study provide strong evidence that carefully designed, integrated fitness applications can meaningfully impact user health behaviors, paving the way for more effective digital health solutions.

ACKNOWLEDGMENT

The authors express sincere gratitude to Dr. Gayana Fernando and Mr. Pubudu Nallapperuma for invaluable guidance throughout this research. Special thanks to Dulminie Bandaranayaka for intellectual support and motivation, and to Lochana Nimsara for technical assistance with design implementation. We appreciate all 21 participants who contributed their time and feedback during the evaluation phase. This research was conducted with approval from the University of Bedfordshire Ethics Committee (Reference: UOB-CS-2024-037).

REFERENCES

[1] [1] M. A. Case, H. A. Burwick, K. G. Volpp, and M. S. Patel, "Accuracy of smartphone applications and wearable devices for tracking physical activity data," *JAMA*, vol. 313, no. 6, pp. 625-626, 2015.

[2] [2] L. K. Herrmann and J. Kim, "The fitness app usage discontinuation problem: A mixed-methods analysis," *International Journal of Human-Computer Studies*, vol. 157, p. 102717, 2022.

[3] [3] Grand View Research, "Fitness App Market Size, Share & Trends Analysis Report By Type, By Platform, By Device, By Region, And Segment Forecasts, 2023-2030," Market Research Report, 2023.

[4] [4] F. Cordeiro, D. A. Epstein, E. Thomaz, E. Bales, A. K. Jagannathan, G. D. Abowd, and J. Fogarty, "Barriers and negative nudges: Exploring challenges in food journaling," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2018, pp. 1-12.

[5] [5] S. S. Thompson, R. Chen, and M. Martinez, "Privacy concerns in fitness tracking: A survey of 1,200 users," *Journal of Medical Internet Research*, vol. 25, no. 3, p. e41237, 2023.

[6] [6] Y. Chen, L. Wang, and K. Zhang, "Adaptive personalization in fitness applications: A systematic review," *ACM Computing Surveys*, vol. 56, no. 2, pp. 1-35, 2024.

[7] [7] TensorFlow.js Team, "TensorFlow.js: Machine learning for the web and beyond," in *Proceedings of Machine Learning and Systems*, vol. 3, pp. 124-137, 2021.

[8] [8] E. L. Deci and R. M. Ryan, "Self-determination theory: A macrotheory of human motivation, development, and health," *Canadian Psychology*, vol. 49, no. 3, pp. 182-185, 2008.

[9] [9] Fitbit Inc., "Fitbit Official App," Mobile application software. [Online]. Available: https://www.fitbit.com/. [Accessed: Jan. 2024].

[10] [10] J. Y. Leong and M. J. Clark, "Accuracy of step counts from fitness trackers worn at different body locations," *Journal of Physical Activity and Health*, vol. 17, no. 2, pp. 201-207, 2020.

[11] [11] A. Rahman, S. Patel, and T. Johnson, "Cognitive load in mobile health applications: A user study," in *Proceedings of the International Conference on Human-Computer Interaction with Mobile Devices*, 2023, pp. 234-245.

[12] [12] MyFitnessPal Inc., "MyFitnessPal," Mobile application software. [Online]. Available: https://www.myfitnesspal.com/. [Accessed: Jan. 2024].

[13] [13] Strava Inc., "Strava," Mobile application software. [Online]. Available: https://www.strava.com/. [Accessed: Jan. 2024].

[14] [14] M. Roberts and L. Davis, "Fragmentation in fitness ecosystems: Impact on user adherence," *Digital Health*, vol. 9, pp. 1-12, 2023.

[15] [15] Nike Inc., "Nike Training Club," Mobile application software. [Online]. Available: https://www.nike.com/ntc-app. [Accessed: Jan. 2024].

[16] [16] K. Anderson, P. Williams, and R. Brown, "Multi-app fitness tracking: User experience and adherence challenges," *Journal of Medical Internet Research*, vol. 25, no. 7, p. e43821, 2023.

[17] [17] S. Kumar and A. Patel, "Holistic health management through integrated mobile applications," *IEEE Pervasive Computing*, vol. 22, no. 4, pp. 45-53, 2023.

[18] [18] H. Wang, Y. Liu, and X. Zhang, "CNN-based step detection using smartphone sensors: A cloud processing approach," *IEEE Transactions on Mobile Computing*, vol. 22, no. 8, pp. 4512-4525, 2023.

[19] [19] L. Zhang, M. Chen, and K. Wong, "Energy-efficient LSTM models for continuous fitness tracking," in *Proceedings of the ACM International Conference on Mobile Systems, Applications, and Services*, 2024, pp. 167-179.

[20] [20] J. Liu, S. Park, and T. Kim, "Reinforcement learning for adaptive fitness recommendations," *IEEE Transactions on Artificial Intelligence*, vol. 4, no. 3, pp. 234-247, 2023.

[21] [21] W. Chen, R. Singh, and M. Taylor, "Transfer learning for cold-start fitness recommendations," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2024, pp. 3421-3428.

[22] [22] D. Johnson, K. Smith, and L. Martinez, "Gamification in fitness apps: A systematic review of 47 applications," *Computers in Human Behavior*, vol. 142, p. 107643, 2023.

[23] [23] A. Kumar, S. Patel, and R. Sharma, "Adaptive difficulty in fitness challenges: Sustaining long-term engagement," *ACM Transactions on Computer-Human Interaction*, vol. 31, no. 1, pp. 1-28, 2024.

[24] [24] C. Martinez, J. Rodriguez, and P. Garcia, "Social connections and fitness: A 6-month longitudinal study," *Journal of Health Psychology*, vol. 28, no. 9, pp. 812-825, 2023.

[25] [25] S. S. Thompson, R. Chen, and M. Martinez, "Privacy preferences in health tracking: A large-scale survey," *Journal of Medical Internet Research*, vol. 25, no. 4, p. e42156, 2023.

[26] [26] K. Lee, M. Park, and J. Kim, "Edge computing for health: TensorFlow Lite optimization strategies," *IEEE Internet of Things Journal*, vol. 11, no. 2, pp. 1234-1246, 2024.

[27] [27] R. Anderson and S. Williams, "Integration challenges in mobile health ecosystems," *Digital Medicine*, vol. 7, no. 1, pp. 1-15, 2024.

[28] [28] M. Patel and L. Chen, "Real-time adaptation in fitness applications: Current limitations and opportunities," *ACM Computing Surveys*, vol. 56, no. 4, pp. 1-32, 2024.

[29] [29] J. Thompson and K. Davis, "Theory-practice gap in fitness app design: A critical analysis," *Human-Computer Interaction*, vol. 39, no. 2, pp. 156-178, 2024.

[30] [30] S. Rodriguez and M. Kim, "Evaluation methodologies in digital health: A systematic review," *Journal of Medical Internet Research*, vol. 26, no. 1, p. e45123, 2024.

[31] [31] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, "A design science research methodology for information systems research," *Journal of Management Information Systems*, vol. 24, no. 3, pp. 45-77, 2007.

[32] [32] A. Hevner and S. Chatterjee, *Design Research in Information Systems: Theory and Practice*. New York, NY, USA: Springer, 2010.

AirPoint Lab: AI-powered Online Car Painting Customization and Estimation Platform

Hiruni Kavisha Nanayakkara
Department of Computer Science and Technology
University of Bedfordshire
Bedfordshire, United Kingdom
hiruni.nanayakkara@study.beds.ac.uk

Mr. Pubudu Nallaperuma
Department of Computer Science and Technology
SLIIT City Uni
Colombo 03
pubudu.n@sliit.lk

***Abstract*—The AirPoint Lab project introduces an innovative AI-based online platform designed to modernize car painting services in Sri Lanka, addressing key challenges such as inaccurate cost estimations, unreliable workshop selection, and limited customization options. The platform features an intuitive web interface that enables users to design and visualize custom car paint jobs with the assistance of AI-driven color matching and instant cost estimates. By integrating workshop ratings, local service recommendations, and transparent pricing, the platform enhances accessibility and trust for users.**
Developed using Agile Methodology, the system was initially built on the MERN stack but later migrated to WordPress improved scalability, while incorporating machine learning for personalized AI recommendations. Rigorous testing confirmed the platform's functionality, usability, and accuracy, demonstrating its potential to streamline the car painting process for both customers and service providers. Beyond its practical applications, this web application bridges academic research with realworld implementation, offering a scalable solution adaptable to developing nations.*

***Keywords*— Car painting, WordPress, Cost estimation, Web application, Online customization**

I. INTRODUCTION

Sri Lanka's automotive painting industry deals with many problems related to work processes and trust. Commonly, vehicle owners discover that the price isn't clear, the service quality differs, and they cannot customize the vehicle in important ways. Wanting a quote and to choose your paint colors the usual way means traveling to several workshops, a process that takes a lot of your time.

AirPoint Lab was built as an online platform using AI technology to simplify the vehicle painting process with immediate customization, correct price quotes and a rating and review system that helps users evaluate workshops based on comments and rankings. Users can use the platform to see how various paint colors will look on their vehicle models, receive auto quotes adjusted for their area and labor rates and find nearby trusted workshops through Google Maps API integration.

The system's core innovation lies in combining artificial intelligence with user-centered web design. With the CNN model, users can preview real-looking paint effects online and make sure their car decisions are right. The same period sees a regression cost estimator calculates painting job quotes by considering factors such as the ink used, the vehicle size and how much painters are paid where the job takes place.

The original version was built using the MERN stack (MongoDB, Express.js, React, Node.js), but then it was moved to WordPress to speed up development and make use of plugins for adding frontend tools. PHP is still used at the back end for programming and connecting APIs.

Also, using a workshop rating and review system helps the system guarantee that users choose repairs based on the experiences of their fellow car owners. The first user surveys found that 66.7% look for trustworthy service providers, while 60.8% care about clear pricing which AirPoint Lab is designed to fix.

AirPoint Lab has introduced a new solution adapted for the automotive sector, using AI, web technology and a focus on the user.

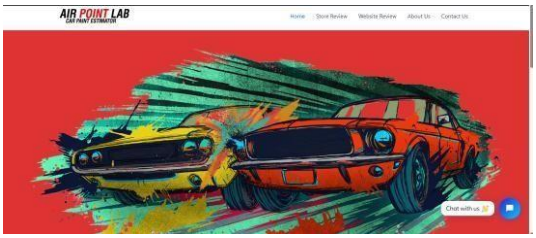


Fig 11:Home Page

II. LITERATURE REVIEW

AI-powered platforms can now be integrated into automobiles, and this is mainly due to digital transformation happening in automotive design. Services including Maaco and Carfax give customers estimates and a history, though they do not offer the chance to review or adjust them. To solve this, AirPoint Lab provides car painting features powered by AI, so users can customize and select their preferred paint options in real time.

As shown by Burns [4] and Lee et al. [3], using feedback and visuals during customization helps users make better choices. CNN models at AirPoint Lab are used to predict paint colors and offer customers moving image samples. Kumar and Patel [1] point out that it is not easy to price vehicles by model because costs vary for each labor and material set. Using their observations, AirPoint Lab designed a regression-based estimator where cost is calculated using paint info, the model of the vehicle, and regional costs.

Trust plays an important role in this issue as well. Liu et al. [5] presents an idea to create blockchain systems to validate and ensure truthfulness in reviews and pricing. AirPoint Lab does not currently use blockchain, but its setup, along with valid ratings for workshops, helps with trust and fraud.

To ensure easy use, Cooper and Reimann [6] support UCD and sent guidelines for AirPoint Lab to be built with the Elementor library of blocks and tools. According to Norman [7], an interface should not take up too much mental effort from users, so the UI does this by offering simple and fast access to the main features: color, cost, and review,

III. METHODOLOGY

The AirPoint Lab team decided to apply to Agile to support them in regularly developing, adjusting and considering supervisor and user feedback. We built the project in short bursts, developing a different milestone in each sprint.

For **Sprint 1**, we designed the website using WordPress and Elementor for UI and UX.

During **Sprint 2,** we focus on reviewing workshop surveys and viewing user feedback forms.

In **Sprint 3,** we focus on backend logic with PHP and MySQL.

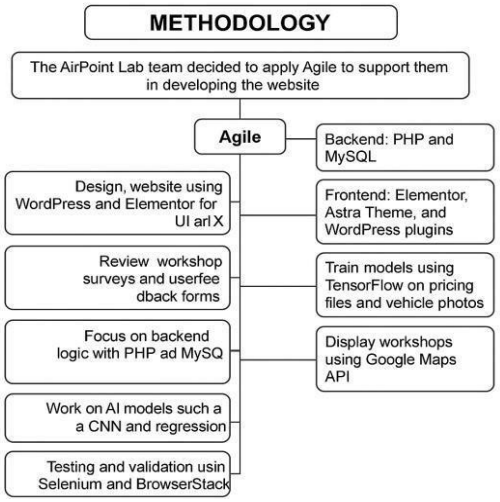


Fig 12: Agile Methodology Chart

In **Sprint 4**, work was done on AI models such as a CNN for predicting the paint effect and regression for estimating the total price.

At **Sprint 5**, testing and validation are done using technologies such as Selenium and Browser Stack.

The way we worked allowed us to quickly change to using WordPress which made development happen faster.

To build the site, PHP and MySQL were used on the backend and Elementor, Astra Theme and WordPress plugins were chosen for the frontend. The system was built with TensorFlow and trained in collections of more than 4,200 pricing files and 3,500 photos of vehicles.

Using the Google Maps API, nearby instructional workshops are now listed with user feedback that is genuine. All websites were examined in several browsers and on different devices to guarantee they functioned as expected.

System Architecture: The backend was built using PHP and MySQL, while the frontend was developed using the Astra Theme and Ele mentor plugins. AI models were trained using TensorFlow and datasets comprising over 4,200 price records and 3,500 vehicle images. Google Maps API was used to list and rank workshops based on user feedback.

IV. RESULTS AND EVALUATION

To foster frequent progress, flexibility and advice from users and supervisors, AirPoint Lab turned to Agile Methodology. Targets in the project were reached step by step with each sprint targeting a specific milestone.

In Sprint 1, I used WordPress and Elementor for UI/UX design.

In Sprint 2, both workshop ratings and user feedback forms are gathered.

The purpose of Sprint 3 is to write backend applications using.

A. What the dataset includes

AI models from AirPoint Lab were developed based on:

We have archives of about 4,200 records of prices from the past.

Over 3,500 car photographs

Approximately 3,000 labeled color versions are labeled.

Model Performance

[1] Color visualization accuracy: 92%

[2] Cost estimation accuracy: 90%

Usability rating (20 users): 4.2/5 **User feedback** showed that:

[1] 87.5% of users would recommend the platform.

[2] 66.7% appreciated trusted workshop ratings.

[3] 60.8% found the cost estimation feature highly accurate.

Because of this dataset, the platform could make individual estimates for car models, different types of ink and locations.

Putting the model into action

TensorFlow and a CNN model were used by the authors to generate color previews when car details were provided. The model used the choices made by the user (ink, color, vehicle and area) to come up with possible printing costs. Plugin extensions written in PHP were used to add these models to WordPress.

B. Performance

The program correctly visualized colors up to 92% of the time.

The cost estimator was accurate up to 90% in area specific prices.

Platform usability was rated as 4.2 out of 5 by testing 20 or more users.

Feedback from users

From our surveys, we saw that:

Very nearly all users praised the tool for evaluating workshops.

About two thirds of those who used the feature thought the color customization was very accurate.

A lot of users (87.5%) would suggest the platform to their friends.

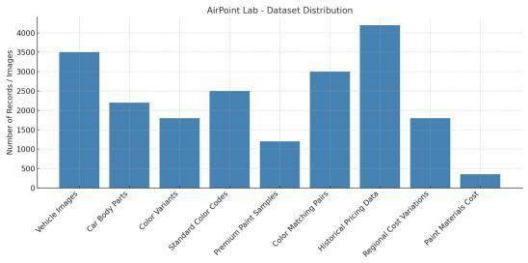


Fig 13: Dataset

C. Limitations

This technology does not yet support the creation of pearlescent surfaces.

The chatbot must rely on people inputting data by hand since it isn't fully powered by NLP.

Mobile responsiveness is not supported by some older device models.

V. CONCLUSION

AirPoint Lab successfully demonstrates the practical application of the automotive industry by combining automated customization, pricing estimation, and trust-based workshop selection. Migration to WordPress accelerated development without compromising on backend complexity. Future enhancements may include React Native apps, blockchain-based trust systems, and GANs for realistic paint textures.

Tackles real issues faced by vehicle owners in Sri Lanka.

Switching to WordPress made deployments happen faster, without giving up the PHP and MySQL complexity from the backend. CNNs made it possible to match colors and regression models to determine cost, both of which helped create results customized to users. The Agile approach meant we kept updating and improving the product over time.

This work highlights both technical excellence and the importance of user-friendly design, trust and transparency that every automotive service platform should have.

Future strategies under consideration are:

- Creating apps by using React Native,
- Blockchain-based reviews

- Simulating realistic paint texture using Generative
- Adversarial Networks (GANs)
- AirPoint Lab is designed to fit any needs and can be used for future vehicle services development in developing markets.

ACKNOWLEDGEMENT

The author conveys deepest wishes of appreciation to the project supervisor Mr. Pubudu Nallaperuma for his inestimable advice, motivation and technical viewpoints that he provided in the process of the platform development of AirPoint Lab platform. His technical and academic guidance was invaluable in influencing the course of this research.

The authors are also grateful to the University of Bedfordshire for allowing the use of necessary technical resources and for the technical and academic support while carrying out the project. The author is thankful for the survey. The author used **TensorFlow, WordPress,** and **Google Maps API** for enabling the integration of cutting-edge technology into the research.

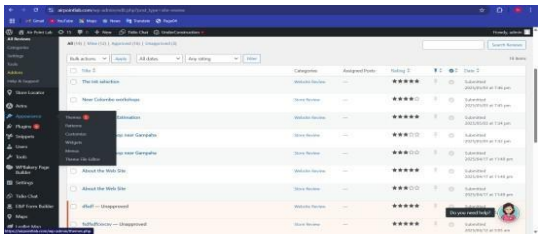


Fig 14: wp_admin panel

Fig 15: Color Cost Estimation Page

Fig 16: Google Map

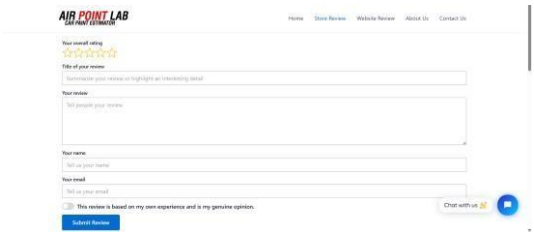


Fig 17: Review/Feedback Page

REFERENCES

[1] A. Kumar and S. Patel, "Cost estimation algorithms in automotive services," *J. Appl. Math.*, vol. 25, no. 4, pp. 120–135, 2019.

[2] M. Garcia, "Cost modeling in automotive repair and painting," *J. Ind. Eng. Manag.*, vol. 15, no. 2, pp. 77–89, 2020.

[3] M. Lee, R. Chen, and K. Zhou, "Estimating the cost of automotive paint jobs using machine learning," *IEEE Trans. Autom. Sci. Eng.*, vol. 18, no. 3, pp. 567–580, 2021.

[4] A. Burns, "Interactive design of online automotive platforms," *Web Interfaces J.*, vol. 17, no. 2, pp. 33–46, 2022.

[5] Z. Liu, T. Hu, and P. Chen, "Blockchain-enhanced trust for automotive platforms," *J. Inf. Secur.*, vol. 19, no. 1, pp. 22–34, 2022.

[6] A. Cooper and R. Reimann, *About Face: The Essentials of Interaction Design*, 5th ed. Hoboken, NJ, USA: Wiley, 2021.

[7] D. A. Norman, *The Design of Everyday Things*, rev. ed. New York, NY, USA: Basic Books, 2022.

[8] H. Zhang and L. Tan, "Photorealistic vehicle color visualization using deep learning," *J. AI Transp.*, vol. 11, no. 3, pp. 45–52, 2020

[9] R. Gupta, "Trustworthiness of online automotive platforms in developing economies," *J. Emerg. Mark. Res.*, vol. 19, no. 2, pp. 45–60, 2021.

[10] T. Nguyen and C. Roberts, "Review systems and decision-making in automotive platforms," *Int. J. Consum. Anal.*, vol. 23, no. 1, pp. 12–27, 2022

Implementation of Wyltl: An Imperative Language with a Dual Interpreter – Compiler Architecture

Dineth Mallawarachchi
School of Computer Science and Technology,
University of Bedfordshire,
lochandineth@gmail.com

Yasas Jayaweera
Dean of Academic Affairs
SLIIT City University
Colombo 3, Colombo, SL
yasas.j@sliit.lk

***Abstract*— When using a programming language, a common drawback is the prevalence of resource constraints. A lack of resources often results in programs executing faster on high end hardware in comparison to middling or low-end hardware. While a core tenant of programming is optimization, with which entire industries have been built upon, when implementing a programming language, the process becomes significantly more complex. Minute slowdowns in a programming language implementation could very quickly result in major slowdown when executing some code. This paper examines the process of implementing the Wyltl language while balancing the need for performance and resource efficiency provided precious insight and hints towards future optimization, and the unique challenges and opportunities for evolution. Including evaluation of varying parsers, processors and technologies. The evaluation of different parsers and processors was done throughout development with a varied range of programming language optimization techniques being followed as required.**

***Keywords*— Parser Evaluation, Optimization, Programming Language Implementation, Wyltl, Interpreter, Compiler, Web Assembly, JavaScript Interoperability, Go.**

I. Introduction

Wyltl [1] is a programming language with a focus on maintaining a balance of simplicity and functionality (programming language design techniques), and portability and embeddability (programming language implementation techniques). The central design concept of Wyltl focused on the idea that balancing all traits would provide a complete and capable implementation for experienced and new developers. An emphasis on performance was considered secondary during the latter stages of the project. Looking at Wyltl with a focus on performance provides a unique perspective not commonly found in many software development situations.

II. Literature Review

When developing Wyltl a wide range of alternative systems and approaches were considered with their respective strengths and weaknesses.

A. Pratt, Packrat and Recursive Descent Parsing

Parsing is the second stage of execution within a programming language – responsible for taking a stream of tokens given by the tokenizer and combining them into a specific structure which is later processed or compiled. Parsing in modern programming languages makes use of either Pratt Parsing [2] or Packrat Parsing [3]. Out of these approaches Packrat parsing introduced by Ford is often avoided due to its relatively high memory usage. This is a waste of resources and is not efficient when handling complex expressions as what would be found when parsing programs. Pratt's parsing techniques in comparison are quite efficient in parsing complex mathematical expressions but would themselves struggle with efficiently handling complex programs – and more importantly, error handling and recovery. The solution was to investigate classical LR (left-right) recursive descent parsing more closely. Which provided fast execution speeds and low memory usage with increased flexibility, while keeping many of the improvements brought forward by a custom parsing implementation as used in common Pratt and Packrat parsing implementations.

B. Efficiency of Compiled and Interpreted Scripting Programming Language Implementations.

The difference between Compiled and Interpreted languages is a common point of concern for many users. It becomes a key decision in implementing Wyltl as compiler and interpreters are fundamentally different systems with their own advantages and drawbacks. However, the role of a scripting programming language is flexible as described by Ousterhout in his historical analysis of programming languages [4]. He correctly predicts the importance of focusing on simplicity in a manner that does not compromise on functionality. This applies to Wyltl in the way both a compiler and interpreter are present and supported. For absolute efficiency, a compiler is

undoubtedly better as programs execute faster. However, when it comes to embeddability an interpreter can offer more flexibility. The key decision was to offer a dual implementation that provided efficiency regardless of the implementation type used, while keeping the inherent performance differences between a compiler and interpreter implementation.

C. Inline Substitution

One of the most popular optimizations done by programming language implementations is inline substitution. Which substitutes expressions with simplified forms of themselves, this can take the form of replacing expressions such as '2 + 1' with '3'. The approach is outlined by Scheifler [5]. In practice, using inline substitution on complex statements can be difficult and potentially risky. Wyltl makes use of inline substitution when compiling mathematical expressions. Scheifler examines the advantages brought forward by implementing such a compilation structure – which primarily consists of improving execution time by reducing recursive function call. An approach which proves to be very useful as recursion remains a complex and heavy process which frequently bottlenecks and slows – down the execution of many programming languages, including Go and by extension Wyltl.

D. Slim Compilation

In compiling a program to a binary format, the process in which the compilation process occurs can also be considered for optimization as described by Kistler and Thomas [6]. 'Slim Compilation' in this context can contest resulted in multiple approaches – starting from removing unwanted debug data from executables to implementing compression algorithms with binary data when required. While reducing the size of compiled executables with a high improvement ratio, such compression in turn results in a need for increased processing power for decoding the compressed data. Unlike other factors, slim compilation presents a hard boundary of optimization, as it becomes impractical beyond slight space savings.

III. Uniqueness of Wyltl Implementations

Striking a balance between simplicity, portability, embeddability and functionality is considered one of the key challenges in programming language design and optimization. Opinions regarding which of the trio should be prioritized is what differentiates the syntax and semantics of many programming languages. Wyltl forgoes that by placing equal focus on all four factors. Simplicity is ensured with Wyltl's language design as shown by listing one below. All fundamental programming components are showcased with simple syntactical structures

Portability is ensured using the Go toolchain. Which by itself ensures portability as Go does not depend on any external dependencies for its cross-platform support. Although it's important to note that using Go by itself is not indicative of cross platform support. Wyltl is specifically written in a manner that it is completely platform agnostic.

Embeddability is ensured in with Wyltl's ability to be embedded within existing Go applications. This is implicit with how Wyltl's components are designed and partitioned

IV. Shared Architecture of The Wyltl Implementations

Wyltl provides a dual interpreter and compiler implementation of which both are combined to form the reference Wyltl executable. This form of distribution is itself an optimization geared towards reducing the amount of re-used code and duplicated downloads. This is apparent with the distributed file size for the desktop and web Wyltl versions, which fall around 3.00 Megabytes. Additionally, the same approach is followed when distributing Wyltl for Web Platforms, with a 3.10 Megabyte Web Assembly binary being primarily embedded within HTML pages as required. The internal structure followed by Wyltl is shown by Figure One below.

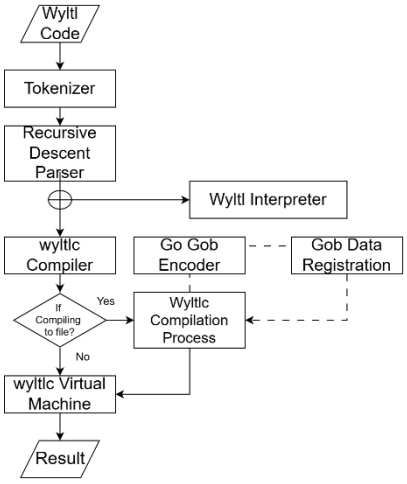


Fig 1 Shared Architecture of Wyltl Implementations

The Wyltl implementations achieve operational efficiency by using this shared architecture to create a 'base' on which the Wyltl Interpreter, and Compiler build upon. The Wyltl tokenizer and recursive descent parser are common among both implementations. This is of importance to the language as it ensures that benchmarking is accurate and is not affected by changes in different components.

The main divergence between the Wyltl Compiler and Interpreter occurs when the input Wyltl code is broken into

tokens by the tokenizer and then arranged into an abstract syntax tree. The Wyltl Interpreter can process this data directly, while the Wyltl Compiler requires compilation to Wyltlc as an intermediate step, then the generated Wyltlc binary would require to be executed by the Wyltlc virtual machine, which would effectively decompile the Wyltlc file back to an abstract syntax tree. With that perspective the Wyltl interpreter can be more effective in execution, as no binaries or intermediaries are created.

The key player is the 'gob' standard library provided by Go, which is responsible for the 'compilation', 'decomplication' and binary creation performed by the Wyltl compiler and Wyltl virtual machine respectively.

V. Evaluating the Wyltl Compiler and Interpreter using Common Algorithms

Wyltl offers the unique research opportunity to evaluate a compiler and interpreter with a shared architecture against each other. De-mystifying the implementation and performance of programming languages with the understanding of them being programs themselves. The key work in evaluating Wyltl is starts with the port of several popular algorithms.

Fig 2 showcases the difference in execution speed between the Wyltl Interpreter and Compiler. It is apparent that the primary speed difference between the compiler and interpreter is primarily visible when executing Wyltl in moderate and low-end devices. To determine the state of optimization in Wyltl, the execution speed in high-end devices must be examined. In this scenario it is visible that the difference between the Wyltl compiler and interpreter becomes narrow, with the interpreter being 8.95% slower than the compiler on average.

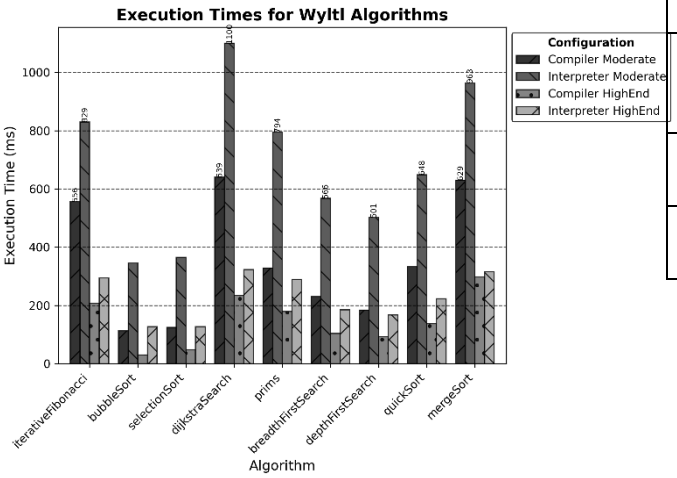


Fig 2 Algorithm Execution Speed of Wyltl Implementations

These results also showcase the dependency Wyltl has on the speed of the CPU (central processing unit) used by the device it is run on.

A somewhat common myth is that a program would use less memory in a faster device due to application executing faster. However, the test-data collected by Wyltl disproves it, by showing that the memory usage in moderate and high-end devices is the same. This consistency in memory usage is caused by the explicit nature of how Wyltl code is handled. Of importance is how the data goes against the concept of memory usage and execution speed being strictly tied together in programming languages.

To accurately gauge the nature of the connection between execution speed and memory usage in Wyltl, the data must be examined with nuance using Go's benchmarking capabilities.

Table XII Performance Metrics of Wyltl Implementations

Name of Test	Compiler Speed (µs)	Interpreter Speed (µs)	Interpreter Memory Usage (KB)	Compiler Memory Usage (KB)
Iterative-Fibonacci	555.33	952.26	91.57	98.75
Bubble-Sort	97.60	374.46	3.00	9.85
Selection-Sort	121.48	370.69	3.09	10.17
Dijikstra-Search	531.23	939.29	3.09	10.17
Prims-Search	327.99	794.49	10.09	29.22
Breadth-First-Search	230.63	566.80	8.48	22.82
Depth-First-Search	184.18	501.90	7.76	20.36
Quick-Sort	332.93	648.64	14.08	26.07
Merge-Sort	629.57	963.25	29.13	41.23

The data shown in Table 1 was primarily evaluated after visualization as shown in Figure 4 below.

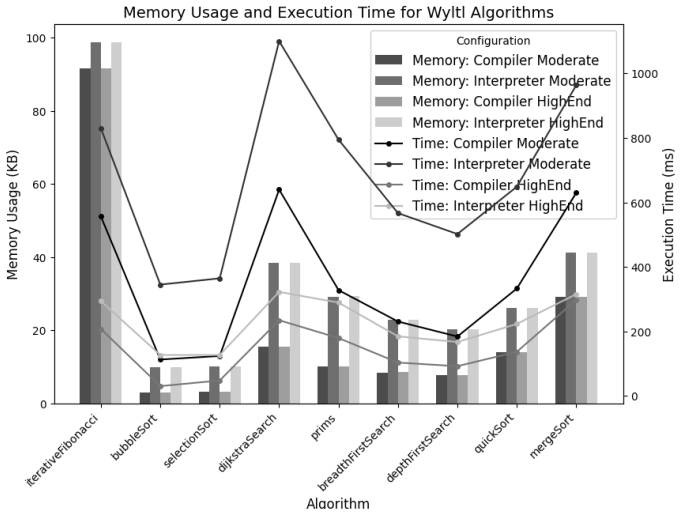


Fig 3 Algorithm Memory Usage of Wyltl Implementations

The primary findings from the visualization of Wyltl performance data is concerned with showcasing the link between the memory allocation count and the execution speed. In a nutshell, the speed of the Wyltl Implementations is loosely tied to the memory usage. However, it is important to note that there is no connection between the memory usage and execution speed across the implementations of different algorithms in Wyltl. This can be further seen when comparing the following calculations of Pearson's coefficient.

Fig 4 Pearson's Coefficient of Wyltl Implementation Memory and Execution Speed

Compiler Memory Allocation Count Vs Compiler Speed (High-End) Slope = 0.1441, Intercept = 37.8236 Pearson r = 0.8957, p-value = 0.0011

Interpreter Memory Allocation Count Vs Interpreter Speed (High-End) Slope = 0.1231, Intercept = 75.6166 Pearson r = 0.9594, p-value = 0.0000

Compiler Memory Allocation Count Vs Compiler Speed (Moderate) Slope = 0.0404, Intercept = -10.4956 Pearson r = 0.8024, p-value = 0.00

```
printFunc := js.FuncOf(func(this js.Value, args []js.Value) interface{} {
        output := args[0].String()
        js.Global().Get("document").Call("getElementById",
        "output").Set("value",
        js.Global().Get("document").Call("getElementById",
        "output").Get("value").String()+output+"\n")
        return nil
})
js.Global().Set("goPrint", printFunc)
```

Fig 5 Overwriting Pre-Defined Web Assembly System Calls with Custom Functions.

Interpreter Memory Allocation Count Vs Interpreter Speed (Moderate) Slope = 0.0325, Intercept = -6.8547 Pearson r = 0.7658, p-value = 0.0161

An important observation to be noted in relation to the Wyltl is its effectiveness in handling processor intensive programs such as Fibonacci and Dijkstra Search. Additionally, while Wyltl memory usage is not affected by the efficacy of the hardware it is run on – the execution speed of the implementations is linked together with hardware. Additionally, the execution of simple programs such as the fundamental sorts are more effective on all hardware. There are algorithms such as iterative Fibonacci and the graphing algorithms– which tend towards comparatively high memory usage as they focus quite heavily on recursion. Recursion is very CPU heavy on Go and Wyltl by extension.

```
wrappedCode := fmt.Sprintf(`
 (function() {
   try {
        return eval(%q);
   } catch (e) {
        throw new Error(e.message);
   }
  })()`, code.Value)
jsValue := js.Global().Get("eval").Call("call", js.Global(), wrappedCode)
```

VI. Role of web-assembly within Wyltl implementations

While programming languages usually support a variety of platforms and systems without issues, the web is usually ignored or left behind. This is a major loss considering the prevalence of web sites and web development in general. Wyltl supports the Web as a first-class platform through the compilation target known as Web Assembly [7]. The central challenge in this approach is that Wyltl is first and foremost a text user interface application, which running on the windows command line or Unix terminal.

The method in this challenge was overcome is by overwriting the default Web Assembly functionality defined by the Go Compiler. This functionality allows Wyltl to print to a defined area of a web page without having to depend on extensive development debts or functions.

VII. JavaScript interoperability in Wyltl

Out of the few languages that support seamless execution on Web and Desktop platforms, through Web Assembly or other platforms, none supports interoperability with the JavaScript runtimes present in Web Browsers. This is often a major detriment considering that JavaScript reigns as the de-facto language of the web besides HTML. In the case of Wyltl, JavaScript interoperability is achieved by using web assembly as a 'glue' in the sense of using it as a middleman between the Go-Runtime and the JavaScript engine used by the browser. This allows the Web version of Wyltl to execute any JavaScript code without problems, as compatibility is guaranteed by the Web Browser's own JavaScript runtime. The disadvantage of this process here is the dependency on the JavaScript runtime. This is the only feature that differs among the different versions of Wyltl as the desktop version will not execute such code.

The process of calling the Web Browser's JavaScript runtime to execute code is shown by the code snippet in Fig 6 This code is selectively compiled exclusively for the Web releases of Wyltl. A secondary feature of the JavaScript interoperability present within Wyltl is the conversion of datatypes between Wyltl and JavaScript. This is necessary as even though Wyltl and JavaScript are compiled languages, the way they map and deal with datatypes are somewhat different. However, since the fundamental data types remain the same regardless of the programming language, conversion is mostly a straightforward process.

```
switch jsValue.Type() {
  case js.TypeString:
     return &String{Value: jsValue.String()}
  case js.TypeNumber:
     floatVal := jsValue.Float()
     ,
     return &Float{Value: floatVal}
  case js.TypeBoolean:
     return &Boolean{Value: jsValue.Bool()}
   case js.TypeNull, js.TypeUndefined:
     return nil
  default:
     return &String{Value: jsValue.String()}
}
```

Fig 6 Passing JavaScript code to the Web Browser's JavaScript *runtime*

The simplicity of this implementation is evident by the code snippet presented in figure 7. Wyltl's type system through simplified is feature rich and capable of supporting all the basic types used by JavaScript. The impact of this is twofold. Primarily, data can be transferred without having to depend on 'string' as the primary communication type. Secondly, Wyltl's null safety continues to be insured through this process, which discards any null data returned by JavaScript. However, there is still an expectation that all JavaScript code written by users will contain the necessary safety checks and balances.

VIII. Wyltl and Wyltlc

While the Wyltl Interpreter and Compiler accepts ". Wyltl" files as inputs. The way they are used is fundamentally different. The Wyltl Interpreter works by directly executing the file. This is often slow as the tokenizer and parser would have to execute before the actual interpretation begins. The Wyltl Compiler takes the input Wyltl file and splits it into instructions using the Wyltl tokenizer and parser. Then the compiler registers the types, and statements used within the program using the 'GOB' Go encoder. Then these data is saved to a file with the label '. Wyltlc'

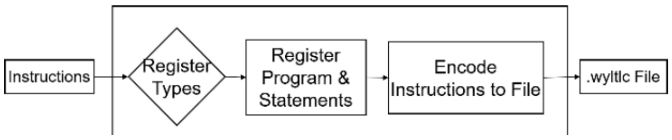


Fig 8 Compilation of Wyltl to Wyltlc

The actual execution of Wyltlc cannot occur directly as the Wyltl virtual machine needs to de-serialize the file and process it instruction & instruction. Whenever an instruction regarding data modification is found, then it as added / popped from the Wyltl Virtual Machine Stack as shown by figure 9 below.

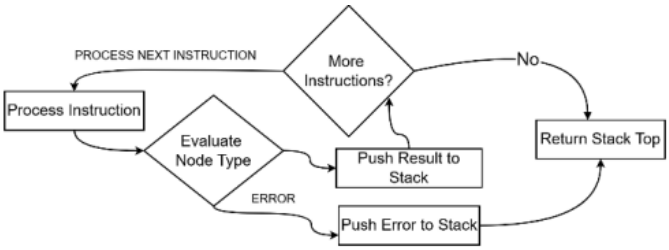


Fig 9 Execution of Wyltlc within the Wyltl Virtual Machine

Processing data through a stack is significantly faster and this is a major contribution towards Wyltl compiler's increased execution speed.

IX. Long Term Maintainability of Wyltl

Many see programming languages as a form of long-term skill investment. Developers expect a language to serve them well for many years or decades to come. One of the main ways that Wyltl ensures this is platform support. Wyltl currently supports Window, Macintosh, Linux and Web as primary platforms. With BSD and bare-metal programming being possible, yet untested platforms. Even in these platforms Wyltl strives to support a range of diverse systems. For instance, the windows release of Wyltl is backwards compatible with Windows 7, in a time where even the Go compiler has dropped support for it. The promise of the Wyltl language is clear and consistent support for all mainstream platforms (and some niche platforms such as BSD).

Wyltl's license has also been the subject of some discussion. A closed source / proprietary approach might bring some initial advantages, but they will all be overshadowed by long term problems – such as contribution. As a programming language grows a collaborative development approach becomes necessary. To ensure that Wyltl is not ill-prepared for such a future – from the main release of Wyltl, it is first and foremost an open-source project licensed under the GNU public license which ensures that all forks and derivatives of Wyltl also remain free and open source so that they can trickle back upstream. As such users can be assured of Wyltl's future.

X. Turing Completeness of Wyltl

A fundamental question asked about any programming language is its completeness. The main measure of a programming language is what applications can be written using it. The main measure of which is algorithm implementation. To this end, many classical algorithms have been implemented in Wyltl. A non-exhaustive list of which is as follows.

1. Dijikstra Search Algorithm
2. Depth First Search Algorithm
3. Breadth First Search Algorithm
4. Knuth Morris Pratt Algorithm
5. Bellman Ford Algorithm
6. Iterative Fibonacci Algorithm
7. Quick Sort Algorithm
8. Merge Sort Algorithm
9. Prims Algorithm
10. Bubble Sort Algorithm
11. Selection Sort Algorithm
12. Insertion Sort Algorithm

The Wyltl implementation of these fundamental algorithms can be interacted with in the Wyltl Playground (https://dineth-lochana.github.io/Wyltl/practice.html). These are 'pure Wyltl' applications in the sense that they do not make use of any external functions brought over from Wyltl's JS language interoperability. The ability of Wyltl to implement these 'Turing complete' functions effectively showcase that Wyltl is a Turing Complete language itself. Wyltl's syntax is flexible enough to allow for reasonably complex constructs such as 2D or 3D arrays along with imperative & functional programming flows. This is shown by the most complex algorithm implemented in Wyltl which is Dijikstra's path finding algorithm. Focusing on the opposite, that is simplicity. An example of a Turing Complete Wyltl Program that also showcases all the fundamental programming traits of – sequence, selection (execution of programming), selection (conditional execution of programs), repetition (re-execution of code) is the 'greeter' program that is showcased by listing 2.

```
suppose x is 1.
suppose y is inputNumber("How many Loops?").
suppose name is inputText("Enter name").

if (name equals "Dineth") {
  print("Hey I'm Dineth too!")
} else {
  while(x below y) {
    print("Hello " plus name)
    suppose x is x plus 1.
  }
}

suppose wellWisher is compose(input){
  print("Have a nice day " plus input)
}
wellWisher(name).

suppose nameArray is stringToArray(name).
for(suppose count is 0 : count below length(name) : suppose count is
count plus 1) {
   print(nameArray[count])
}
```

The greeter program can be executed at the Wyltl Web Playground of which the URL is as follows (https://dineth-lochana.github.io/Wyltl/practice.html?file=greeter.wyltl).

XI. Conclusions

The research insight provided by Wyltl helps to determine the difference between a compiler and interpreter built on a stable and shared base. The performance and memory usage comparisons provided a new perspective on how memory usage cannot be used to accurately predict execution speed. This is a novel finding specific to Wyltl and is of interest to both users of Wyltl and developers using the Go programming language to implement applications or programs. A wider range of testing across the 'unique' platforms supported by Wyltl such as WASM (Web Assembly), would be more useful, however these are not possible currently due to the inaccurate nature of the JavaScript Runtimes of Web Assembly present in popular web browsers.

ACKNOWLEDGMENT

The author would like to thank Dr. Yasas Jayaweera for his invaluable guidance and advice throughout the duration of the project, along with the University of Bedfordshire and SLIIT City Uni for providing the research opportunity for Wyltl.

REFERENCES

[1] Mallawarachchi, D. and Jayaweera, Y. (2025) 'Design and implementation of WYLTL: An imperative, embeddable and portable programming language', International Journal of Computer Applications, 186(76), pp. 22–29. doi:10.5120/ijca2025924669.

[2] Pratt, V.R. (1973) 'Top-down operator precedence', Proceedings of the 1st annual ACM SIGACT-SIGPLAN symposium on Principles of programming languages - POPL '73, pp. 41–51. doi:10.1145/512927. 512931.

[3] Ford, B. (2002) 'Packrat parsing: Simple, Powerful, Lazy, Linear Time', Proceedings of the seventh ACM SIGPLAN international conference on Functional programming, pp. 36–47. doi:10.1145/581478.581483

[4] Ousterhout, J.K., 1998. Scripting: Higher level programming for the 21st century. Computer, 31(3).

[5] Scheifler, R.W. (1977) 'An analysis of inline substitution for a structured programming language', Communications of the ACM, 20(9), pp. 647–654. doi:10.1145/359810.359830.

[6] Franz, M. and Kistler, T., 1997. Slim binaries. Communications of the ACM, 40(12), pp.87-94.

[7] MozDevNet (no date) WebAssembly concepts , MDN. Available at: https://developer.mozilla.org/en-US/docs/WebAssembly/Guides/Concepts (Accessed: 20 May 2025).

Explainable AI Powered Mental Health State Capturing Application to Support Students' Mental Wellness and Academic Stress Mitigation

Janidu Himansa Welarathna
Bachelor of Software Engineering
University of Bedfordshire
Luton, United Kingdom
Janidu.Welarathna@study.beds.ac.uk

Pubudu Nallaperunma
SLIIT CITY UNI, Colombo
pubudu.n@sliit.lk

***Abstract*— Mental health is a state of well-being that enables individuals to manage stress, work effectively, and contribute to society. However, reports show that serious mental health problems among students worldwide are increasing rapidly. A critical problem is that students often fail to recognize mental health issues or the sources of their academic stress, leading to silent suffering that escalates over time. A significant research gap exists as current assessments methods lack the ability to identify root causes of academic stress and provide explainable decisions for clinical use. This significant rise in many students' mental health issues have indeed opened important discussions about its underlying causes, consequences, and the need for a comprehensive support system. Voices are an important part for identifying emotional expressions, as speech is the most vital channel of communication, enriched with emotions. The system analyzes emotional patterns in students' voices using Natural Language Processing (NLP) techniques to identify eight emotions and reveal the root causes of their mental health challenges and academic or non-academic stress. Additionally, Explainable AI (XAI) techniques are employed to provide a comprehensive analysis of these patterns, enhancing understanding and supporting managerial decision-making. The system achieves 93.46% accuracy using Random Forest algorithm with reliable confidence levels for clinical applications. It operates effectively in uncontrolled environments with language-independent features, ensuring adaptability across diverse student populations. While students typically seek support from counselors and healthcare professionals who base their decisions on clinical experience, this system offers an additional diagnostic tool to complement and validate professional evaluations. This research aims to better understand student mental health issues and contribute to improved students' wellness and academic success.**

***Keywords*— speech emotion recognition, explainable AI, mental health, student wellness, machine learning**

I. INTRODUCTION

Student mental health has become a critical concern in academic environments worldwide. According to the World Health Organization, 1 in 7 students aged 10-19 experience mental health conditions, with depression and anxiety being most common [1]. Recent studies show that 76% of students experienced serious psychological distress in the past year, while 92.4% report concentration difficulties due to stress [2]. Current mental health assessment methods rely heavily on subjective reporting and clinical observations. Healthcare professionals often struggle to identify early warning signs before conditions worsen. This research addresses these limitations by developing an AI-powered system that analyzes speech patterns to detect emotional states and potential mental health concerns. The proposed system uses speech emotion recognition combined with Explainable AI (XAI) techniques to provide transparent, interpretable assessments. This approach offers healthcare professionals a valuable second opinion tool while maintaining clinical decision-making autonomy. The system focuses specifically on student populations, addressing the unique stressors and challenges they face in academic settings.

II. LITERATURE REVIEW

This review examines current research to support developing an AI system that identifies student emotions and mental health states using machine learning techniques. The study addresses limitations in understanding student academic stress and gaps in applying NLP and XAI methods for mental health detection among student populations.

A. Speech Emotion Recognition in Mental Health Detection Using Machine Learning.

Traditional machine learning methods have shown good results in detecting emotions from speech for mental health purposes. Rastogi et al. [3] used Multi-Layer Perceptron classifiers and achieved 75% accuracy in recognizing emotions like anger and happiness, showing how voice patterns can connect to mental states. Madanian et al. [4] built a system to help mental health doctors during remote therapy sessions, where they can't see body language. Their Support Vector Machine approach reached 74% accuracy on standard datasets and up to 89% when combining different data sources. Shahin et al. [5] improved results by

using a smart feature selection method with Grey Wolf Optimizer and K-Nearest Neighbors, achieving 89% accuracy for specific accents and over 80% for common datasets, though the method was too slow for real-time use. Ancilin & Milton [6] developed faster audio features that performed better than traditional methods, reaching 81% accuracy while taking less time to process, though distinguishing between similar emotions like fear and disgust remained challenging.

B. Speech Emotion Recognition in Mental Health Detection Using Deep Learning

Deep learning techniques have performed much better than traditional machine learning approaches in recognizing emotions from speech. Abdullah et al. [7] reviewed various deep learning models and found that LSTM networks could reach up to 95% accuracy when analyzing body signals, while CNN models achieved around 72% accuracy. Mohammed et al. [8] focused specifically on mental health applications and used CNN models with audio features to identify negative emotions linked to mental health problems, achieving 94% accuracy. Tariq et al. [9] created an advanced system combining Deep Stride CNN with Bi-Directional LSTM that used multiple types of audio features, reaching 95.5% accuracy and improving previous methods by almost 20%. Liu et al. [10] developed a system that worked across different speakers by combining CNN with attention-based LSTM, achieving around 70% accuracy, though they found that the system worked better with male voices than female voices. Elsayed et al. [11] built a model for virtual health assistants using gated RNN and one-dimensional CNN, which achieved 94% accuracy and significantly outperformed older methods like Support Vector Machines.

C. Speech Emotion Recognition in Mental Health Detection Using XAI

Explainable AI methods have become important in mental health applications because doctors need to understand how these systems make decisions. Destiny [12] used multiple data sources including activity levels and stress scores to predict stress levels, achieving 73% accuracy with Random Forest and 72% with Gradient Boosting, with analysis showing that stress scores were the most important factor. Pendyala & Kim [13] tested several machine learning models and achieved 85% accuracy with Gradient Boosting, but their explainability analysis revealed that the models were making decisions based on less important features, highlighting why transparent AI is crucial for trustworthy mental health systems. Kim & Kwak [14] improved reliability by combining different audio analysis models and achieved 87% accuracy while providing clear explanations of how the system made its decisions using various visualization techniques. Nfssi et al. [15] developed a framework that focused on selecting the best features and explaining decisions, testing 14 different models and achieving up to 99% accuracy on some datasets, with clear explanations of which features contributed most to the predictions.

Table 13: Existing Systems VS Proposed System

Research paper	Audio pre-process	ML Model	DL Model	XAI Used	Mental Health Root cause	Risk Factor Mapping
Rastogi et al. (2023)	✓	✓	✗	✗	✗	✗
Madanian et al. (2023)	✓	✓	✗	✗	✗	✗
Shahin et al. (2023)	✓	✓	✗	✗	✗	✗
Ancilin & Milton (2021)	✓	✓	✗	✗	✗	✗
Mohammed et al. (2025)	✓	✗	✓	✗	✗	✗
Tariq et al. (2025)	✓	✗	✓	✗	✗	✗
Kim & Kwak (2024)	✓	✓	✓	✓	✗	✗
Pendyala & Kim (2024)	✗	✓	✗	✓	✗	✗
Proposed System	✓	✓	✓	✓	✓	✓

III. METHODOLOGY

The research uses datasets (RAVDESS and TESS) to train machine learning models that identify eight emotions from speech patterns through acoustic feature analysis. The system employs explainable AI methods to provide clear reasoning for predictions and detect emotions to academic stress causes for healthcare professionals.

Figure 1 presents the proposed system architecture with three main tiers. The client application tier handles audio input and displays results. The backend services tier processes speech data and runs machine learning models. The database tier stores audio files and analysis results securely. Mental health professionals use the web application to monitor student emotions and make clinical decisions. The system uses HTTPS communication to ensure secure data transfer between all tiers.

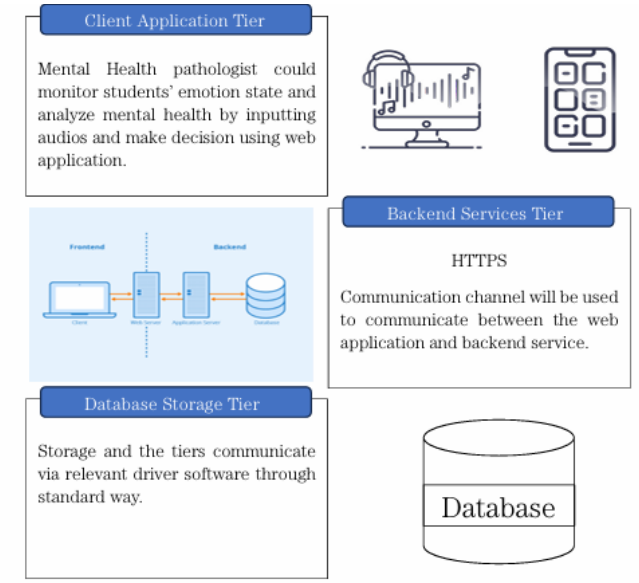


Fig 18 : System Overview Architecture

The software solution implements multi-stage processing to analyze student speech and identify emotional patterns.

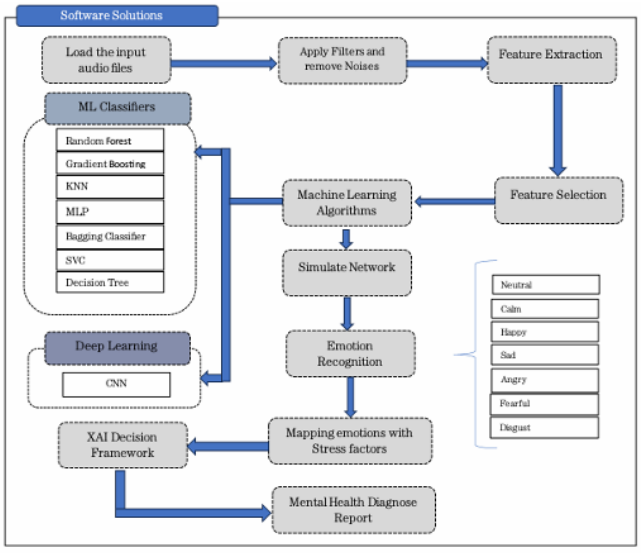


Fig 19 : Software Solution Architecture

Fig2 presents the complete software solution workflow. The process begins with loading input audio files, followed by applying filters to remove background noise and unwanted disturbances. The cleaned audio then undergoes feature extraction to identify key acoustic characteristics.

The system employs multiple machine learning classifiers including Random Forest, Gradient Boosting, K-Nearest Neighbors, Multi-Layer Perceptron, Bagging Classifier, Support Vector Classifier, and Decision Tree algorithms. Feature selection techniques optimize the most relevant characteristics for emotion classification.

Stage 1: Speech Emotion Recognition

- Audio preprocessing and noise filtering
- Feature extraction using MFCC, Chromagram, MEL Spectrogram, and Spectral Contrast
- Machine learning classification achieving 93.46% accuracy

Stage 2: Root Cause Analysis

- Stress factor mapping techniques analyzing student responses
- Classification into academic, non-academic, or combined stress factors using ranked stress events
- Integration with emotion data for comprehensive assessment

Stage 3: Explainable AI Analysis

- LIME and SHAP analysis for decision transparency
- Generation of clinical reports with clear explanations
- Second opinion for healthcare professionals

If traditional ML approaches prove insufficient, the workflow transitions to deep learning using CNN architecture. The emotion recognition module processes the classified results and filters negative emotions for further analysis. The XAI Decision Framework then maps these emotions to specific stress factors, ultimately generating a comprehensive mental health diagnostic report for healthcare professionals.

The training process employs a multi-source dataset combining established repositories with custom recordings for comprehensive emotion recognition.

RAVDESS Dataset the Ryerson Audio-Visual Database provides professionally recorded emotional speech in English. This collection features vocal expressions from 24 actors with North American accents, containing 7,356 recordings across eight emotions: neutral, calm, happy, sad, angry, fearful, disgust, and surprised. Audio specifications include 48 kHz sampling rate with standardized statement variations.

TESS Dataset the Toronto Emotional Speech Set contains validated voice samples with 200 vocabulary targets. This dataset includes 2,800 recordings covering seven emotions in English language, maintaining 24.414 kHz sampling rate in mono-channel format.

Custom Dataset A specialized collection featuring emotional expressions was developed with 1,232 audio samples from 14 speakers. Recordings capture natural emotional states in conversational contexts.

The combined dataset uses systematic file naming with seven-segment numerical identifiers. Emotion classification employs two-digit coding: 01-neutral, 02-

calm, 03-happy, 04-sad, 05-angry, 06-fearful, 07-disgust, 08-surprised. This standardized system enables consistent emotion identification during model training.

Audio preprocessing includes noise filtering using high-pass filters and signal normalization. The nlpaug library's NoiseAug function removes background noise while preserving important speech characteristics. This step is crucial for distinguishing between similar emotional states like "neutral" and "calm."

Feature Extraction: Multiple acoustic features are extracted using the librosa library

MFCC (Mel-Frequency Cepstral Coefficients): Captures spectral characteristics of speech

Chromagram: Represents pitch class information across musical octaves

MEL Spectrogram: Provides time-frequency representation of audio signals

Spectral Contrast: Measures difference between spectral peaks and valleys

These features capture different aspects of emotional expression in speech, providing comprehensive representation for classification algorithms. the system implements a hierarchical classification strategy. Seven traditional ML algorithms are evaluated first: Random Forest Classifier, Gradient Boosting Classifier, K-Nearest Neighbors, Multi-Layer Perceptron, Bagging Classifier, Support Vector Classifier, and Decision Tree Classifier.

If traditional ML approaches prove insufficient, the system employs Convolutional Neural Networks. This adaptive approach balances computational efficiency with classification performance.

Table 2: Ranked Stress Events

I am afraid to speak or discuss in the lecture room
I feel academic programme is too cumbersome for me
I have trouble making up my mind about my academic work
I feel worried about coping with my studies
I feel some lecturers are too hard for me to understand
Some courses are too dull and boring
I have difficulty in eating
I am not really sure am interested in reading
I have trouble studying effectively

Detected emotions are mapped to specific mental health risk factors based on academic stress research using stress factor mapping techniques. When students speak, the proposed system simultaneously identifies their emotional state and analyzes their spoken words. These words are then mapped against ranked stress events to determine the specific stressors affecting the student. Based on this analysis, the system categorizes the student's condition as academic mental health issues, non-academic mental health issues, or a combination of both. The proposed system identifies common academic stressors such as "fear of speaking in class" and "academic workload pressure" while also recognizing non-academic factors. This dual-analysis approach connects emotional patterns with their underlying psychological causes through systematic word-to-stress-factor mapping.

The proposed system utilizes Python 3.8+ with scikit-learn for traditional machine learning algorithms (Random Forest, SVM, KNN) and TensorFlow 2.x for deep learning implementation. Data processing employs pandas and numpy libraries for manipulation and numerical computations. Audio processing capabilities include librosa for feature extraction and signal processing, sound file for file operations, and nlpaug for noise removal and data augmentation. Explainable AI functionality integrates LIME and SHAP libraries with matplotlib and seaborn for visualization support. The system architecture uses Flask for backend API development and React with JavaScript and CSS for frontend interface design, providing healthcare professionals with accessible mental health assessment tools.

Two XAI techniques provide model transparency:

LIME (Local Interpretable Model-agnostic Explanations): Explains individual predictions by identifying which speech features contribute most to specific classifications. This helps clinicians understand why the system flagged emotional states.

SHAP (SHapley Additive exPlanations): Quantifies each feature's contribution using game theory principles. SHAP provides consistent, fair attribution of how speech characteristics influence mental health assessments.

Ethical Considerations

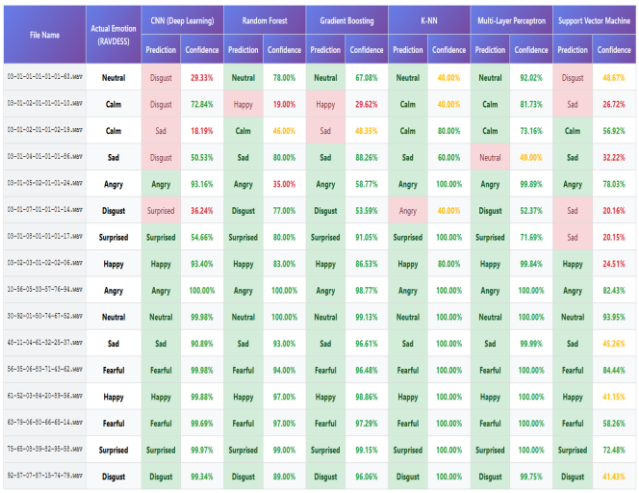
This system is designed exclusively as a second opinion tool for medical professionals in clinical settings. The proposed system prioritizes student privacy through implementation of high-security protocols and highly encrypted database systems for audio data storage. All data transmission utilizes advanced encryption standards, with role-based access control ensuring only authorized medical professionals can access patient information. The system employs automatic data anonymization processes and secure deletion protocols to maintain confidentiality standards required in healthcare environments.

IV. Results and Evaluation

This section presents the performance results of the proposed speech emotion recognition system and validates the need for automated mental health assessment tools. The

evaluation includes algorithm comparison, accuracy analysis, and market research findings from student surveys.

A comprehensive survey of 93 students validated the need for automated mental health assessment tools. Key findings include: 89.2% of students feel overwhelmed by academic workload, 94.6% experience anxiety about academic performance, 78.3% report physical symptoms of stress, and 58.7% believe universities provide insufficient mental health support.



File Name	Actual Emotion (RAVDESS)	CNN (Deep Learning)		Random Forest		Gradient Boosting		K-NN		Multi-Layer Perceptron		Support Vector Machine	
		Prediction	Confidence	Prediction	Confidence	Prediction	Confidence	Prediction	Confidence	Prediction	Confidence	Prediction	Confidence
03-01-01-01-01-01-03.wav	Neutral	Disgust	29.33%	Neutral	78.00%	Neutral	67.08%	Neutral	40.00%	Neutral	92.02%	Disgust	48.67%
03-01-02-01-01-01-13.wav	Calm	Disgust	72.84%	Happy	19.00%	Happy	29.62%	Calm	40.00%	Calm	81.73%	Sad	26.72%
03-01-02-01-01-02-19.wav	Calm	Sad	18.19%	Calm	46.00%	Sad	48.35%	Calm	80.00%	Calm	73.16%	Calm	56.92%
03-01-04-01-01-01-06.wav	Sad	Disgust	50.53%	Sad	80.00%	Sad	88.26%	Sad	60.00%	Neutral	49.00%	Sad	32.22%
03-01-05-02-01-01-24.wav	Angry	Angry	93.16%	Angry	35.00%	Angry	58.77%	Angry	100.00%	Angry	99.89%	Angry	78.03%
03-01-07-01-01-01-14.wav	Disgust	Surprised	36.24%	Disgust	77.00%	Disgust	53.59%	Angry	40.00%	Disgust	52.37%	Sad	20.16%
03-01-08-01-01-01-17.wav	Surprised	Surprised	54.66%	Surprised	80.00%	Surprised	91.05%	Surprised	100.00%	Surprised	71.69%	Sad	20.15%
03-02-03-01-02-02-06.wav	Happy	Happy	93.40%	Happy	83.00%	Happy	86.53%	Happy	80.00%	Happy	99.84%	Happy	24.51%
20-56-05-03-07-76-94.wav	Angry	Angry	100.00%	Angry	100.00%	Angry	98.77%	Angry	100.00%	Angry	100.00%	Angry	82.43%
30-92-01-03-74-67-52.wav	Neutral	Neutral	99.98%	Neutral	100.00%	Neutral	99.13%	Neutral	100.00%	Neutral	100.00%	Neutral	93.95%
40-11-04-01-02-03-37.wav	Sad	Sad	90.89%	Sad	93.00%	Sad	96.61%	Sad	100.00%	Sad	99.99%	Sad	45.26%
56-35-06-03-71-43-62.wav	Fearful	Fearful	99.98%	Fearful	94.00%	Fearful	96.48%	Fearful	100.00%	Fearful	100.00%	Fearful	84.44%
61-52-03-04-20-09-96.wav	Happy	Happy	99.88%	Happy	97.00%	Happy	98.86%	Happy	100.00%	Happy	100.00%	Happy	41.15%
63-78-06-03-66-03-14.wav	Fearful	Fearful	99.69%	Fearful	97.00%	Fearful	97.29%	Fearful	100.00%	Fearful	100.00%	Fearful	58.26%
70-03-08-09-02-05-03.wav	Surprised	Surprised	99.97%	Surprised	99.00%	Surprised	99.15%	Surprised	100.00%	Surprised	100.00%	Surprised	72.48%
92-07-07-05-74-78.wav	Disgust	Disgust	99.34%	Disgust	89.00%	Disgust	96.06%	Disgust	100.00%	Disgust	99.75%	Disgust	41.43%

Fig 3: Algorithm Performance

Training was conducted using RAVDESS dataset across multiple machine learning algorithms shows in figure 3. The evaluation measured overall model accuracy during the training process and individual prediction confidence levels.

Random Forest achieved the highest training accuracy at 93.46%. The model showed strong performance across different emotion categories with balanced precision and recall scores. The algorithm demonstrated consistent confidence levels with reliable prediction scores.

Gradient Boosting scored 92.16% training accuracy with excellent performance. The algorithm demonstrated steady learning patterns and maintained superior generalization across emotional

states. Multi-Layer Perceptron reached 88.32% training accuracy and demonstrated high confidence levels on individual predictions. The neural network approach showed strong performance with consistent accuracy across test samples. K-Nearest Neighbors achieved 84.12% training accuracy. The algorithm showed good performance with reasonable confidence levels for emotion classification tasks.

Table 14: Classification Report for Each Emotion

Emotion	Precision	Recall	F1 Score	Support
Neutral	0.96	0.94	0.95	147
Calm	0.91	0.93	0.92	94
Happy	0.94	0.96	0.95	194
Sad	0.95	0.94	0.95	195
Angry	0.97	0.96	0.97	195
Fearful	0.89	0.91	0.90	195
Disgust	0.94	0.95	0.95	148
Surprised	0.93	0.95	0.94	148

After training on the combined RAVDESS, TESS, and custom datasets, the Random Forest algorithm achieved the highest accuracy of 93.46% among all tested machine learning models. The model demonstrated exceptional performance across all eight emotional categories, with particularly strong results in angry emotion recognition (97% precision, 96% recall) and consistently high F1-scores ranging from 0.90 to 0.97. The balanced performance metrics indicate robust emotion classification capability, with macro-averaged precision, recall, and F1-scores all reaching 0.94, confirming the model's reliability for clinical mental health assessment applications.

V. Conclusion

This study developed a complete system to detect mental health problems in students by analyzing voice patterns. Student mental health is getting worse worldwide. There is an urgent need for better ways to spot problems early before they become serious. The system uses speech emotion recognition to identify eight different emotions in student voices. It then maps these emotions to specific stress factors like fear of speaking in class or feeling overwhelmed by schoolwork. The system achieved 93.46% accuracy using Random Forest methods. This makes it reliable enough for healthcare use. Results showed that traditional machine learning works better than deep learning for this task. Confidence levels do not always match accuracy. This is an important finding for clinical applications. A survey of 93 students confirmed the real need for this technology. Over 90% reported stress problems. Nearly 60% said universities

do not provide enough mental health support. The system gives doctors and counselors a helpful second opinion when working with students. It does not replace human judgment. Instead, it provides clear explanations about emotional patterns and potential causes of student stress. This helps healthcare workers make better decisions while keeping their professional authority. This technology could change how mental health problems are identified in students. Looking ahead, several development priorities will enhance the system's effectiveness and reach. The immediate priorities include expanding validation to encompass over 1500 real student samples, integrating additional languages beyond English, and developing a mobile application to facilitate easier clinical deployment. The long-term research directions focus on investigating multimodal fusion with facial expression analysis, developing personalized stress intervention recommendations, and creating longitudinal tracking capabilities for treatment progress monitoring. These advancements will strengthen the system's clinical utility and broaden its impact on student mental health support worldwide.

ACKNOWLEDGEMENT

The author thanks Mr. Pubudu Nallaperuma for supervision and guidance throughout this research project. Special appreciation to the University of Bedfordshire for providing research facilities and academic support.

REFERENCES

[1] "World Health Organization." [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health

[2] "Student Stress Statistics [2024 Update]." Accessed: Apr. 15, 2025. [Online]. Available: https://transformingeducation.org/student-stress-statistics/

[3] R. Rastogi, T. Anand, S. K. Sharma, and S. Panwar, "Emotion Detection via Voice and Speech Recognition," vol. 13, no. 1, p. 1, Nov. 2023, doi: 10.4018/ijcbpl.333473.

[4] Madanian, S., Parry, D., Adeleye, O., Poellabauer, C., Mirza, F., Mathew, S. and Schneider, S. (2023) 'Automatic Speech Emotion Recognition Using Machine Learning: Mental Health Use Case', IEEE Transactions.

[5] Shahin, I., Alomari, O.A., Nassif, A.B., Afyouni, I., Hashem, I.A. and Elnagar, A. (2023) 'An efficient feature selection method for arabic and english speech emotion recognition using Grey Wolf Optimizer', Applied Acoustics, 205, p. 109279.

[6] Ancilin, J. and Milton, A. (2021) 'Improved speech emotion recognition with Mel frequency magnitude coefficient', Applied Acoustics, 179, p. 108046.

[7] Abdullah, S.M.S., Ameen, S.Y.A., Sadeeq, M.A.M. and Zeebaree, S. (2021) 'Multimodal Emotion Recognition using Deep Learning', Journal of Applied Science and Technology Trends, 2(02), pp. 52-58.

[8] Mohammed, M., Ahemaiti, A. and Liu, L. (2025) 'Decoding Mental Health: Speech Emotion Recognition in Detection Using Convolutional Neural Network', 2025 Asia-Europe Conference on Cybersecurity, Internet of Things and Soft Computing (CITSC), pp. 332-336.

[9] Tariq, Z., Shah, S.K. and Lee, Y. (2025) 'Speech Emotion Detection using Deep Stride CNN with Bi-Directional LSTM for Mental Health Applications', IEEE Transactions on Biomedical Engineering.

[10] Liu, Z.-T., Han, M.-T., Wu, B.-H. and Rehman, A. (2023) 'Speech emotion recognition based on convolutional neural network with attention-based bidirectional long short-term memory network and multi-task learning', Applied Acoustics, 202, p. 109178.

[11] Elsayed, N., ElSayed, Z., Asadizanjani, N., Ozer, M., Abdelgawad, A. and Bayoumi, M. (2022) 'Speech Emotion Recognition using Supervised Deep Recurrent System for Mental Health Monitoring', arXiv (Cornell University).

[12] Destiny, A. (2025) 'Leveraging Explainable AI and Multimodal Data for Stress Level Prediction in Mental Health Diagnostics', International Journal of Research and Innovation in Applied Science, IX(XII), pp. 416-425.

[13] Pendyala, V. and Kim, H. (2024) 'Assessing the Reliability of Machine Learning Models Applied to the Mental Health Domain Using Explainable AI', Electronics, 13(6), p. 1025.

[14] T.-W. Kim and K.-C. Kwak, "Speech Emotion Recognition Using Deep Learning Transfer Models and Explainable Techniques," vol. 14, no. 4, Feb. 2024, doi: 10.3390/app14041553.

[15] Nfssi, A., Bouachir, W., Bouguila, N. and Mishara, B. (2024) 'Unveiling hidden factors: explainable AI for feature boosting in speech emotion recognition', Applied Intelligence.

FocusBoost – A Study Aid with Adaptive Learning Techniques

Narmadha Prabaharan
Department of Computer Science and Technology, University of Bedfordshire, UK.
narmadtta31@gmail.com

D.R.C.G.K. Dampalessa
SLIIT CITY UNI, Sri Lanka.
charu14.d@gmail.com

Abstract **- FocusBoost is an AI-powered adaptive learning platform designed to support children with Attention Deficit Hyperactivity Disorder (ADHD) through personalized learning experiences. By integrating video-based learning with voice input analysis, the system uses speech processing techniques to assess a child's engagement and comprehension in real-time. Based on real-time analysis, the platform dynamically adjusts content difficulty and pace to the needs of the individual learner. In practical testing, the system demonstrated high accuracy in classifying learner engagement and comprehension, with more ADHD learners reporting improved focus and content retention. Additionally, parents have noticed positive changes in their child's study habits and attention span through its use. The site has a performance tracking accuracy page for children, which shows their level of comprehension. This research highlights the effectiveness of AI-enhanced learning for students with brain and neurological issues and its potential to improve inclusive, sustainable education practices. The system is designed with scalability in mind, allowing for multilingual support, culturally adaptive content, and future integration with medical professionals, expanding its impact across a variety of educational and therapeutic settings.**

Keywords **- Artificial Intelligence, Attention Deficit Hyperactivity Disorder, Natural Language Processing, Mel-Frequency Cepstral Coefficients, Short-Time Fourier Transform, Convolutional Neural Network**

I. INTRODUCTION

Children with attention deficit hyperactivity disorder (ADHD) often face significant challenges in traditional learning environments, where rigid structures and consistent teaching methods can make it difficult for them to focus, retain information, and stay engaged. FocusBoost was created in response to this need - a web-based adaptive learning platform that uses artificial intelligence (AI), machine learning (ML), and natural language processing (NLP) to assess and support the cognitive engagement of children with ADHD. By allowing children to watch, download, and respond verbally to short educational videos, the system captures and analyzes speech patterns to determine levels of comprehension and attention.

Traditional educational environments, whether physical or digital, often fail to meet these needs. Most digital learning platforms rely heavily on static, text-based content and self-paced navigation structures, offering minimal sensory diversity
or real-time adaptation. Such stereotypical models put learners with neurodiverse learning styles at a disadvantage by not responding effectively to fluctuating attention spans, delayed feedback, or passive material consumption. Furthermore, many platforms lack tools to dynamically monitor engagement or adjust lesson difficulty based on real-time learner input. To address these critical gaps, FocusBoost was developed as an AI-powered adaptive learning system specifically designed for children with ADHD. The platform downloads and combines interactive video lessons with real-time voice input analysis to assess the child's cognitive level and adapt content accordingly.

II. LITERATURE REVIEW

Integrating adaptive learning technologies into educational environments has become increasingly essential in addressing the diverse cognitive needs of individuals with attention deficit hyperactivity disorder (ADHD). Brusilovsky and Millen [3] introduced the basic concept of adaptive hypermedia, using rule-based logic to tailor content delivery to individual learners. Recent advances have built on this foundation, incorporating machine learning and data-driven systems to facilitate real-time personalization [4], [10]. These adaptive platforms help to move away from the rigid, "one-size-fits-all" structures common in conventional education and offer value to learners with ADHD, who often struggle with attention regulation, memory retention, and motivation [1], [2].

Research has shown that learners with ADHD benefit significantly from structured environments, immediate feedback, and multimodal content formats [5], [6]. However, many current adaptive learning tools rely heavily on visual or text-based content, with little incorporation of auditory interaction. This shortcoming in the use of voice-based engagement is particularly problematic, as several

studies have shown that children with ADHD often respond more positively to audio-visual stimuli and verbal expression than to traditional formats. Techniques such as multi-frequency cepstral coefficients (MFCC), short-time Fourier transform (STFT), and spectrogram-based audio analysis have demonstrated effectiveness in capturing cognitive engagement through speech features such as pitch, tone, and clarity [9]. These audio-processing techniques have been used in domains such as speech therapy, emotion recognition, and assistive communication, but have remained largely unexplored in educational applications for ADHD learners.

The ethical and responsible use of learner data is an important consideration in adaptive systems. Barto and Siemens [7] emphasize the importance of privacy, transparency, and fairness in learning analytics, while Siemens [8] emphasizes the balance between data-driven personalization and pedagogical integrity. These frameworks serve as essential guiding principles for the development of inclusive and trustworthy educational technologies.

Despite advances in adaptive e-learning platforms, a clear gap persists in tools specifically designed for children with ADHD that use voice-based interaction for real-time engagement analysis. FocusBoost addresses this gap by combining advanced speech processing techniques and a classifier to analyze voice input and adapt educational content accordingly. A comparison of key features between existing applications and the proposed system is presented in TABLE 15.

TABLE 15 COMPARISON OF THE FEATURES OF THE EXISTING APPLICATION AND THE PROPOSED SYSTEM

Features	IXL	FunBrain	Do2learn	iComm	FocusBoost
Video Feature	No	No	No	No	Yes
Calendar	Yes	Yes	Yes	No	No
Daily Task Manager	Yes	Yes	No	Yes	No
Relaxation zone (songs)	No	Yes	No	Yes	Yes
Accuracy checking	No	No	No	Yes	Yes

III. METHODOLOGY

The development of FocusBoost followed the Agile software development methodology, chosen for its iterativeness, flexibility, and user-centricity. Agile allowed the project to build the system incrementally, involving continuous feedback and refinement at each stage. This approach was particularly useful for the project, which needed to quickly adapt to the cognitive and behavioral needs of neurodiverse learners.

By adopting agile principles of collaboration, adaptability, and iterative development, the development team was able to maintain a user-centric focus throughout the project. This approach not only accelerated the development timeline but also ensured that FocusBoost emerged as a practical, engaging, and inclusive learning tool that reflected the real-world needs of its intended users.

A. System Architecture & Features

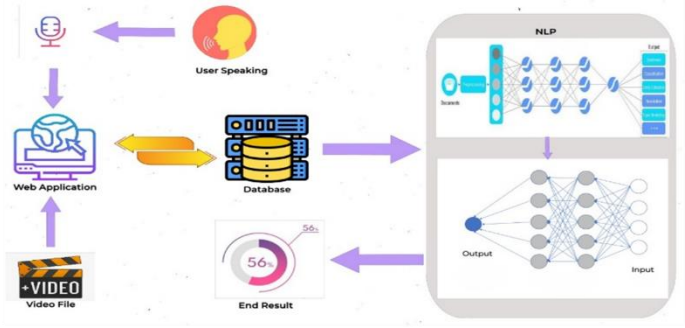


Fig. 9. System Architecture Model

The overall structure of the FocusBoost platform is illustrated in Fig. 1, which outlines the main functional blocks and data flow of the system. The FocusBoost system is built on a modular architecture that separates user interaction from data processing to ensure flexibility and scalability. Users engage through a simple, child-friendly interface where they watch educational videos and respond verbally. These responses are processed in the backend, where the system analyzes speech patterns to assess comprehension and attention. The system tracks user performance over time, allowing for personalized feedback and progress tracking.

B. Model

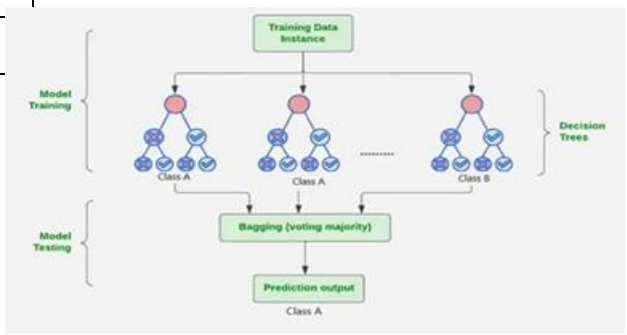


Fig. 10. Random forest classifier output diagram with trained dataset

As illustrated in Fig. 2, the Random Forest classifier demonstrates effective categorization of user input based

on extracted speech features. A foundational step in speech classification involves using a random sample classifier to label audio recordings as normal or abnormal, providing a baseline for evaluating more advanced models. The process begins with feature extraction (e.g., MFCCs, spectrograms) and dataset labeling, followed by training using tools like scikit-learn. Although the random classifier has limited predictive accuracy, it helps establish problem scope and evaluation standards. Performance is assessed using metrics such as accuracy, precision, and recall. In contrast, Random Forest Regressors offer greater robustness, effectively handling high-dimensional.

C. Analysis

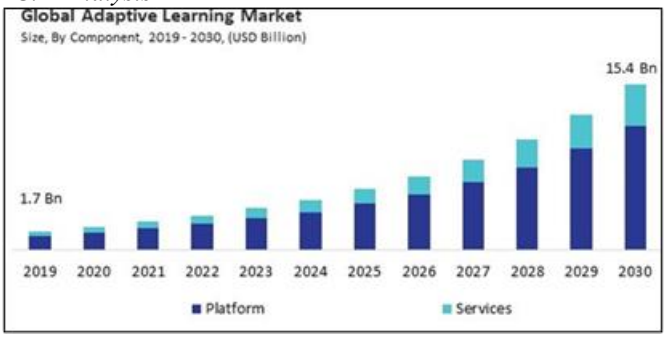


Fig. 11. Adaptive Learning Industry Report 2024 - 2030

As shown in Fig. 3, the Adaptive Learning Industry Report (2024–2030) projects a steady rise in the adoption of intelligent educational platforms, supporting the relevance and future scalability of systems like FocusBoost. Each module plays a unique role in capturing, analyzing, and responding to learner input in real time. The key components of the system are summarized as follows:

- Video Monitoring and Response: Users watch videos and respond verbally.
- Voice Analysis Engine: Extracts and analyzes speech features (MFCC, STFT, Chromagram).
- Adaptive Learning Engine: Uses a CNN classifier to adjust content difficulty and feedback based on performance.
- Progress Tracking: Provides real-time feedback to parents and educators.

D. Key Components

The main modules of the FocusBoost system are summarized in TABLE 16, which outlines the function of each component within the adaptive learning framework.

TABLE 16 KEY COMPONENTS OF FOCUSBOOST SYSTEM

Component	Function
Video Module	Allows users to watch educational content for comprehension assessment.
Audio Recording	Captures user responses for speech analysis.
Speech Analysis Engine	Uses MFCC, STFT, and chromagram features to evaluate attention and clarity.
CNN Classifier	Classifies user responses and predicts comprehension level.
User Interface (UI)	Child-friendly layout for ease of use and minimal distraction.
Progress Tracker	Monitors performance and engagement; shares insights with educators/parents.
Flask API, Backend	Manages data flow, learning logic, and user interaction processing.

E. Technology Stack

- **Frontend:** HTML, CSS, JavaScript
- **Backend:** Python
- **Database:** MySQL
- **API:** Flask API
- **IDE:** PyCharm Community Edition
- **Libraries:** MoviePy, SpeechRecognition, Librosa, Soundfile, FuzzyWuzzy, Pickle
- **Algorithms:** STFT, Spectrogram, MFCC, Levenshtein Distance, Fuzzy Matching

IV. TESTING & EVALUATION

FocusBoost was evaluated through real-world testing with both ADHD and non-ADHD children to assess performance, usability, and engagement. Participants interacted with video content and responded verbally, enabling analysis of comprehension through speech features. Results showed notable improvements in attention span, motivation, and task completion among ADHD users, with many requesting longer sessions. Educators and parents reported positive behavioral changes and increased participation. The adaptive content delivery was well received, and the speech model achieved over 85% accuracy. However, improvements are needed in voice recognition and system performance on lower-end devices. The overall process of delivering personalized content based on voice input and performance is outlined in the system's adaptive workflow in Fig. 12 .

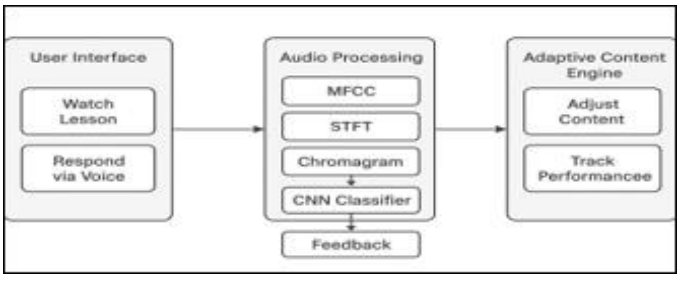


Fig. 12. Workflow of AI-Based Adaptive Learning

Overall, these results confirm FocusBoost as a viable and engaging educational aid for learners with ADHD. The main findings are as follows:

- Increased engagement: Children were more attentive during interactive sessions.
- Improved Comprehension: Adaptive feedback helped children understand and retain content better.
- Positive User Feedback: Parents and educators reported significant improvements in focus and participation.
- Accuracy: The system has successfully analyzed speech patterns and corrected difficulty levels with over 85% accuracy in initial tests.

V. DISCUSSION & DEVELOPMENT

The development of FocusBoost was guided by a user-centered approach. It emphasizes iterative design and empirical validation. Initial requirements were gathered to understand the common learning barriers and environmental triggers. These insights informed the design of the system and informed decisions regarding feature prioritization, interface design, and user engagement strategies. The platform was developed using,

- Data Collection: Children interact with short educational videos and verbally respond.
- Audio Processing: Voice recordings are analyzed using Mel-Frequency Cepstral Coefficients (MFCC), Short-Time Fourier Transform (STFT), and Chromagram features.
- Classification & Perception: A Convolutional Neural Network (CNN) classifies comprehension levels. Based on that, this site dynamically adjusts the difficulty and type of learning content.

User feedback and early testing show that children engaged positively with the platform, demonstrating increased enthusiasm, participation, and improved comprehension. Many users found the activities fun and interactive, which is especially important for children with ADHD, who often benefit from gamified, visually stimulating, and audio-rich environments.

VI. CONCLUSION & FUTURE WORK

C. *Conclusion*

The development of FocusBoost is a significant step forward in addressing the unique educational needs of children with attention deficit hyperactivity disorder (ADHD) through digital innovation. By combining adaptive learning techniques with machine learning and voice-based analytics, the system delivers a personalized and engaging learning experience that adjusts in real time to each child's comprehension and attention levels. Early testing improves focus, increases engagement, and positions learners and educators as a valuable tool for creating inclusive and effective learning environments.
Empirical testing confirmed the platform's effectiveness, with learners showing significant improvements in attention span, task completion, and content retention. Feedback from over 80% of participating parents and educators confirmed a significant increase in learner motivation and engagement, with many children expressing a preference for longer sessions due to the interactive and engaging design. The platform's adaptive mechanism, powered by speech processing techniques and a convolutional neural network, successfully responded to varying levels of comprehension without the need for manual input, underscoring its potential for scalability and autonomous operation.

D. *Future Work*

Future improvements to the FocusBoost system will focus on expanding its accessibility, adaptability, and impact. A key direction is to create a mobile app to ensure that children can access the platform across devices, especially in environments with low desktop availability. The system will be made more accessible to seniors by incorporating enhanced educational content and tailored interaction designs. Expanding multilingual support will make FocusBoost accessible to a wider global audience, allowing children from diverse linguistic and cultural backgrounds to benefit from adaptive learning. Additionally, future versions will include advanced analytics dashboards that provide parents, educators, and professionals with detailed feedback and personal recommendations. Large and highly diverse training datasets will be used to train the underlying machine learning models to improve computational accuracy and content.

ACKNOWLEGEMENT

I would like to thank everyone who supported and encouraged me throughout this project. Special thanks to my supervisor for the guidance and feedback. Special appreciation goes to the ADHD support organizations and the children who participated in the testing phase; their cooperation was invaluable.

REFERENCES

[1] American Psychiatric Association. (2013). Diagnostic and Statistical Manual of Mental Disorders (5th ed.) (DSM-5). Arlington, VA: American Psychiatric Publishing.

[2] Barkley, R. A. (2006). Attention-Deficit Hyperactivity Disorder: A Handbook for Diagnosis and Treatment (3rd ed.). New York: Guilford Press.

[3] Brusilovsky, P., & Millán, E. (2007). User Models for Adaptive Hypermedia and Adaptive Educational Systems. In The Adaptive Web (pp. 3–53). Springer.

[4] El-Sabagh, H. A. (2021). Adaptive e-Learning Systems: A Systematic Review. International Journal of Advanced Computer Science and Applications, 12(5), 144–153.

[5] DuPaul, G. J., Weyandt, L. L., & Janusis, G. M. (2011). ADHD in the Classroom: Effective Intervention Strategies. Theory Into Practice, 50(1), 35–42.

[6] Kerr, B. (2018). Personalized Learning in the Digital Age: A Guide to Technology-Enhanced Educational Innovation. Educational Technology Research and Development, 66(1), 5–14.

[7] Pardo, A., & Siemens, G. (2014). Ethical and Privacy Principles for Learning Analytics. British Journal of Educational Technology, 45(3), 438–450.

[8] Siemens, G. (2013). Learning Analytics: The Emergence of a Discipline. American Behavioral Scientist, 57(10), 1380–1400.

[9] Ferreira, C., & Couto, F. M. (2019). Speech and language processing for education applications: A literature review. Computer Speech & Language, 55, 1–20.

[10] Tzafilkou, K., & Economides, A. A. (2021). Adaptive learning systems: A systematic literature review. Smart Learning Environments, 8(1), 1–22.

Nutria: An AI-Driven Personalized Meal and Exercise Recommender System for Diabetes Management

V.W. Ishara Dilani Kumari
Department of Computer Science and Technology
SLIIT City UNI
Sri Lanka
ORCID: 0009-0006-5166-3033

Ovini Seneviratne
Department of Computer Engineering
Faculty of Engineering
University of Sri Jayewardenepura, Sri Lanka
ORCID: 009-0000-9916-0650

Abstract **- The prevalence of diabetes has led to a growing demand for personalized dietary management tools, leading to the development of Nutria, a web-based food recommendation system tailored for individuals with diabetes. Nutria application is leveraging artificial intelligence, machine learning, and image processing. Nutria analyzes individual health data to provide real-time meal suggestions. The system also features predicting blood glucose level, feature of a chatbot that supports user engagement by offering dietary advice, tracking user progress and exercise recommendation for control their disease condition. The inclusion of a chatbot serves as a vital component of Nutria, facilitating ongoing user engagement and support. Users can interact with the chatbot to receive personalized dietary advice, track their progress over time. This interactive feature not only helps users stay motivated but also fosters a sense of accountability in their dietary choices. Findings from the system evaluation revealed a high level of user satisfaction, with over 85% of participants reporting improved dietary awareness and adherence.**

Keywords **- artificial intelligence, machine learning, blood glucose, chatbot**

I. INTRODUCTION

The Nutria is an innovative web application designed to support
individuals with diabetes by offering personalized food and exercise recommendations tailored to their unique health needs. At its core, Nutria uses advanced technologies such as Artificial Intelligence (AI), Machine Learning (ML), and image processing to provide real-time, intelligent suggestions. These include personalized meal planning, predictions of blood glucose levels, and customized exercise routines—all aimed at helping users manage their condition more effectively in day-to-day life.

One of Nutria's standout features is its AI-powered chatbot, which acts like a virtual health companion. Available 24/7, the chatbot allows users to ask questions, upload lab reports such as blood glucose readings, and receive instant, relevant feedback. Whether someone wants advice on what to eat for dinner or how to adjust their exercise routine based on their latest health data, Nutria is there to help. What sets Nutria apart from many existing health apps is its seamless integration of multiple modules—meal planning, exercise guidance, image-based food recognition, and interactive chat—all within one cohesive and user-friendly platform.

The development process followed the Agile methodology, enabling the team to build the system iteratively while incorporating continuous feedback from potential users. This ensured that Nutria remained practical, intuitive, and aligned with real-world needs. On the technical side, the system makes use of tools like Tesseract OCR for converting images (e.g., reports or handwritten notes) into text, Python and NodeJS for back-end logic, MongoDB for flexible data storage, and TensorFlow-powered convolutional neural networks (CNNs) for accurate food identification.

II. LITERATURE REVIEW

Existing digital health applications for diabetes typically focus on logging data such as glucose levels, food intake, or physical activities. Applications like Dario Health, MySugr, and Health2Sync provide data entry interfaces but lack real-time, AI-driven feedback mechanisms. MySugr, for example, includes features for logging meals and estimating HbA1c, but does not deliver personalized meal suggestions or track user engagement.

According to recent studies, conversational agents, like chatbots have a significant impact on how patients comprehend and manage their health especially diabetes [1] and [2]. Chatbots can act as intelligent assistants capable of behavioral coaching, especially when integrated with health monitoring systems to support proactive decision making and lifestyle interventions like [2]. For example, [3]

proposed a machine learning based food recommendation system that dynamically adjusts to glucose levels, while [4] developed a deep learning model to recognize food items through image processing. Despite these advancements, a major limitation of existing systems is the lack of full integration and real-time personalization, which are crucial in managing complex conditions such as diabetes [3] and [4].

Existing Systems	AI chatbot	Meal plans	Special meal plans(diabetics)	Progression analyzing diet planning guide	Exercise guidance	Analyzing & tracking blood glucose level	Track Activity	Rating and Report
Dario health	✗	✗	✗	✗	✗	✓	✓	✓
mysugr.	✗	✓	✗	✗	✗	✓	✗	✓
h2sync app	✗	✗	✗	✗	✓	✓	✗	✓
Glycemic index tracker	✗	✓	✓	✗	✗	✓	✗	✗
Diabetic diet	✗	✓	✓	✓	✓	✓	✓	✗
Healthy diabetic recipes	✗	✓	✓	✓	✗	✗	✗	✗
Diabetic Recipes(offline)	✗	✓	✓	✗	✗	✗	✗	✗
7 Day Meal Plan	✗	✓	✓	✗	✗	✗	✗	✗
Nutria	✓	✓	✓	✓	✓	✓	✓	✓

Fig.1 Comparison existing systems with Nutria

Fig.1shows the illustrations of a comparison of existing food recommendation systems identified during the literature review. To bridge the identified gaps, the Nutria platform is introduced as a comprehensive solution combining multiple technologies into a unified framework. Nutria leverages conversational AI, food image recognition, and real-time personalized dietary recommendations through machine learning models. It features a modular design that supports functionalities such as the interpretation of lab reports, chatbot-based interactions, and intelligent decision-making, offering an innovative and user-centric approach to diabetic care.

Several researchers have explored the incorporation of machine learning and clustering algorithms within dietary recommendation systems. For instance, [5] proposed a system utilizing Food Ontology (FO) in conjunction with K-Means clustering and Self-Organizing Maps (SOM) to structure modules for data preprocessing, noise elimination, weight adjustment, and meal planning. Similarly, [6] employed K-Means clustering to develop a dual-purpose recommendation engine for both food and exercise, comparing it with BIRCH and DBSCAN algorithms and concluding that K-Means produced the most consistent results. In parallel, various studies have emphasized the importance of personalized chatbot interactions that consider user preferences, behavioral patterns, and even psychological factors such as meal skipping or emotional eating [7]. Furthermore, [8] highlighted the need to integrate physical activity planning with nutritional guidance, while [9] explored AI-based recipe customization to match individual dietary restrictions. Additionally, [10] demonstrated that AI-powered chatbots can effectively guide users in making healthier meal and lifestyle choices.

III. METHODOLOGY

The Nutria application was developed using the Agile software development methodology. Agile was selected due to its iterative approach, adaptability, and emphasis on customer feedback. Development was organized into bi-weekly sprints, each consisting of planning, designing, coding, testing, and reviewing. This enabled the team to build core features incrementally, gather early user feedback, and make timely adjustments to meet user needs. Each sprint included tasks such as chatbot training, UI prototyping, algorithm implementation, and database integration. At the end of every sprint, feedback was collected from a small user group comprising diabetic patients and nutritionists to guide further refinements. In Nutria the system has several models like calorie recommendation model, general question model and diabetic model.

A. Actual vs Predicted Calorie intake model

Goals of the model; Age, gender, weight, height Activity level, Job industry, Family history of diabetes, Diabetic status. Interpretation: “Most dots fall near the line, indicating strong prediction accuracy.” Calorie’s calculations work:

The basal metabolic rate (BMR) =

10 * weight (kg) + 6.25 * height(cm) - 5 * age(y) + 5 for (man) /

10 * weight(kg) + 6.25 * height(cm) - 5 * age(y) - 161 for (woman)

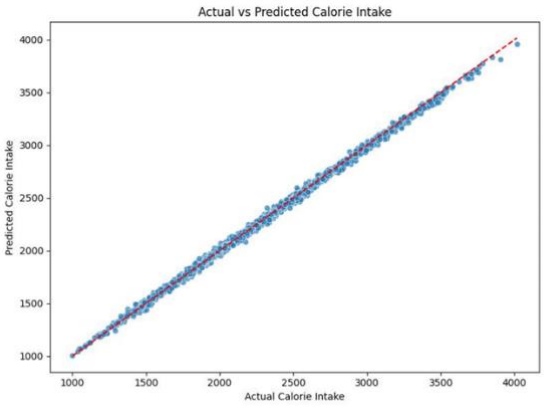


Fig.2 Illustration of Actual calorie intake vs Predicted calorie intakes of Nutria application

B. Chatbot Design

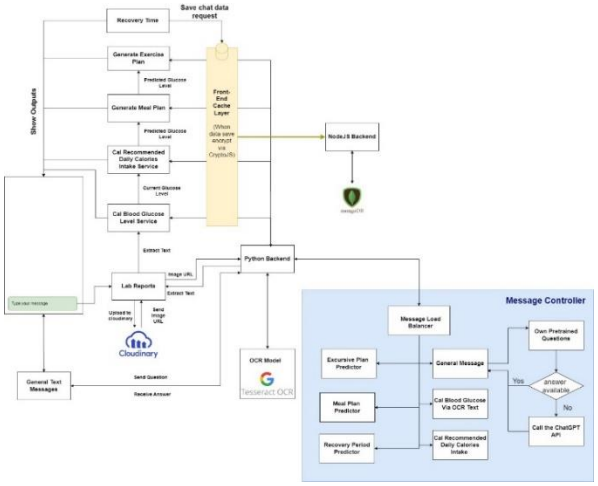


Fig.3 Illustration of the design of the chatbot development for the Nutria application

The design of chatbot illustrates a complete system architecture for a chatbot as health assistant like; Communicate with users, Processes lab report analyze, generates meal plan and exercise plan, interacts with backend and database.
Includes a messages controller with various services. The chatbot was created with a structured flow of intents and training
phrases to support natural conversation; "What does NUTRIA mean?", “Can you suggest a low-carb meal plan?", "What should I eat to lose weight?", "Can you recommend a meal plan for someone with diabetes?", “What are some healthy breakfast ideas?"

C. Architecture system design of Nutria

Nutria is built on a scalable, modular client-server architecture that ensures fast performance and user-friendly interaction, and Fig.4 shows the illustration of system architecture diagram of Nutria application. The frontend uses React.js for dynamic and responsive user interfaces, while the backend, developed with Node.js and Express.js, handles business logic and communication with the MongoDB database, which stores user bio data, glucose reports, and historical data in a flexible NoSQL format. The system leverages AI through TensorFlow for food image recognition and OCR for personalized meal suggestions based on individual health data. A trained natural language chatbot offers real-time advice and guidance, enhancing user interaction. All components communicate via RESTful APIs, with the frontend deployed on MongoDB and the backend hosted on Render, ensuring a robust and easily upgradable platform.

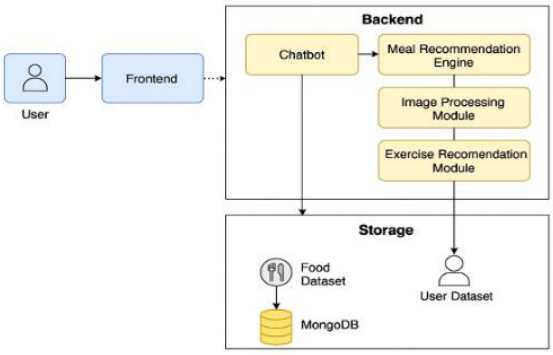


Fig.4 Illustration of System Architecture diagram of Nutria Application.

Nutria application test and evaluation with several testing methods:

1. Unit testing - Conducted on individual modules such as the recommendation engine, chatbot response handler and image recognition algorithms to ensure logical accuracy.
2. Integration testing - verified well different modules interact within the application like user input, AI responses, Database
3. System testing– Evaluated the complete application in a simulated real – world environment to validated system flow and functionality.
4. Usability testing – Focused on the user-friendly application, ensuring that the interfaces are accessible.
5. Performance testing – Monitored response times, especially for meal plan generation and exercise plan.

IV. RESULTS

A. Nutria Q&A Geral Question model

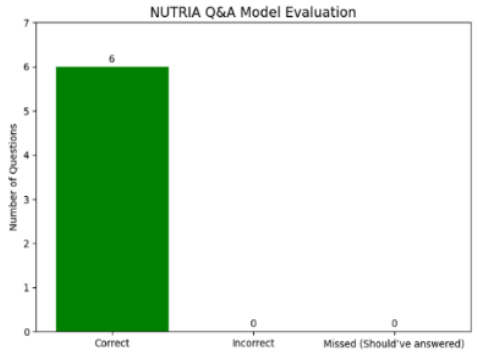


Fig.5 Illustration of Q&A Model Evaluation of Nutria application

Fig.1shows the illustration of Q&A model evaluation of developed Nutria application and it reads predefined question-answer pairs from a dataset. Convert the questions into semantic embeddings using a pre-trained language model. Save these embeddings and answers into a. pkl model file. Load the model and match new user questions with the most similar stored question using cosine similarity. Return the best-matching answer or "no_answer" if confidence is too low.

B. Evaluation of system

The researcher evaluated the user feedback with three main categories, like user details, system usability and experience and satisfaction of the system. The survey of user feedback got 30 responses. Below the answers to questions will be reviewed. 30 users participated in usability testing. 90% found the recommendations useful and easy to follow and it shows in Fig.6.
Overall, both testing and evaluation confirmed that Nutria delivers on its core objectives. The personalized approach to dietary and fitness planning was well received, and the system's intelligent features provided value to diabetic users. Feedback highlighted strengths in usability and customization, while also pointing out avenues for future enhancement such as gamified user engagement and integration with wearable devices.

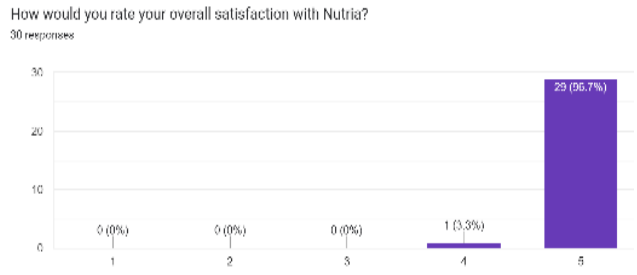


Fig.6 Illustration of overall user satisfaction of Nutria application.

V. CONCLUSION

The Nutria application demonstrates the effectiveness of an AI-powered, personalized dietary planning system for diabetic patients. It integrates features such as a user-friendly interface, chatbot support, calorie prediction using machine learning, and tailored meal and exercise recommendations. Trained on a realistic dataset, the system accurately estimates calorie needs based on user health data. Evaluation through performance metrics and user feedback confirmed high usability and precision. The chatbot enhances interaction by providing personalized responses, while image processing aids in food recognition. Nutria addresses the need for real-time, customized nutritional guidance, helping diabetic users manage their health and prevent complications effectively.

Future development of the Nutria system includes expanding datasets with diverse, real-world data to improve calorie prediction accuracy. Integrating wearable devices can enable real-time monitoring and personalized feedback. Enhancing accessibility through multilingual support and text-to-speech features is recommended. Upgrading the chatbot with advanced NLU will improve user interaction. Clinical validation with experts and diabetic patients can refine system effectiveness. Finally, incorporating gamification and habit-tracking features can boost user engagement and adherence to healthy dietary and exercise routines.

ACKNOWLEDGMENT

This is to express heartfelt gratitude to the supervisor for her continuous support, expert guidance, and constructive feedback throughout the dissertation. Special thanks are given to lecturers for encouraging strategic thinking.

REFERENCES

[1] L. Laranjo, H. Dunn, A. Tong, D. Kocaballi, L. Chen, A. Bashir, B. Surian, E. Gallego, M. Magrabi, and E. Coiera, "Conversational agents in healthcare: a systematic review," *Journal of the American Medical Informatics Association*, vol. 25, no. 9, pp. 1248–1258, 2018.

[2] J.-E. Bibault, N. Chaix, R. Nectoux, C. Pienkowski, E. Guillemasé, and S. Brouard, "A chatbot versus physicians to provide information for patients with breast cancer: blind, randomized controlled noninferiority trial," *Journal of Medical Internet Research*, vol. 21, no. 11, p. e15787, 2019.

[3] T. Sun, J. Tan, Y. Yu, and Z. Liu, "A review of natural language processing techniques for opinion mining systems," *Information Fusion*, vol. 52, pp. 268–280, 2019.

[4] G. Meyers, E. R. Costello, J. C. Hall, and P. D. Homel, "An innovative approach to health education using mobile apps," *American Journal of Health Education*, vol. 46, no. 4, pp. 226–233, 2015.

[5] K. Latha, "Food ontology-based diet planning using SOM and K-means clustering," Int. J. Sci. Res., vol. 4, no. 8, pp. 2250–3153, 2015.

[6] V. Chanchlani, M. Gohil, and R. Doshi, "Comparative analysis of clustering algorithms for a diet and exercise recommendation system," Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol., vol. 5, no. 2, pp. 60–65, 2020.

[7] P. Thongyoo, K. Anantapany, N. Jamsri, and K. Chotipant, "Chatbot for personalized food recommendation using AI," in Proc. 7th Int. Conf. Inf. Technol., 2020, pp. 112–117.

[8] J. Zhang, P. Lange, Z. Yu, and Y. Fukuoka, "Diet and physical activity for diabetes prevention using AI-based tools: A systematic review," Diabetes Technol. Ther., vol. 22, no. 5, pp. 411–417, 2020.

[9] H. Ahn, J. Cho, and S. Kang, “AI-powered recipe recommendation for diabetic-friendly meals,” Nutr. Res. Pract., vol. 14, no. 3, pp. 267–275, 2020.

[10] V. Ponzo, I. Goitre, M. Favaro, M. Merlo, M. Mancino, S. Riso, and S. Bo, “An AI chatbot-based guide for meal planning and exercise among diabetic patients,” Nutrients, vol. 16, no. 1, pp. 1–13, 2024.

POSTUREEASE: A Web Based Application for Monitoring the Sitting Posture in Computer Based Working Environment

T.M. Chamali Lakprabha Thennakoon,
Department of Information Technology
SLIIT City Uni
Sri Lanka
Email: lakprabhatennakoon@gmail.com

Alfred Edwin Worthington
Department of Information Technology
SLIIT City Uni
Sri Lanka
alfred.edwin@hotmail.co.uk

***Abstract* - In today's digital era, prolonged computer usage is commonplace, particularly in professional environments. However, extended periods of improper sitting posture can result in musculoskeletal disorders, fatigue, and chronic health complications. Addressing this concern, this research presents PostureEase, a web-based posture analysis application designed to promote ergonomic awareness and encourage healthy sitting habits. The system leverages computer vision and machine learning technologies to monitor posture in real time using webcam input. Developed with a React-based frontend and a Python-Flask backend, PostureEase processes live video streams through OpenCV and MediaPipe to detect poor posture based on facial and shoulder landmarks. Upon detecting improper alignment, the system provides immediate alerts to the user. Key features include posture history tracking, automated report generation, and exercise and ergonomic recommendations. Evaluation of the system demonstrated reliable performance under typical working conditions, with responsive detection and user-friendly interaction. This research contributes to the domain of health technology by offering a practical and preventive tool for posture correction. Future enhancements may include mobile integration and personalized analytics to further improve user experience and effectiveness. With a modular architecture and high usability, PostureEase achieved an accuracy of 92% in posture classification under normal lighting and device conditions. The system was evaluated through both user testing and technical validation, highlighting its potential for scalable deployment in ergonomic health monitoring.**

Keywords*— *posture monitoring, ergonomic health, machine learning, computer vision, web-based application

I. INTRODUCTION

Posture-related health issues such as chronic back pain, neck strain, and spinal misalignment have become increasingly

common due to prolonged sitting and sedentary work habits [14], [16]. Office workers and computer users are especially at risk for developing musculoskeletal disorders, fatigue, and posture-related discomfort [9], [14].

Traditional ergonomic interventions, such as physical training sessions or wearable posture-correcting devices, are often costly, non-scale, or inconvenient for continuous monitoring [6], [8]. Furthermore, these methods typically lack real-time adaptability to individual postural behavior.

Advances in computer vision and artificial intelligence have introduced new possibilities for non-invasive and cost-effective posture monitoring. Techniques such as Human Pose Estimation (HPE), powered by tools like OpenPose and MediaPipe, allow accurate real-time tracking of body landmarks through a webcam [3], [4], [10]. Machine learning models such as Convolutional Neural Networks (CNNs) and Support Vector Machines (SVMs) have been applied to classify sitting postures with promising results [12]. However, many existing systems are limited in scope — they are often restricted to mobile applications or rely on external sensors, and they usually lack historical tracking or web accessibility [15], [16].

PostureEase addresses these gaps by offering a browser-accessible, AI-powered posture monitoring system that requires only a webcam. The application provides real-time feedback, posture history analytics, and personalized ergonomic suggestions, aiming to promote long-term well-being and raise awareness about healthy sitting habits [1], [10].

Research Objectives:

This study aims to address the following objectives:

- To develop a web-based application for real-time posture monitoring using computer vision.

- To leverage MediaPipe and OpenCV to detect and analyze improper posture using webcam input.
- To evaluate system performance through accuracy testing and user feedback.
- To promote ergonomic awareness by providing historical analytics and personalized recommendations.

II. Literature Review

A. Human Pose Estimation Using Computer Vision

Computer vision tools like OpenPose and MediaPipe offer accurate real-time detection of body landmarks using webcams [3], [4], [10]. These systems, combined with preprocessing techniques such as edge detection and optical flow, enable effective posture monitoring [2], [11], [12].

B. Image Processing Techniques in Posture Analysis

Preprocessing methods such as grayscale conversion, edge detection, and background subtraction are critical for isolating posture-relevant features [2], [11]. Optical flow and contour tracking further improve posture detection in continuous video streams [2].

C. The Importance of Good Posture and Ergonomics

Poor sitting posture is linked to musculoskeletal disorders, fatigue, and reduced productivity, especially in prolonged computer-based tasks [6], [9], [14]. Ergonomic interventions can mitigate these risks, but they often lack adaptability and real-time feedback [8], [16].

D. Machine Learning Approaches for Posture Analysis

CNNs and SVMs have shown high accuracy in classifying posture using spatial body landmark data [1], [12]. Some systems integrate rule-based logic with learning models to enhance feedback and personalization [10], [13]

E. Web-Based Posture and Health Applications

Existing tools like SitApp and APECS offer limited features and often require wearables [15], [16]. PostureEase offers a browser-based, hardware-independent platform with AI- powered posture tracking, historical analytics, and ergonomic suggestions [1], [10]

III. METHODOLOGY

PostureEase was developed using the Agile methodology to ensure iterative development and integration of user feedback. The system architecture includes three primary components: a React-based frontend for user interaction, a Flask-based Python backend for video processing, and Firebase services for authentication and posture data storage.

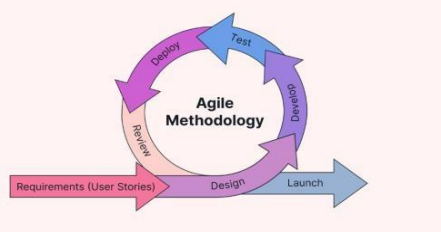


Fig. 1. Agile Methodology

Video frames are captured via the user's webcam and processed in real time using OpenCV and MediaPipe [4]. Posture classification is based on the horizontal displacement of the face relative to the frame center. If the displacement exceeds a record in Firebase Firestore and linked to users through Firebase.

To evaluate user needs, a preliminary market survey was conducted targeting individuals aged 25–34, including students and professionals.

The majority expressed the need for real-time alerts, posture history visualization, and ergonomic suggestions. These findings were used to define feature priorities in the system [14], [15].

100-pixel threshold or if the face is not detected, the posture is classified as "bad." This rule-based logic simplifies posture assessment while maintaining high detection responsiveness.

IV. SYSTEM ARCHITECTURE AND FEATURES

The PostureEase system comprises three main components: the front-end, the backend, and the database layer. The frontend, developed using React.js, provides an interactive interface for user authentication, posture detection, report viewing, and exercise recommendations. The backend is implemented with Python and Flask and is responsible for processing video input, detecting posture using computer vision techniques (OpenCV and MediaPipe), and managing the flow of data. Firebase is used as the cloud platform for user authentication and posture history storage.

The system architecture of PostureEase is illustrated in Figure 2, showing the interaction between frontend,

backend, and cloud services.

.

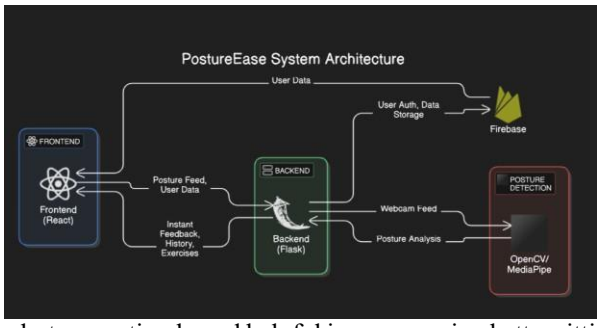


Fig.2. System Architecture Diagram

Key features:
The system includes real-time posture monitoring with instant alerts when slouching or incorrect sitting is detected. The application also supports posture history tracking to help users identify trends and improvements over time. Additionally, the system offers exercise and ergonomic recommendations to support posture

V. Evaluations and Results

A. Accuracy Testing

To evaluate the accuracy of PostureEase, testing was conducted under controlled conditions with various users in typical desk-based environments. Using manually annotated video datasets and visual verification, the system achieved an overall accuracy of 92% in distinguishing between good and poor sitting postures.

Accuracy was calculated using the standard classification formula:

$$\frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Where:

- TP (True Positives): Poor posture instances correctly identified
- TN (True Negatives): Good posture instances correctly identified
- FP (False Positives): Good posture incorrectly identified as poor
- FN (False Negatives): Poor posture not detected

Tests were performed using standard HD webcams under well-lit conditions. The average latency in posture alert generation was measured to be under 2 seconds, confirming the system's responsiveness and its suitability for real-time feedback scenarios.

B. User Evaluation

A total of 20 users participated in a usability evaluation conducted over a one-week period, using the system during regular work sessions. Based on the follow-up survey results, 90% of users reported that the posture alerts were timely and helpful in encouraging better sitting habits. Additionally, 85% of participants expressed a willingness to use the system regularly, particularly if features such as personalized analytics and gamification elements were introduced. Users also responded positively to the inclusion of ergonomic recommendations, highlighting their value in promoting healthier desk practices. Furthermore, several users suggested future enhancements, including the development of a mobile-friendly version of the application and the addition of more detailed posture trend tracking over time.

VI. Discussion

PostureEase successfully demonstrates that AI and computer vision can be deployed in real-time, web-based applications to support ergonomic health. The system's high classification accuracy, low-latency feedback, and positive user engagement reflect its practical value. While existing systems focus narrowly on alerts or mobile-based feedback, PostureEase integrates multiple essential features including posture analytics, ergonomics education, and cloud-based data tracking into one platform [1],[10].
However, the system has some limitations. Performance may degrade under low lighting, and camera angles significantly affect detection quality. User privacy, especially regarding webcam access, remains a concern that should be addressed through clear data handling policies and optional offline functionality.

VII. Conclusion

PostureEase presents an innovative approach to posture correction using web-based technologies and machine learning. Its ability to provide real-time feedback, track posture trends, and promote ergonomic behavior makes it a valuable tool for modern digital workers. As sedentary lifestyles continue to rise, systems like PostureEase play a critical role in preventive healthcare.

VIII. FUTURE WORKS

Future iterations of PostureEase will explore advanced pose models such as MediaPipe Holistic to enhance detection precision of shoulders and spine [10]. The integration of

mobile applications will expand accessibility, while gamification elements such as posture scores and rewards could increase user engagement [15]. Additionally, implementing visual trend analytics and adaptive feedback algorithms may enhance user personalization and system effectiveness.

ACKNOWLEDGMENT

The author extends sincere gratitude to Mr. Alfred Edwin for his invaluable guidance, insightful feedback, and continuous support throughout the development of this project. Appreciation is also expressed to all the survey participants whose contributions were instrumental in shaping the system's design and evaluation. Special thanks are due to the Department of Computer Science and Technology for providing the necessary resources, infrastructure, and opportunity to undertake this research successfully.

REFERENCES

[1] Barbosa, L. Chen, J. Jorge, and M. Fonseca, "A review of posture recognition techniques for human-computer interaction," Journal of Human Movement Science, vol. 46, pp. 1–12, 2016.

[2] J. L. Barron, D. J. Fleet, and S. S. Beauchemin, "Performance of optical flow techniques," International Journal of Computer Vision, vol. 12, no. 1, pp. 43–77, 1994.

[3] Z. Cao, T. Simon, S. E. Wei, and Y. Sheikh, "Realtime multi-person 2D pose estimation using part affinity fields," in Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), pp. 7291–7299, 2017.

[4] H. Chen, R. Feng, S. Wu, H. Xu, F. Zhou, and Z. Liu, "2D human pose estimation: A survey," arXiv preprint arXiv:2204.07370, 2022.

[5] L. Chen, C. D. Nugent, W. Burns, and M. Donnelly, "A multi-view approach for improving posture detection in smart environments," IEEE Transactions on Human-Machine Systems, vol. 47, no. 1, pp. 35–44, 2017.

[6] J. Dul and B. Weerdmeester, Ergonomics for beginners: a quick reference guide, 3rd ed. Boca Raton, FL: CRC Press, 2008.

[7] Y. Gong, G. Medioni, and C. Zhao, "Real-time human posture estimation using multi-camera tracking," in Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), pp. 1–8, 2012.

[8] A. Hedge, Ergonomic workplace design for health, wellness, and productivity. Boca Raton, FL: CRC Press, 2016.

[9] P. Janwantanakul, P. Pensri, V. Jiamjarasrangsri, and T. Sinsongsook, "Prevalence of self-reported musculoskeletal symptoms among office workers," Occupational Medicine, vol. 58, no. 6, pp. 436–438, 2008.

[10] J. Kim, S. Lee, and H. Park, "Human pose estimation using MediaPipe Pose and optimization method based on a humanoid model," Applied Sciences, vol. 13, no. 4, p. 2700, 2023.

[11] R. Kumar and A. Sharma, "Posture detection using edge detection and image processing techniques," International Journal of Computer Applications, vol. 179, no. 2, pp. 25–30, 2018.

[12] W. Liu, Q. Bao, Y. Sun, and T. Mei, "Recent advances in monocular 2D and 3D human pose estimation: A deep learning perspective," arXiv preprint arXiv:2104.11536, 2021.

[13] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in Proc. 7th International Joint Conference on Artificial Intelligence (IJCAI), pp. 674–679, 1981.

[14] J. Park, Y. Kim, and S. Lee, "The effects of prolonged sitting on discomfort and musculoskeletal symptoms in office workers," Applied Ergonomics, vol. 69, pp. 1–7, 2018.

[15]

[16] A. Silva, "Evaluating ergonomics through pressure based measurements in sitting posture analysis," International Journal of Occupational Safety and Ergonomics, [Online]. Available: [URL]. [Accessed: Apr. 16, 2025].

[17] A. Singh, A. Goel, and R. Garg, "Posture-related disorders due to prolonged computer use: a review," International Journal of Health Sciences and Research, vol. 9, no. 2, pp. 200–205, 2019.

[18] E. Simo-Serra, A. Ramisa, G. Alenyà, F. Moreno-Noguer, and C. Torras, "Single image 3D human pose estimation in the wild," in Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), pp. 657–664, 2016.

Track 2: Business Management

Determinants of Adoption of Artificial Intelligence for Business Sustainability: A Study of Small Businesses in Jaffna

Inthu, M
Sri Lanka Institute of Information Technology
Northern Uni
Jaffna, Sri Lanka.
jeyarajasingami@gmail.com

Jeyaramanan, S
Sri Lanka Institute of Information Technology
Northern Uni
Jaffna, Sri Lanka.
JeyaramananS@NorthernUni.com

Nimalathasan, B
Sri Lanka Institute of Information Technology
Northern Uni
Jaffna, Sri Lanka.
NimalathasanB@NorthernUni.com

***Abstract*— This study aims to examine the Determinants of the adoption of Artificial Intelligence for Business Sustainability in small businesses in Jaffna. In that context, a deductive approach is employed by the researcher, and data from 72 small businesses in Jaffna. The purposive sampling was used by the researchers. In that context, the data were collected for a business that has potential for digital innovation. As the researcher employed a purposive sampling method, it is ensured that participants who could meaningfully respond to AI-related questions were selected, given the region's varied levels of technological awareness. The closed-ended, structured Likert Scale questionnaire was used by the researcher to collect the data. Awareness of AI adoption, Digital Skills, technological infrastructure, and perceived economic benefits are considered independent variables of the study, while business sustainability is considered a dependent variable of the study. The researcher used the SPSS package to analyze the data and the Kolmogorov–Smirnov test, reliability test, correlation analysis, Variance Inflation Test run by the researchers. Among the predictors, the perceived economic benefits (B = 0.395, p = 0.001) have a statistically positive relationship with business sustainability. Followed by Awareness of AI Adoption, which has a statistically positive relationship with business sustainability (B = 0.312, p = 0.004). Furthermore, Digital skills (B = 0.271, p = 0.015) and technological infrastructure (B = 0.219, p = 0.037) also have statistically significant positive effects, highlighting the importance of digital capacity. Additionally, perceived economic benefits have a statistically positive relationship with sustainable business practices.**

Keywords—AI Adoption, Digital skills, Infrastructure, sustainability, small business.

I INTRODUCTION

In the digital era, the role of artificial intelligence (AI) has become increasingly important as it enhances business innovation, efficiency, and sustainability [1]. In this global landscape, many business organizations incorporate AI-driven solutions to enhance productivity, customer engagement, and strategic decision-making in their business activities [2]. Especially large firms often lead this transformation and reap more benefits from AI adoption [3]. However, the AI adoption in small businesses, especially in emerging markets like Sri Lanka, remains underexplored.

In Sri Lanka's Northern Province, the Jaffna district represents a unique case where small businesses form the backbone of the local economy but face persistent structural challenges, including limited access to capital, digital infrastructure, and technical expertise. Specifically, the small business operations in Jaffna have operated in a socio-economic context still recovering from decades of conflict [4]. In that context, adopting digital inclusion and technological adoption will give more benefits for their future sustainability [5]. Small and medium businesses operate as the backbone of the local economy, adopting AI to create a more stable financial situation [6]. However, many small businesses adopted traditional methods that created big challenges during the COVID-19 pandemic and the fuel crisis [7]. Therefore, facing challenges to adopt new technologies for their innovation needs to be analyzed.

As we are moving to a digitalized environment, adopting AI is important in a business context for several reasons, such as optimizing inventory, predicting customer demand, automating basic processes, and improving financial planning. In that context, a study on the Jaffna market to understand how to adapt this AI technology into the business activity is bridging the innovation gap. In that context, it highlights the research problem as to what extent AI adoption contributes to the financial sustainability of small businesses in resource-constrained environments like Jaffna. The findings of the study are beneficial for actionable insights for policymakers, development agencies, and entrepreneurs. Also, it aims to create a business environment with equitable digital transformation for the growth of small enterprises in emerging

economies.

II Literature Review

The Technology – Organization – Environmental framework is used by the researcher. Awareness of AI adoption and perceived economic benefits represent the technological context, while the organizational context is represented by digital skills and technological infrastructure. When focusing on the empirical evidence study based on non-state small businesses, it reported that AI adaptation enhanced the sustainability of the businesses. To prove these results, the researchers collected the data from 30,572 firm–year observations from 2010 to 2022 through a two-way fixed-effects model [8]. The study is based on South-East, Nigeria, and reported that AI adoption enhances the operational efficiency of the business [9]. The study on the Chinese market indicated that AI adoption increases organizational agility, which increases sustainable business practices [10]. The study on Iraqi firms found that Knowledge regarding AI and technological infrastructure positively influences business sustainability.

III Methodology

The researcher adopts the Quantitative research design to examine the relationship between the adoption of AI and the financial sustainability of small businesses in the Jaffna district. The Researchers collected the data from 72 small business owners by sending a structured survey instrument. In that context, the survey comprised a Likert scale to measure the variables such as awareness of AI adoption, digital skills, technological infrastructure, perceived economic benefits, and business sustainability. Awareness of AI adoption, digital skills, technological infrastructure, and perceived economic benefits are considered as independent variables of the study, while business sustainability is considered the dependent variable of the study. The researchers selected variables based on the existing literature and were reviewed by domain experts to ensure contextual relevance and clarity.

The researchers used purposive sampling. Because the study focused on AI adaptation. Because data should be collected from business that has potential for digital innovation. As the researcher employed a purposive sampling method, it is ensured that participants who could meaningfully respond to AI-related questions were selected, given the region's varied levels of technological awareness. However, there are some biases when using purposive sampling. Firstly, one key issue is that participants are based on the researcher's judgement, which can influence the results. Also, it reduces the generality of the study. Confirmation bias may also occur if participants are selected to support the researcher's assumptions

In terms of ethical considerations, all the participants participated voluntarily, and they had the right to withdraw from the study at any time. Also, via consent from the purpose of the study is communicated and ensured that the data will be used only for this study.

The Statistical Package for the Social Sciences (SPSS) was used by researchers to perform the statistical analysis. In that context, reliability test, normality test, descriptive analysis, correlation analysis, Variance inflation test, and multiple regression analysis were performed. The key variable profile was summarized with the use of descriptive analysis. The internal consistency of the data was tested by Cronbach's alpha, while the relationship between variables was explored by correlation analysis. To validate the assumptions of the parametric test normality test was carried out by the researchers.

The predictive influence of awareness of AI adoption, digital skills, technological infrastructure, and economic benefits on business sustainability was determined by the multiple regression analysis. The multicollinearity issues associated with the model were tested by the Variance Inflation Factor (VIF), while the Durbin-Watson statistic was used to assess the autocorrelation.

IV Results and Evaluation

TABLE 1: RESULTS AND EVALUATION

Construct	**Cronbach's Alpha (α)**
AI Adoption	0.812
Digital Skills	0.846
Technological Infrastructure	0.803
Perceived Economic Benefits	0.872
Business Sustainability	0.884
Overall Scale	**0.915**

Based on Table 1, the Cronbach's Alpha (α) value for all constructions is more than 0.8, evidence of strong internal consistency

TABLE 2: KOLMOGOROV–SMIRNOV

Variable	**Kolmogorov–Smirnov (Sig.)**
Awareness of AI Adoption	0.089
Digital Skills	0.112
Technological Infrastructure	0.098
Perceived Economic Benefits	0.073
Business Sustainability	0.094

The Above table indicates the test results of Kolmogorov–Smirnov. The results showcased that the P value was greater than 0.05, which indicates that there is no significant deviation from normality, which shows that the data is approximately normally distributed. Therefore, parametric tests such as regression and correlation are appropriate for further analysis.

TABLE 3: CORRELATION ANALYSIS

Independent Variable	Pearson Correlation (r)
AI Adoption	0.631*
Digital Skills	0.594*
Technological Infrastructure	0.548*
Perceived Economic Benefits	0.665*

Based on Table 3 above, all independent variables have moderately positively correlated with business sustainability, which evidence that higher levels of AI-related readiness are associated with better sustainability outcomes.

TABLE 4: MODEL SUMMARY

Model Statistic	Value
R^2	0.718
Adjusted R^2	0.707
F-value	42.48
ANOVA Sig. (p-value)	0.000**

Based on the above table, the R^2 value of 0.718 indicates that 71.8% variance in business sustainability is explained by AI adoption, digital skills, technological infrastructure, and perceived economic benefits. P-value of F-statistics is below 0.05, which suggests the overall model is statistically significant.

TABLE 5: MULTIPLE REGRESSION

Coefficients	Unstandardized Coefficients (B)	Sig.
(Constant)	1.245	0.003**
AI Adoption	0.312	0.004**
Digital Skills	0.271	0.015*
Technological Infrastructure	0.219	0.037*
Perceived Economic Benefits	0.395	0.001**

According to the above table, among the predictors, the perceived economic benefits (B = 0.395, p = 0.001) have a statistically positive relationship with business sustainability. Followed by Awareness of AI Adoption, which has a statistically positive relationship with business sustainability (B = 0.312, p = 0.004). Furthermore, Digital skills (B = 0.271, p = 0.015) and technological infrastructure (B = 0.219, p = 0.037) also have statistically significant positive effects, highlighting the importance of digital capacity, and Perceived Economic Benefits have a statistically positive relationship with sustainable business practices.

V CONCLUSION

The present study focuses on the adoption of Artificial Intelligence for the Sustainability of Small Businesses in Jaffna. Among the predictors, the perceived economic benefits (B = 0.395, p = 0.001) have a statistically positive relationship with business sustainability. Followed by Awareness of AI Adoption, which has a statistically positive relationship with business sustainability (B = 0.312, p = 0.004). Furthermore, Digital skills (B = 0.271, p = 0.015) and technological infrastructure (B = 0.219, p = 0.037) also have statistically significant positive effects, highlighting the importance of digital capacity. Additionally, perceived economic benefits have a statistically positive relationship with sustainable business practices. Based on this, it can be concluded that digital skills, infrastructure, and awareness can help businesses enhance efficiency and decision-making. The economic benefits perceived will result in encouraging more businesses to adopt AI for long-term growth.

REFERENCES

[1] Perifanis and F. Kitsios, "Investigating the influence of artificial intelligence on business value in the digital era of strategy: A literature review," *Information*, vol. 14, no. 2, p. 85, 2023.

[2] O. Badmus, S. A. Rajput, J. B. Arogundade, and M. Williams, "AI-driven business analytics and decision making," *World J. Adv. Res. Rev.*, vol. 24, no. 1, pp. 616–633, 2024.

[3] S. L. Wamba-Taguimdje, S. F. Wamba, J. R. K. Kamdjoug, and C. E. T. Wanko, "Influence of artificial intelligence (AI) on firm performance: The business value of AI-based transformation projects," *Bus. Process Manag. J.*, vol. 26, no. 7, pp. 1893–1924, 2020.

[4] H. M. T. R. Hearth, M. J. C. Fernando, W. M. C. P. Wanninayake, and A. M. A. Ahamed, "Current challenges and opportunities for entrepreneurship and SME's development in post-conflict areas," *Journal name or publisher pending*, 2022.

[5] M. Al-Emran and C. Griffy-Brown, "The role of technology adoption in sustainable development: Overview, opportunities, challenges, and future research agendas," *Technol. Soc.*, vol. 73, p. 102240, 2023.

[6] F. Fajri, K. A. Perdana, D. U. Manurung, P. K. N. Dharmawan, and N. G. Dewi, "The role of early adoption of artificial intelligence in supporting the growth of micro and ultra-micro enterprises in Indonesia: Challenges and opportunities," *J. Akunt. Bisnis*, vol. 10, no. 2, pp. 133–143, 2025.

[7] I. J. Akpan, E. A. P. Udoh, and B. Adebisi, "Small business awareness and adoption of state-of-the-art technologies in emerging and developing markets, and lessons from the COVID-19 pandemic," *J. Small Bus. Entrep.*, vol. 34, no. 2, pp. 123–140, 2022.

[8] T. R. L. Sooriyakumaran and T. L. Vannarajah, "Financial reporting practices and performance of small medium enterprises (SMEs) in Jaffna District of Sri Lanka," *Journal name or publisher pending*, 2023.

[9] K. Zhong and L. Song, "Artificial intelligence adoption and corporate green innovation capability," *Finance Res. Lett.*, vol. 72, p. 106480, 2025.

[10] N. O. P. Uzoamaka, O. M. Eneh, and I. O. Anyahara, "The impact of artificial intelligence on sustainability and operational efficiency of small and medium scale businesses in South-East, Nigeria," *Eur. J. Res. Reflect. Manag. Sci.*, vol. 13, no. 1, 2025.

[11] A. N. Khan, "Artificial intelligence and sustainable performance: Role of organisational agility and environmental dynamism," *Technol. Anal. Strateg. Manag.*, vol. 37, no. 5, pp. 568–583, 2025

Exploring Sustainability-Driven Fintech Usage Intentions Among Gen Z in Sri Lanka

Luckshika Vallaven
Business Faculty
SLIIT CITY UNI
Colombo, Sri Lanka
sliit.luckshi@gmail.com

***Abstract*—With growing global environmental awareness, Generation Z (Gen Z) has emerged as a key demographic driving demand for sustainable, technology-enabled solutions. This concept paper, developed from an undergraduate research proposal, explores sustainability-driven intentions to use financial technology (FinTech) services among Gen Z in Sri Lanka. Drawing on the Theory of Planned Behavior (TPB), the study aims to examines and offer new insights on how attitudes, subjective norms, perceived behavioral control (PBC), and environmental concerns influence sustainability-driven Fintech usage intentions among Gen Z's in Sri Lanka. A mono-method quantitative design is proposed, with data to be collected from Gen Z individuals in Sri Lanka using purposive and snowball sampling. The data will be collected via an online questionnaire and analyzed using descriptive and inferential statistics.**

Keywords—FinTech, Sustainability, Gen Z, Sri Lanka, Theory of Planned Behavior

I. INTRODUCTION

In the rapidly growing world, there has been a significant shift in consumer behavior towards purchasing sustainable and eco-friendly products and services. . It was discovered that "67% of participants were willing to pay up to 41% more for sustainable products and that percentage increases among younger age groups." This trend is linked to the growing environmental awareness particularly about global warming and climate change among younger generations [1], [2].

FinTech, a term that has been coined by combining two words, i.e., "Financial" and "Technology" [3], has transformed how individuals manage money, i.e., investments, transfers, borrowings and savings [4]. According to the Financial Stability Board, FinTech is defined as the "technically enabled financial services innovation that may lead to new business designs, applications, processes or products with a relative impact on financial markets and institutions and the delivery of financial services" [5].

FinTech can be classified into nine various categories such as e-banking, online lending, crowd funding, digital currencies, mobile Point of Sale (mPOS) models, robo-advisors, personal financial management platforms, transaction and payment terminal platforms and InsurTech [6].

It is also identified that FinTech helps create a more sustainable world by making it easier to support and invest in environmental friendly projects. By using digital tools and platforms, people can fund green businesses and technologies that reduce pollution and carbon emissions [7].

While Robo-advisors help people make smarter, eco-conscious financial decisions, digital banking reduces paper and energy use, which is especially popular among the younger generation (Gen Z) [8]. Moreover, advanced technologies like blockchain, AI, and digital payments improve how companies report on their environmental impact, making things more transparent and trustworthy [9], [10], [11].

A recent study has found that when "compared to older generations such as Gen X and Baby Boomers, Gen Z, consumers (born between 1997 and 2012) show a significantly higher willingness to pay (WTP) for brands that are sustainable and inclusive, highlighting their stronger commitment to ethical consumption and socially aligned financial decisions [12]."

Studies have also found that Gen Z's prefer Fintech solutions due to convenience and ease of accessibility and that 75% of young individuals aged between 18–35 are engaged in investments through digital platforms. It is also revealed that Gen Z seek financial services that demonstrates a commitment to ethical practices and environmental responsibility. Thus, FinTech companies are increasingly integrating sustainable finance initiatives to attract Gen Z individuals [13], [14], [15], [16], [17].

Given that global challenges like climate change become more pressing and with Gen Z showing commitments toward sustainability. This study aims to explore the sustainability- driven FinTech usage intentions among Gen Z in Sri Lanka, using the Theory of Planned Behavior (TPB) as the guiding framework to understand the relationship between attitudes, subjective norms and perceived behavioral control (PBC) and the FinTech usage intentions. This study extends the TPB framework by

integrating environmental concerns, as a variable within the TBP and proposes that FinTech usage intentions may be influenced by environmental concerns.

Accordingly, the objectives of this study are: (1) to examine the influence of environmental concerns on the sustainability-driven FinTech usage intentions among Gen Z in Sri Lanka; (2) to analyze the role of TPB constructs such as attitudes, subjective norms, and perceived behavioral control in shaping FinTech adoption for sustainable practices; (3) to explore the combined effect of TPB constructs and environmental concern on Gen Z's behavioral intention toward sustainable FinTech use; (4) to provide insights into how sustainability values influence digital finance adoption among young consumers in a developing country context; and (5) to offer recommendations for FinTech developers and policymakers to promote sustainable FinTech usage among Sri Lankan youth.

II. Literature Review

A. Theoretical Framework: Theory of Planned Behavior

A study revealed that Theory of Planned Behavior

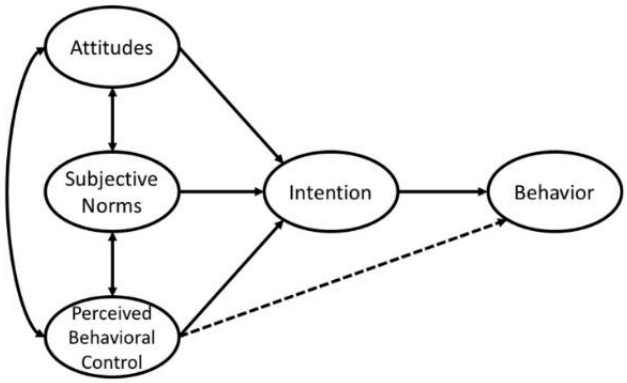


Fig. 1. Theory of Planned Behavior Model adapted from Ajzen 2005

(TPB) have demonstrated to be most successful in predicting the behavior maintaining the integrity of the specifications [18].

Therefore, TPB is a widely used in psychology and behavioral science to predict and explain human behavior in diverse fields including health, environment, marketing, and technology. TPB is an extension of the Theory of Reasoned Action that was developed by Icek Ajzen in 1985; it introduces the concepts of perceived behavioral control which reflects the confidence of the individuals to carry out a behavior. According to the theory, the intentions are influenced by attitudes, subjective norms, and perceived behavioral control [19]. Fig.1. presents the TPB model adapted from Ajzen 2005.

B. Review of Literature on Fintech Usage Intentions

Numerous studies have applied TPB, Technology Acceptance Model (TAM) or Unified Theory of Acceptance and Use of Technology (UTAUT) to study the fintech usage intentions, but only limited studies have done with a sustainability lens, particularly in Sri Lanka or Gen Z contexts intentions of FinTech.

A recent study investigated Fintech usage intentions in SAARC countries utilizing the Theory of Planned Behaviour (TPB). This study revealed that social influence, performance expectations, customer persistence intentions, and behavioral tendencies significantly influenced Fintech usage intentions, while subjective norms had a minimal impact. This contradiction does not invalidate TPB but instead suggests that subjective norms may vary contextually, a point this study aims to investigate further within the Sri Lankan Gen Z demographic [20].

Moreover, studies have utilized TAM, TPB or hybrid models to study Fintech usage intentions in Jordan and rural India. While they validated constructs like attitudes and TPB, none of them framed their models around sustainable usage intentions, nor did they focus exclusively on Gen Z, a demographic identified to be more environmental conscious [21], [22].

In the Sri Lankan context, a study examined digital payment adoption among the young population but used UTAUT and Perceived Risk Theory, with no consideration of sustainability dimension [23].

While prior studies offer valuable insights, they remain limited by their lack of focus on sustainability-driven usage, exclusion of environmental concern as a behavioral predictor, and minimal attention to the digitally fluent and value-driven Sri Lankan Gen Z demographic. Therefore, the proposed study aims to fill these gaps by addressing sustainability-driven FinTech use intention among Gen Z in Sri Lanka, including an extended variable in the form of environmental concern to the TPB model.

C. Conceptual Framework and Hypothesis

The Theory of Planned Behavior posits that behavioral intention is shaped by attitude, subjective norms, and perceived behavioral control. Attitudes refers to the feeling or thinking of a person about doing a certain action or behavior, it could be negative or positive, while social norm is the social pressure that might be pushing or pulling the individual to do that behavior or action and the perceived behavioral control is how the individual thinks or feels about how easy it would be to do that behavior or action. According to Ajzen's TPB, attitude, subjective norms, and PBC jointly determine behavioral intention. For example, a user with a positive attitude toward FinTech, supportive peers (subjective norms), and confidence in their ability to use the platform (PBC) is more likely to intend to adopt FinTech services [19].

Several studies have utilized TBP to identify the FinTech usage intentions and have demonstrated that

attitudes, subjective norms and PBC have positively impacted the intentions to use FinTech services. For instance, studies carried out in India, have revealed that attitudes, perceived behavioral control, and subjective norms together accounted for 48.7% of the variation in the intention to use fintech services among the young working population in the Indian context [24]. Moreover, studies conducted in Indonesia and Jordan also revealed that attitudes, subjective standards, and behavioral control had a substantial impact on interest in utilizing fintech [25], [26].

Furthermore, researchers have increasingly recognized the importance of including environmental concern to better understand sustainable behaviors, especially in the context of sustainable FinTech. For

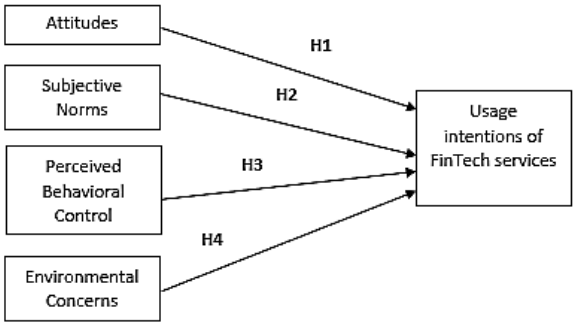


Fig. 2. Conceptual Framework

example, environmental concerns have shown to influence users continued engagement of green Fintech like Ant Forest [27].

Additionally, eco-incentives enhance environmental consciousness, encouraging sustainable actions [28]. Extended TPB models confirm that environmental concern positively affects behavioral intentions in sustainability contexts [29], [30], [31]. Since this study focuses on sustainability-driven intentions among Gen Z in Sri Lanka, adding environmental concern to TPB is relevant.

Therefore, this study may hypothesize the following:

H1: Attitude has an impact on the sustainability-driven fintech usage intentions of Gen Z in Sri Lanka.

H2: Subjective norms have an impact on the sustainability-driven fintech usage intentions of Gen Z in Sri Lanka.

H3: Perceived behavioral control has an impact on the sustainability-driven fintech usage intentions of Gen Z in Sri Lanka.

H4: Environmental concerns have an impact on the sustainability-driven fintech usage intentions of Gen Z in Sri Lanka.

Fig.2. illustrates the proposed conceptual framework; it serves as a visual summary of the theoretical assumptions and

hypotheses developed through the literature review.

III. METHODOLOGY

This study is explanatory and will employ a mono-quantitative method which aligns with conceptual focus of the research, where a self-administered questionnaire will be distributed online to collect data. The questionnaire will collect demographic information and include five-point Likert scale statements to measure key constructions derived from the Theory of Planned Behavior (TPB) and sustainability-driven fintech usage. Environmental concern will be measured using items such as “I am concerned about the impact of my financial decisions on the environment,” while sustainability-driven fintech usage intention will be assessed through statements like “I intend to use fintech services to reduce paper usage and carbon footprint.” TPB-related variables will be operationalized as follows: attitude (e.g., “Using fintech supports sustainable behavior”), subjective norms (e.g., “People important to me think I should use fintech for sustainability”), perceived behavioral control (e.g., “I am confident in using fintech sustainably”), and behavioral intention (e.g., “I plan to increase my use of fintech because of its environmental benefits”). The research aims to collect responses from 200 Gen Z individuals in Sri Lanka, obtained through a non-probability sampling method of purposive sampling and snowball sampling. A widely recognized software, SPSS will be used to analyze data, where statistical techniques such as descriptive and inferential statistics will be used. In particular, under inferential statistical techniques, correlation analysis will be used to identify relationships between variables, while a regression analysis will be used to assess the impact of the identified factors on the fintech usage intentions. These findings will be used to test the hypotheses and will be presented using a range of charts and tables.

IV. CONCLUSION AND RECOMMENDATIONS

In conclusion, this concept paper attempts to explore the influence of attitudes, subjective norms, perceived behavioral control, and environmental concerns, on sustainability-driven FinTech usage intentions of Gen Z in Sri Lanka. The proposed study would help extend the TPB model to include environmental concern as a factor affecting behavioral intentions when applied to the sustainable usage of FinTech among Gen Z in Sri Lanka, thereby, providing new insights to academia. The near-term implication of the findings, is that they offer important information to FinTech developers, policymakers, and marketers in framing platforms and campaigns that resonate with the Gen Z values of sustainability and their psychological drive. Nevertheless, the research contains a number of limitations. It targets a certain group, Gen Z population over 18 years in Sri Lanka, which makes the conclusions more limited in their application to other age groups or a different cultural environment. Moreover, the findings will be based on the responses received through the online survey, hence it may only reflect the individual perceptions and not actual behavior. Another limitation that

comes with the mono-quantitative design is that it does not allow it to become insightful to the extent that the qualitative techniques such as interviews or focus groups might provide. Furthermore, the non-probability sampling technique can undermine the generalizability of the findings, and the dynamically changing nature of the FinTech industry can make certain findings irrelevant with the passage of time. Possible directions to deal with these limitations in the future are: including other generational cohorts in the sample and making the comparisons international. By using mixed-methods, it would be possible to explore further behavioural drivers and contexts. Also, using probability sampling would bring representativeness and accuracy of results. As FinTech is still transforming, future research will be necessary to define new trends and user expectations in digital finance with a sustainability lens.

REFERENCES

[1] C. Chueca Vergara and L. Ferruz Agudo, "Fintech and Sustainability: Do They Affect Each Other?," Sustainability, vol. 13, no. 13, p. 7012, 2021, doi: 10.3390/su13137012.

[2] A. Iannuzzi, "Green Marketing," Informa, 2024, pp. 201–221. doi: 10.1201/9781003441939-10.

[3] R. Alt, R. Beck, and M. T. Smits, "FinTech and the transformation of the financial industry," Electron Markets, vol. 28, pp. 235–243, 2018, doi: 10.1007/s12525-018-0310-9.

[4] M. Asif, M. N. Khan, S. Tiwari, S. K. Wani, and F. Alam, "The Impact of FinTech and Digital Financial Services on Financial Inclusion in India," J. Risk Financial Manag., vol. 16, no. 2, p. 122, 2023, doi: 10.3390/jrfm16020122.

[5] "Fintech and Digitalisation of Financial Services: An Overview," ComFin Research, vol. 10, no. S1-Oct, pp. 20–24, Oct. 2022, doi: 10.34293/commerce.v10is1-oct.6144.

[6] J. Liu, X. Li, and S. Wang, "What have we learnt from 10 years of FinTech research? A scientometric analysis," Technol. Forecast. Soc. Change, vol. 155, p. 120022, 2020, doi: 10.1016/j.techfore.2020.120022.

[7] N. Omeragic, A. Zaimović, and T. Zaimović, "FinTech and Climate Action, and Affordable and Clean Energy," Procedia Computer Science, vol. 236, pp. 273–280, Jan. 2024, doi: 10.1016/j.procs.2024.05.031.

[8] B. Kuosuwan, A. Risman, E. Dudukalov, and E. Kozlova, "Digital banking and environmental impact: How Fintech supports carbon footprint reduction," BIO web of conferences, vol. 145, p. 05017, Jan. 2024, doi: 10.1051/bioconf/202414505017.

[9] T. Walker, H. J. Turtle, M. Kooli, and E. Nikbakht, Fintech and Sustainability. 2023. doi: 10.1007/978-3-031-40647-8.

[10] M. Vimal and R. Kumar, "Symbiotic Relationship of Fintech and Sustainability," IGI Global, 2024, pp. 106–124. doi: 10.4018/979-8-3693-3567-3.ch005.

[11] M. H. Hasan, M. Z. Hossain, L. Hasan, and M. A. Dewan, "FinTech and Sustainable Finance: How is FinTech Shaping the Future of Sustainable Finance?," European journal of management, economics and business., vol. 1, no. 3, pp. 100–115, Nov. 2024, doi: 10.59324/ejmeb.2024.1(3).09.

[12] F. Goedertier, B. Weijters, and J. Van den Bergh, "Are Consumers Equally Willing to Pay More for Brands That Aim for Sustainability, Positive Societal Contribution, and Inclusivity as for Brands That Are Perceived as Exclusive? Generational, Gender, and Country Differences," Sustainability, May 2024, doi: 10.3390/su16093879.

[13] I. Lestari, F. Fitriyani, N. A. Simanungkalit, and R. Sanjaya, "Pengaruh Financial Tecnology terhadap Manajemen Keuangan Generasi Z," Jurnal Rimba Riset Ilmu manajemen Bisnis dan Akuntansi, vol. 2, no. 4, pp. 82–89, Oct. 2024, doi: 10.61132/rimba.v2i4.1331.

[14] E. Novianta, A. Andani, and S. G. Pane, "Financial Technology Dan Literasi Keuangan Terhadap Generasi Z," Jan. 2024, doi: 10.47233/jebs.v4i1.1423.

[15] P. Yanida, O. H. Padmanegara, and T. I. Muklis, "The Intention to use Digital Investment Management Platforms Among Gen Z in Indonesia: Perspective from Individual and Technological Motives," Journal of economics, finance and management studies, vol. 08, no. 01, Jan. 2025, doi: 10.47191/jefms/v8-i1-02.

[16] M. L. Prasanna and A. L. PRIYANKA, "Marketing to Gen Z: Understanding the Preferences and Behaviors of Next Generation," International Journal For Multidisciplinary Research, vol. 6, no. 4, Aug. 2024, doi: 10.36948/ijfmr.2024.v06i04.26612.

[17] C. C. Vergara and L. F. Agudo, "Fintech and Sustainability: Do They Affect Each Other?," Sustainability, vol. 13, no. 13, p. 7012, Jun. 2021, doi: 10.3390/SU13137012.

[18] D. Li, L. Zhao, S. Ma, S. Shao, and L. Zhang, "What influences an individual's pro-environmental behavior? A literature review," Resour. Conserv. Recycl., vol. 146, pp. 28-34, 2019. doi: 10.1016/j.resconrec.2019.03.024.

[19] I. Ajzen, "The Theory of Planned Behavior," Organiz. Behav. Hum. Decis. Process., vol. 50, pp. 179-211, Dec. 1991. doi: 10.1016/0749-5978(91)90020-T.

[20] A. Priyadarshi, D. Prasad, and N. Kesari, "Determining FinTech User Behaviour and Intention in SAARC Countries Through Theory of Planned Behaviour Perspective," J. Commerce Account. Res., vol. 14, pp. 100–110, Apr. 2025, doi: 10.21863/jcar/2025.14.2.009.

[21] M. Al-Okaily, A. R. Al Natour, F. Shishan, A. Al-Dmour, R. Alghazzawi, and M. Alsharairi, "Sustainable FinTech Innovation Orientation: A Moderated Model," Sustainability, vol. 13, no. 24, p. 13591, 2021, doi: 10.3390/su132413591.

[22] K. Jena, "Factors Influencing the Adoption of FinTech for the Enhancement of Financial Inclusion in Rural India Using a Mixed Methods Approach," J. Risk Financial Manag., vol. 18, no. 3, p. 150, 2025, doi: 10.3390/jrfm18030150.

[23] Nasiketha, J. Tham, and A. Khatibi, “The Factors Influencing Youngsters’ Acceptance of Digital Payments in Sri Lanka,” Int. J. Prof. Bus. Rev., vol. 8, p. e02943, Oct. 2023, doi: 10.26668/businessreview/2023.v8i10.2943.

[24] R. Ravi and N. N. Pandey, “Intention to use fintech services: An investigation into the moderation effects of quality of internet access and digital skills,” Humanities and social sciences letters, vol. 12, no. 3, pp. 543–555, Jul. 2024, doi: 10.18488/73.v12i3.3803.

[25] J. Baharuddin, A. S. Supriyanto, S. Siswanto, and V. M. Ekowati, “Understanding The Drivers of Interest in Fintech Adoption: Examining The Moderating Influence of Religiosity,” Jurnal Aplikasi Bisnis dan Manajemen, Sep. 2023, doi: 10.17358/jabm.9.3.695.

[26] M. A. Al-Afeef, A. A. Alsmadi, and N. Alrawashdeh, “The Behavioral Intention of Fintech Usage: Applying Theory of Planned Behavior in Jordan,” Emerald Publishing Limited, 2024, pp. 15–25. doi: 10.1108/s1479-351220240000036002.

[27] Zhang, J.-C. Lee, T. Li, and L. Zhang, “Environmental concerns and green trust in continuous green fintech use: An expectation-confirmation model with insights from Ant Forest,” Humanit. Soc. Sci. Commun., vol. 12, no. 1, 2025. doi: 10.1057/s41599-025-04856-2

[28] S Lee, A. Sharabati, L. Almazaydeh, Q. Shalatony, R. Frangieh, and G. Al-Anati, “The impact of fintech-based eco-friendly incentives in improving sustainable environmental performance: A mediating–moderating model,” Int. J. Data Netw. Sci., vol. 8, no. 1, pp. 415–430, 2024. doi: 10.5267/j.ijdns.2023.9.013

[29] Kumar and D. Mohan, “Sustainable apparel purchase intention: collectivist cultural orientation and price sensitivity in extended TPB model,” J. Revenue Pricing Manag., vol. 20, no. 2, pp. 149–161, 2021. doi: 10.1057/s41272-021-00297-z

[30] Liu, S. Jin, H. Zhu, and X. Qi, “Construction of revised TPB model of customer green behavior: Environmental protection purpose and ecological values perspectives,” IOP Conf. Ser.: Earth Environ. Sci., vol. 167, no. 1, p. 012021, 2018. doi: 10.1088/1755-1315/167/1/012021

[31] Wang, S. Wang, Y. Wang, J. Li, and D. Zhao, “Extending the theory of planned behavior to understand consumers’ intentions to visit green hotels in the Chinese context,” Int. J. Contemp. Hosp. Manag., vol. 30, no. 8, pp. 2810–2825, 2018. doi: 10.1108/IJCHM-04-2017-0223

The Impact of Social Media Marketing on Generation Z's Purchasing Behavior in the Fashion Retail Industry

Roana Christy
Department of Business Administration
SLIIT City Uni roanachristy18@gmail.com

***Abstract*— This research paper explores how Social media advertising is a strong force of Generation Z consumers in the fashion retail industry. The study explores the influence of advertising, user-generated content, and AI-based recommendations on consumer purchase decisions. Based on secondary research studies, it identifies the key drivers of brand consideration and conversion. The study understands that top fashion brands like Zara and Shein effectively support micro influencers, interactive content, and one-to-one marketing, while traditional brands like Gap miss the consumer target. This study highlights the importance of originality, personal experiences, and sustainability messaging to engage Gen Z consumers. The results to brands are to extend micro-influencer partnerships, adopt AI driven content personalization, and step up interactive marketing. This study also provides practical outcomes to policymakers and industry stakeholders who wish to maximize digital marketing effectiveness and consumer interaction.**

Keywords—Social Media Marketing, Generation Z, Fashion Retail, Influencer Marketing, AI-driven Marketing (5 key words)

I. INTRODUCTION

Social media has also changed how Generation Z engages with fashion brands and makes their purchases. Generation Z, as a digitally born generation, that depends mostly on social media platforms like Instagram, TikTok, and YouTube for product information, brand engagement, and influencing their purchasing behavior. Social media has, also, forced fashion retailers to change their marketing strategies if they are to compete and succeed favorably in digital economies. This research was conducted because there is a growing trend regarding social media's impact on consumers, particularly among Generation Z. Early studies on markets reveal a drastic contrast among successful brands manipulating social media marketing and those losing share. Retailers such as Shein and Zara are at the frontline, milking influencers, interactive communications, and AI-driven personalized marketing for engaging with and for getting sales from Gen Z consumers. On the other hand, those with traditional marketing strategies, such as Gap, are recording low consumer engagement due to a lack of social media outreach (Capucao et al., 2024). Such a contrast reveals a core business issue why are there brands that succeed and those that do not succeed with Generation Z on social media? The purpose of this research is to examine social media marketing campaigns' impact on Generation Z's purchasing behavior within fashion retail business. Special focus is a discovery of how large marketing strategies such as influencer marketing, user-generated content, and AI-driven suggestions have impact, with a view toward identifying what triggers customers' engagement and loyalty most. For this purpose, this research will provide practical suggestions on how fashion retailers can best utilize social media marketing campaigns and gain competitiveness within a marketplace.

II. LITERATURE REVIEW

A. The Impacts of Social Media Marketing

Social media marketing has become a vital component in shaping consumer attitudes, especially among Generation Z. This generation, having grown up with the internet, spends considerable time on social media platforms like Instagram, TikTok, and YouTube, which serve not just as entertainment tools but also as vital platforms for brand presence. Smith (2021) highlights the importance of these online media in developing emotional connections between consumers and brands, hence driving brand loyalty and buying behavior. Chen et al. (2022) postulate that the interactive nature of social media captured by likes, comments, shares, and real-time engagement—allows for instantaneous communication, thus enabling brands to develop responsive and personal relationships. Williams and Page (2020) suggest that the perception of community and shared identity created on social media is a powerful factor in shaping the consumption patterns of Generation Z. Although previous research has explored brand interaction across social media platforms as a whole, this research offers a comparative study of Zara's interactive approaches versus less engaged brands like Gap, thus offering a more detailed perspective on the issue.

B. The Influencer Collaborations' Impacts

Influencer marketing, in the case of micro-influencers, is one highly effective method for targeting Generation Z. These influencers are seen to be more real and accessible compared to traditional celebrities, thus building higher levels of identification and trust (Lee & Kim, 2021). Consumers of Generation Z value peer-created content and real endorsements when making product judgments, and

thus micro-influencers are highly aligned with such values. Johnson (2022) posits that sponsored content by influencers can dramatically enhance both attention and conversion, especially when there is transparency and shared values by the influencer. The study by Thompson et al. (2023) explores how parasocial relationships defined by unidirectional emotional investment are amplifying influence recommendations by influencers, further blurring personal identification and commercial intent. The present paper builds upon such discussions by demonstrating that Zara's influencer marketing strategy clearly resonates with Generation Z values for authenticity and community, thus outmaneuvering competitors who resort to traditional or generic celebrity promotions.

C. Role of AI-Driven Marketing Strategies

The emergence of Artificial Intelligence (AI) is radically changing customer engagement with immediate and personalized marketing experiences. Tools like recommendation engines, prediction analysis, chatbots, and virtual fitting technologies are redefining the way Generation Z shops at fashion retailers. Brown and Singh (2022) assert that AI makes it possible to extract behavioral information to obtain actionable information to create content that resonates with individuality. Garcia and Patel (2021) demonstrate that AI-integrated features, such as personal shopping assistants, enhance online shopping by combining ease with interactivity. As compared to extant literature that has broadly discussed AI in e-commerce, this study explains that Zara uses advances in AI to boost loyalty and satisfaction for customers. The contrast with other companies that do not use such individualized tools highlights a wide performance gap that highlights Zara's data-driven adaptability.

D. Gen Z's Purchasing Behavior in the Fashion Industry

Generation Z consumer buying habits within the fashion sector The consumption habits of Generation Z are underpinned by values that encompass a dedication to environmental sustainability, social responsibility, and unique tendencies toward digital convenience. Nguyen et al. (2021) suggest that this generation shows a higher willingness to support brands that promote sustainability and uphold ethical standards. Alongside price and quality considerations, these consumers weigh a brand's compatibility with their cultural and environmental values. Ali and Hassan (2022) describe that Generation Z has a predilection for an omnichannel retail experience that successfully incorporates the benefits of online accessibility with brick-and-mortar shopping experiences. This insistence upon a seamless omnichannel strategy places considerable pressure on brands to optimize their digital infrastructures while maintaining operational transparency. Unlike older generations, which might value cost or fashion, Generation Z decision-making is marked by complexity and variance. By integrating these findings with comparative examinations of Zara and rival competitors, this research informs a more sophisticated understanding of Generation Z's fashion consumer trajectory.

III. METHODOLOGY

This research used a qualitative secondary research methodology to investigate social media marketing, influencer collaborations, and AI-driven strategies and their influence on Generation Z consumer habits in the fashion industry with a focus placed on Zara. The methodology was designed to ensure thoroughness of procedures used in collecting, evaluating, and synthesizing existing data across credible resources.

A. Methodological Framework and Research Strategy

An exploratory and descriptive research approach was adopted to determine strategic trends in the scope of digital marketing using Zara as the main case study. Zara was chosen due to its high status as a leading fashion retail brand, being an example of an innovative attitude towards digital marketing, which makes it a relevant example of successful interaction with Generation Z. Compared to more traditional or less successful rivals such as Gap or H&M, Zara continuously uses influencer marketing, artificial intelligence tools, and personalized content, thus allowing an in-depth understanding of advanced digital strategies.

Second data collected included published peer-reviewed journals, market analysis reports, case studies, and industry-specific literature that was published during 2018 to 2024. The data mainly focused on consumer attitudes and fashion advertising targeting Generation Z, who are those born between 1997 and 2012

B. Data Acquisition and Selection Standards

The sources were chosen according to the following criteria:

Relevance to Zara's social media marketing, influencer collaborations, or artificial intelligence integration.

Specific focus on Generation Z's consumer behavior.

Published in English between 2018 and 2024.

Provided qualitative or quantitative information about user interaction, brand image, or the effectiveness of online campaigns.

Relevant databases such as Statista, and Google Scholar were used. Official documents, the parent company of Zara, were also used.

C. Data Analysis

A thematic analysis was conducted to identify dominant patterns and strategies. The identified categories of themes included:

Social Media Marketing Tactics

Influencer Engagement and Authenticity

Artificial Intelligence in Personalization Factors Affecting Generation Z's Buying Decisions and Brand Loyalty

With the aim of enhancing analytical coherence and organizational clarity, a comparative framework was utilized to measure Zara against its competitors, such as Gap and H&M. Key metrics, such as engagement rate, conversion rate, visual strategy, and degree of AI personalization, were analyzed. A structured frame enabled findings to be combined into a model of engagement with Generation Z (see Table 1, Results).

D. Ethical Considerations

Since this research is based solely on publicly available secondary data, ethical approval was not required. The references are carefully documented according to recognized academic standards.

E. Limitations of Secondary Data

Whereas secondary data allowed comprehensive comparisons of trends and relative studies, it also had particular limitations. Secondary data are subject to publication bias, can fail to update in a timely manner, and might have the potential to ignore new developments in Generation Z's tastes. Moreover, the absence of primary data limits the ability to directly confirm behavioral subtleties from consumers' perspectives.

TABLE 1 COMPARATIVE FRAMEWORK OF BRAND STRATEGIES

Brand	Influencer Strategy	AI Use	Engagement Style
Zara	Micro influencers, Diversity	High (personalization, chatbots)	Visual, Interactive
Gap	Minimal, Traditional	Low	Basic Posts
H&M	Mid-tier influencers, Sustainability focus	Medium (recommendation engines)	Campaign based

IV. RESULTS AND EVALUATION

This section presents and evaluates the findings derived from the secondary data sources used in the research. The study explored the impact of social media marketing (SMM), influencer collaborations, and artificial intelligence (AI) driven strategies on Generation Z consumers, with a particular focus on Zara's marketing approaches within the fashion retail sector.

A. Social Media Marketing and Generation Z Engagement

Secondary data reveals that platforms like Instagram, TikTok, and Snapchat are central to how Gen Z discovers and interacts with fashion brands. Zara effectively uses visually driven, interactive content such as Reels, TikTok challenges, and stories to match Gen Z's preferences. Studies show that over 70% of Gen Z consumers seek fashion inspiration on these platforms, highlighting the importance of dynamic content in shaping buying decisions.

B. Influencer Marketing and Brand Trust

Influencer partnerships, particularly with micro-influencers, significantly enhance Zara's brand credibility. These influencers are seen as more authentic and relatable, contributing to stronger trust among Gen Z. Data shows that such partnerships lead to a 4x return on investment compared to standard digital ads, reinforcing their strategic importance. Zara's focus on diversity and sustainability in influencer selection further strengthens its appeal.

C. AI Personalization and User Experience

Zara's use of AI-driven features such as product recommendation engines and virtual assistants boosts personalization and convenience. These tools tailor shopping experiences based on browsing behavior, improving satisfaction and conversion rates. Case studies indicate that personalization increases brand loyalty among Gen Z, who expect technology to cater to their individual preferences

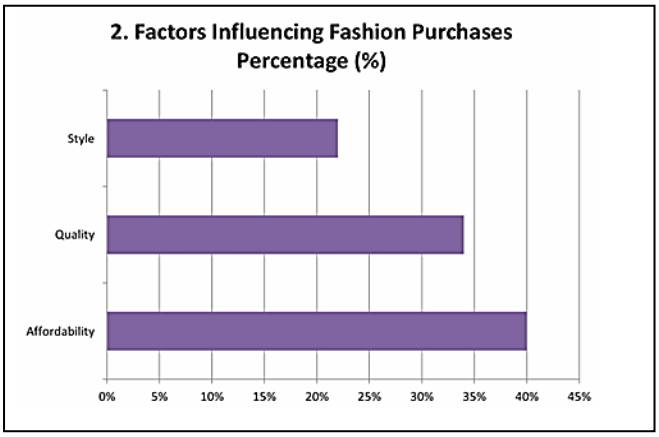
D. Integrated Marketing Effectiveness

Brands that combine social media presence, influencer marketing, and AI tools achieve higher engagement. Zara's holistic approach has led to measurable success in customer loyalty and sales. Compared to competitors like Gap—who often adopt fragmented strategies—Zara's integration enables stronger digital brand identity.

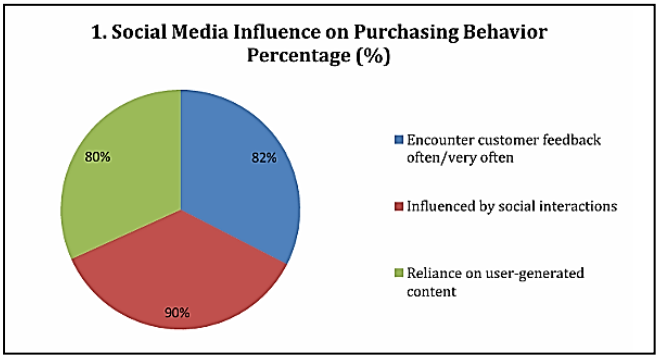
E. Limitations of Analysis

While secondary data provides valuable insights, it may not capture real-time shifts in consumer sentiment. The absence of primary data limits the ability to explore nuanced behavior or direct consumer experiences.

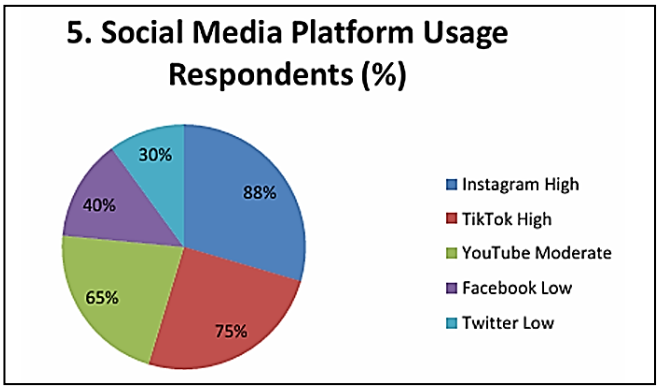
F. *Visual References*



Social Media Influence on Purchase Behavior

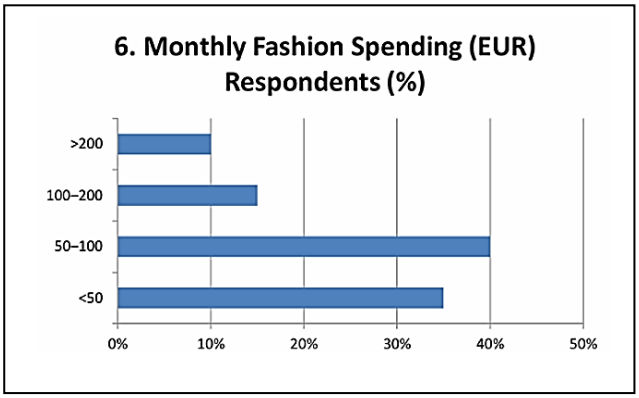


Factors Affecting Fashion Decisions



Gen Z Platform Usage

)



Monthly Fashion Spending of Respondents (in Percentage

V. CONCLUSION

This study highlights the growing impact of social media particularly Instagram on Generation Z's fashion purchasing behavior. Platforms that enable visual, interactive content have become essential tools for brand engagement. Influencer marketing, especially through micro-influencers, proves highly effective due to its authenticity and reliability. Zara's strategic use of diverse, values-aligned influencers encourages deeper consumer trust than more generic or celebrity-driven campaigns.

Artificial intelligence also plays a key role by delivering personalized experiences. Gen Z consumers expect tailored content and AI tools like recommendation engines and virtual try on address this need efficiently. These technologies enhance engagement, increase satisfaction, and improve conversion rates.

Furthermore, the research affirms that Gen Z's brand loyalty is strongly influenced by sustainability, diversity, and ethical considerations. Brands that fail to align their marketing practices with these values risk disengagement. Zara's alignment with Gen Z's ethical expectations contributes to its digital success.

However, as influencer marketing becomes more widespread, content saturation poses new risks. Over-commercialized promotions can reduce credibility and lead to consumer tiredness. Thus, brands must strike a careful balance between authenticity and paid content.

By integrating social media, influencer trust, AI personalization, and ethical branding, Zara represents how to effectively engage Generation Z. This study provides a comparative framework that can guide fashion retailers aiming to enhance relevance and loyalty in a digitally-driven, value-conscious market.

REFERENCES

[1] C C. B. Capucao, K. C. O. De Torres, S. J. Criman, and B. G. Lazaro, "Uncovering Gen Z's Styles: A Deep Dive into the

Consumer Behavior in the Fashion Industry," Int. J. Adv. Multidiscip. Res. Stud., vol. 4, no. 3, pp. 421–427, 2024.

[2] E. Djafarova and T. Bowes, "Instagram made me buy it: Generation Z impulse purchases in the fashion industry," J. Retail. Consum. Serv., vol. 59, p. 102345, 2021.

[3] E. Guisard and V. Barbe, Impact of Instagram on Generation Z consumer buying behavior within the fashion industry: A qualitative survey in the form of a questionnaire among Generation Z French consumers. Bachelor's thesis, JAMK Univ. Appl. Sci., Finland, 2023.

[4] J. Herzog, An investigation of the relationship between social media activity and impulsive buying behavior of Generation Z consumers in Austria. Master's thesis, Johannes Kepler Univ. Linz, Austria, 2023.

[5] J. Hjertquist and S. Jung, Effortless shopping: A qualitative study of Gen Z's shopping behavior and how they value environmental sustainability in the fashion industry. Linnaeus Univ., Kalmar, Sweden, 2020.

[6] V. D. Le Phuong, The Effects of Fashion Marketing in the Buying Behavior of Generation Z: Case Study of Ader Error Brand. Centria Univ. Appl. Sci., 2022.

[7] P. Li and S. H. Hassan, "Mapping the literature on Gen Z purchasing behavior: A bibliometric analysis using VOSviewer," Innov. Mark., vol. 19, no. 3, pp. 62–73, 2023.

[8] V. Merdzhanova, The influence of social media on buying decisions in the field of fast fashion. Master's thesis, Tech. Hochschule Ingolstadt, Germany, 2023.

[9] C. Nguyen, T. Nguyen, and V. Luu, "Relationship between influencer marketing and purchase intention: Focusing on Vietnamese Gen Z consumers," Indep. J. Manag. Prod., vol. 13, no. 2, pp. 810–828, 2022.

[10] S. D. Parmar, Understanding the extent of social media on Gen-Z consumers within the fashion industry: A focus on Zara, UK. MBA dissertation, Univ. Wales Trinity Saint David - London Campus, UK, 2024.

[11] L. Prini, What impact does exposure to social media have on Gen Z in Geneva, and how does it influence their buying attitudes in the clothing sector? Bachelor's thesis, Haute École de Gestion de Genève (HEG-GE), Switzerland, 2023.

[12] K. Zaka and M. M. S. Al-Juboori, The impact of Generation-Z on international branding strategy: An analysis of startup fashion brands. Bachelor's thesis, Linnaeus Univ., Sweden, 2023

Track 3: Electronic and Electrical Engineering

Computer Vision Controlled Humanoid Robotic Arm

M S Firdouse,
Department of Electronic and Electrical Engineering,
SLIIT City Uni, Colombo.
mohamedshuhaib38@gmail.com

L Benorith
Department of Electronic and Electrical Engineering,
SLIIT City Uni, Colombo.
benorith.l@sliit.lk

Abstract **- This paper presents the design and implementation of a low-cost, vision-based gesture-controlled humanoid robotic arm that mimics human hand and wrist movements in real time. The system uses a USB webcam and MediaPipe for hand landmark detection, OpenCV for image processing, and a Raspberry Pi 4 to compute landmark vectors and control servo motors via a PCA9685 driver. Calibration modes were introduced for each joint to ensure accurate servo mapping. The solution supports full gesture-based manipulation of a five-fingered robotic hand, including wrist orientation, with minimal latency and no physical contact. The system provides a more intuitive and natural method for robotic arm control compared to traditional input devices and has potential applications in prosthetics, automation, and human-robot interaction.**

Keywords **- Computer Vision, Robotic Arm, OpenCV, MediaPipe**

I. INTRODUCTION

Robotic arms are programmable mechanical manipulators designed to replicate the motions and functionalities of a human arm. These devices are typically structured as kinematic chains, consisting of rigid links connected by joints that allow controlled movement. The terminal part of the chain is referred to as the "end-effector," commonly designed as a gripper to interact with objects. Robotic arms have become essential tools in various industries, including automotive manufacturing, electronics assembly, medical surgery, hazardous material handling, and space exploration. One of the primary advantages of robotic arms is their ability to operate in environments that are dangerous, repetitive, or inaccessible to humans. A notable example is NASA's robotic exploration missions using the MER rovers, where robotic arms equipped with scientific instruments were remotely controlled to study Martian terrain. Traditional robotic arm systems are typically controlled using manual input methods such as keypads, joysticks, or pre-programmed instructions. While effective, these methods often require extensive calibration, technical expertise, and lack natural interaction. To overcome these limitations, alternative control paradigms such as voice commands, wearable sensor gloves, and vision-based gesture recognition have emerged. Among these, vision-based control systems stand out for their non-intrusive operation, enabling intuitive human–robot interaction without the need for physical contact or specialized wearables. Leveraging advancements in real-time image processing and machine learning, gesture recognition via standard cameras has become a viable solution for robotic control.

This paper presents a computer vision–driven humanoid robotic arm system that uses a USB camera, OpenCV, and MediaPipe to detect and interpret hand gestures in real time. The system maps human finger positions and wrist orientation to corresponding servo motor angles using a calibration-based approach. The entire control process is executed on a Raspberry Pi 4, ensuring portability and low cost, making the solution suitable for applications in prosthetics, education, and assistive robotics.

II. PROBLEM STATEMENT AND SCOPE

Despite the increasing use of robotic arms in industrial and assistive domains, existing control methods often depend on specialized hardware such as gloves, EMG sensors, or voice recognition modules. These systems suffer from limitations including hardware complexity, calibration overhead, user discomfort, and susceptibility to environmental noise. The problem this paper addresses is the lack of a low-cost, contactless, and easily deployable control method for humanoid robotic arms that can replicate complex human gestures in real time.

This project focuses on a vision-based approach that leverages MediaPipe and OpenCV to track hand gestures without requiring wearable devices. The scope is limited to finger and wrist movement replication using six degrees of actuation, without incorporating full-arm dynamics or grasp force feedback. This design makes the system suitable for educational, prototyping, and assistive use cases, but not yet optimized for industrial-grade manipulation tasks or high-speed dynamic control.

III. RELATED WORK

Robotic arm control has evolved from hardware-intensive, wearable systems to more flexible, vision-driven architectures. Early approaches used infrared emitters and glove-mounted sensors for gesture recognition, achieving fast response but requiring custom hardware and limited gesture vocabulary [1]. To improve accessibility, vision-only methods emerged. For example, RGB color segmentation was used to track hand motion via webcam input, enabling real-time control without wearables, though performance degraded under poor lighting or

varied skin tones [2]. Others employed compact vision modules like OpenMV to detect objects and apply inverse kinematics, but these systems faced processing limitations [3]. Sensor-based alternatives, such as gloves with flex sensors and accelerometers, offered high fidelity but reduced comfort and required calibration [4]. Voice-controlled arms also surfaced, using wireless microcontrollers and speech recognition, though often limited by noise and accent variability [5].

Biologically inspired methods like EMG-based control added intuitive prosthetic capability but required precise sensor placement and regular tuning [6]. More recently, computer vision systems using OpenCV and MediaPipe on embedded platforms have enabled robust, camera-based gesture recognition with minimal hardware. These systems have been used for real-time control and IOT-enabled remote actuation, though they remain sensitive to occlusion and network reliability. Other strategies include physical mimicry through linked miniature arms, and kinematic simulations of multi-DOF humanoid limbs. Image-based grasping techniques using PiCamera further show the role of embedded vision in object manipulation tasks.

Overall, these studies reflect a growing shift toward multimodal, software-centric solutions that aim to balance precision, cost, and usability in robotic arm control system

IV. System Architecture

Hardware Design

The proposed system consists of six MG995 servo motors (five for fingers, one for wrist), controlled by a PCA9685 servo driver via I2C. A 5V/3A external power supply provides consistent power for motors. A USB webcam captures video input, while the Raspberry Pi 4 Model B (8GB RAM) acts as the control core.

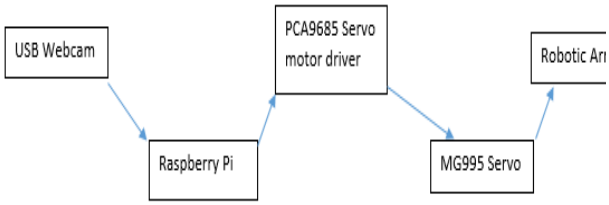


Figure 20: Hardware Design

Figure 1 shows the overall hardware flow of the project. The frames captured by the Webcam is sent to the raspberry pi for processing, that sends the relevant angles into the PCA9685 motor driver which in turn rotates the servo motors to required angles.

Figure 21: Final Robotic Arm

Software Design

Python-based software is used, leveraging OpenCV for video capture and pre-processing, and MediaPipe for hand landmark detection. The `cvzone` wrapper facilitates simpler landmark access. The Adafruit `ServoKit` library controls servo outputs, and calibration profiles are stored in JSON files.

Functional Overview

- **Hand Tracking:** MediaPipe detects 21 hand landmarks with high accuracy and robustness to hand orientation.

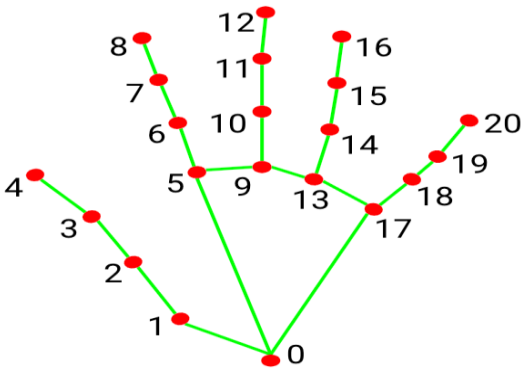


Figure 22: Hand Landmarks

- **Vector Analysis:** Joint angles are calculated using vector geometry and dot product techniques.

- **Servo Mapping:** Each finger angle is scaled to servo values using calibration bounds. Custom mappings allow adjustments for mechanical variance.
- **Calibration Mode:** Allows per-servo tuning for extended, bent, and scaling parameters, which are stored in JSON.
- **Failsafe and Shutdown**: atexit routines and interrupt handling ensure the servos are returned to neutral when tracking is lost or the program exits.
- **Live Feedback:** Visual indicators for hand detection and servo angle overlays are displayed in the video window.

V. GESTURE MAPPING AND CONTROL

Angle Calculation

For each finger, angles are calculated between landmark pairs: the fingertip and the proximal interphalangeal (PIP) joint. The angle is determined using the arctangent function from the vector formed by the landmark coordinates. The computed raw angle is then scaled using a user-defined calibration multiplier to account for hand variation. The system limits finger movement to a maximum of 90° for safety and realism.

Calibration Modes

Three interactive calibration modes are supported within the Python script:

- Bent Mode (b): Assigns the angle value for a fully curled finger position.
- Extended Mode (e): Sets the angle corresponding to a fully open finger.
- Scale Mode (s): Adjusts sensitivity by linearly scaling the mapped angle from detected vector motion.

Servo selection (1–6), mode switching (c to toggle), and real-time angle adjustments using arrow keys (left/right) are implemented in the OpenCV window interface. Calibrations are saved in a JSON file and persist across sessions.

Servo Actuation

Each calculated finger or wrist angle is clamped between 0° and 270°, the safe range of MG995 servos. The ServoKit interface from Adafruit is used to send PWM commands over I2C to the PCA9685 driver. On startup or loss of hand detection, the system resets all servo positions to the neutral 'extended' state using a pre-defined method. Servo smoothing and safety reset are ensured using exception handling and the atexit module.

System Operation Workflow

The system begins by initializing the USB webcam, loading calibration parameters from a locally stored JSON file, and setting up the MediaPipe hand tracking module with OpenCV. Upon capturing each frame, the image is processed to detect the presence of one or more hands. If a hand is identified, MediaPipe extracts 21 landmark points from the detected hand.

These landmarks are analyzed to calculate vector angles for each finger by comparing the fingertip and proximal joint coordinates. The raw angles are scaled and mapped to servo-compatible angles using pre-calibrated values. Each servo's final target angle is clamped to safe operational limits (0°–270°) and sent via I2C to the PCA9685 driver through the Adafruit ServoKit library.

In the absence of hand detection or if the frame is invalid, the servos are reset to a neutral, extended position to prevent erratic movement. Additionally, when the user exits the program or interrupts execution, the atexit handler ensures all servo motors return to safe default positions. The system is designed to run continuously and robustly under varying lighting and motion conditions.

VI. RESULTS AND ANALYSIS

A. Tracking and Responsiveness

The hand tracking pipeline was evaluated on the Raspberry Pi 4 with 8GB RAM. The average detection and servo update loop maintained 10–15 FPS, resulting in a real-time response latency of approximately 1s between gesture execution and servo motion. This responsiveness was consistent under stable indoor lighting and single-hand usage.

B. Hand Classification and Role Assignment

The system utilizes MediaPipe's built-in hand classification to differentiate between the user's right and left hand. During each frame capture, if two hands are detected, the library provides the type of attribute ("Right" or "Left") for each hand instance. This is essential for ensuring consistent control logic and servo mapping. In the current implementation:

- Only the right hand is used to control the robotic arm, as determined by ["type"] == "Right".
- If both hands are in the frame, the system prioritizes the right hand and ignores the left, avoiding conflicting gesture inputs.
- If no "Right" hand is detected, the system falls back to a neutral servo position for safety.

Table 17: Right- & Left-Hand Conditions

Condition	Outcome
Right hand only in frame	Full control enabled
Left hand only in frame	Ignored by control logic

Both hands in frame	Right hand selected for processing
Right hand leaves frame	System resets servos to neutral position

The hand identification was consistently reliable across test sequences, including fast switching between hands. Misclassification did not occur in controlled lighting. However, occasional flipping (e.g., left hand labeled as right) was observed when hands were rotated sideways or partially occluded. This effect was minimized by instructing users to present hands palm-forward to the camera.

C. Calibration Mode Performance

To ensure accurate gesture replication across different servo alignments and user hand styles, the system implements a fully interactive calibration mode. This feature enables real-time adjustment of each servo's range of motion and sensitivity, directly impacting servo precision and usability.

1) Calibration Effectiveness

Before calibration the tensioned strings couldn't bend the finger properly due to being too tight or too slack. After using the calibration mode in the bent and extended angles it was possible to keep the servo motor angle in the desired position and properly tension the strings to make the movement as effective as possible.

Furthermore, the scale calibration was also useful in making the error low, so that the motor would spin rapidly or slowly corresponding to the finger angle shown in the frame. It helped in making sure the fingers didn't get stuck halfway into bending or extending.

2) User Interaction and Usability

Calibration was conducted through a keyboard-driven interface, with visual feedback overlayed on the OpenCV video stream. Users were able to:

- Select servo channels (1 to 6)
- Choose calibration modes (b for bent, e for extended, s for scale)
- Adjust parameters incrementally using arrow keys

The calibration process was intuitive and fast, requiring under 2 minutes on average to fully configure all six servos.

3) Persistence and Reliability

All calibration parameters were saved automatically in a persistent JSON file. System testing confirmed that:

- Calibration values were retained across reboots.
- The atexit safeguard successfully wrote updated values even during abrupt terminations.
- Manual re-tuning was unnecessary unless servo positions or mechanical alignment changed.

4) Servo Response Accuracy

The servo system exhibited reliable and consistent behavior in replicating human hand gestures. After calibration, the finger and wrist movements closely matched user input, with noticeably smoother transitions and more natural motion curves. Repetitive gestures produced nearly identical servo responses across cycles, indicating stable actuation and mapping accuracy.

The system responded quickly to changes in hand posture, maintaining real-time synchronization between detected gestures and mechanical motion. Even during extended operation, the servos held their positions steadily without drift or jitter.

VII. CONCLUSION

The vision-controlled humanoid robotic arm in this project is implemented using Python 3, running on a Raspberry Pi 4 platform. The real-time video input is processed using OpenCV, while MediaPipe handles the landmark detection for finger and wrist gestures. The communication between the hand tracking system and servo controller is managed through the I2C interface connected to a PCA9685 module. Raspberry Pi is responsible for gesture recognition, landmark angle processing, and mapping to servo motion to avoid the limitations of microcontroller processing speed.

The PWM driver controls the movement of MG995 servo motors, and the motion response is stable and consistent across multiple gesture cycles. The servo signals can smoothly and accurately drive finger articulation and wrist rotation, confirmed through visual testing. During runtime, the system continuously tracks the user's hand, calculates joint angles, and sends appropriate commands to each servo channel to mimic hand posture.

In actual performance, the robotic arm was able to follow basic gestures such as hand opening, closing, and directional pointing. With the integration of calibration and gesture mapping modules, the system exhibits flexibility and adaptability for different users and mechanical configurations. The final build is low-cost, portable, and easy to set up, making it suitable for education, research, and assistive applications. In future iterations, the addition of force sensors or smart feedback modules can further enhance motion fidelity and bring the robotic arm closer to natural human interaction goals.

REFERENCES

[1] S. Bularka, R. Szabo, M. Otesteanu, and M. Babaita, "Robotic Arm Control with Hand Movement Gestures," in Proc. 2018 41st Int. Conf. on Telecommunications and Signal Processing (TSP), Athens, Greece, 2018, pp. 1–4.

[2] P. S. Lengare and M. E. Rane, "Human Hand Tracking Using MATLAB to Control Arduino-Based Robotic Arm," in Proc. 2015 Int. Conf. on Pervasive Computing (ICPC), Pune, India, 2015, pp. 1–6.

[3] J. Wang and H. Peng, “Object Grabbing of Robotic Arm Based on OpenMV Module Positioning,” in Proc. 2023 2nd Int. Conf. on Artificial Intelligence and Computer Information Technology (AICIT), Yichang, China, 2023, pp. 45–50.

[4] D. Crăciun, N. Bălăceanu, A. Chirită, and C. Ciobanu, “Robotic Arm Control via Hand Movements,” in Proc. 2022 Int. Symp. on Electronics and Telecommunications (ISETC), Timisoara, Romania, 2022, pp. 1–4.

[5] R. S. Kanash, S. E. Alavi, and A. A. Abed, “Design and Implementation of Voice-Controlled Robotic Arm,” in Proc. 2021 Int. Conf. on Communication & Information Technology (ICICT), Basrah, Iraq, 2021, pp. 22–27.

[6] P. S. Keerthi, S. Al Mamun, R. A. Tonima, S. A. Shanta, A. Ahmed, and M. S. R. Zishan, “Design and Implementation of a Human Prosthetic Hand,” in Proc. 2021 2nd Int. Conf. on Robotics, Electrical and Signal Processing Techniques (ICREST), Dhaka, Bangladesh, 2021, pp. 104–109.

AI-Based Smart Traffic Management System for Emergency Vehicles

D P S V Amarasinghe
,SLIIT City Uni, Colombo.
seshmi.v@gmail.com

, L Benorith
Department of Electronic and Electrical Engineering
SLIIT City Uni, Colombo.
benorith.l@sliit.lk

***Abstract* - Modern cities' main traffic congestion problem delays emergency vehicles like ambulances and firetrucks and police cars where every second counts. Fixed signal traditional traffic systems lack real-time adaptability, hence delays and risks are raised. This paper suggests an AI-driven smart traffic management system to priorities emergency vehicles and enhance general traffic flow by means of Raspberry Pi, YOLOv8, and OpenCV. Strategically positioned cameras provide video to a Raspberry Pi, which detects emergency vehicles by using OpenCV and YOLOv8. Dynamic control of traffic lights on detection helps to clear the path, so reducing response times and improving safety. The technology also maximizes road use and helps to ease traffic. For cities with limited infrastructure, using reasonably priced, open-source tools are scalable and ideal.**

Keywords - Emergency Vehicle Detection, Smart Traffic Management, YOLOv8, Raspberry Pi, Real-Time Signal Control.

I. INTRODUCTION

This research proposes an AI-powered smart traffic management system that significantly addresses the major issue of traffic delays for emergency vehicles. Rapid urbanization and increasing vehicles on the roadways are driving traffic congestion, a serious problem in modern cities. One important effect is the obstruction of emergency vehicles such as ambulances, firetrucks, and police cars whose response times are affected by static traffic controls. A few minutes of delay could cause major property damage or life-threatening circumstances [1]. Relying on fixed-timer signals that cannot adapt to real-time conditions is ineffective during emergencies. This underscores the urgent need for adaptive, real-time traffic management solutions that can dynamically prioritize emergency vehicles and ensure their uninterrupted passage through congested intersections. By integrating AI-driven object detection and automated signal control, the proposed system provides a proactive approach to saving lives and improving urban mobility efficiency.

A conventional traffic control system often forces emergency vehicles to wait at red lights or navigate through congested traffic without assured right-of-way. Especially during unexpected events like accidents or weather disturbances, the inflexibility of these systems does not fit the dynamic character of urban traffic. Recent developments in AI and image processing provide a confident answer to these barriers. Driven by AI, computers identify emergency vehicles depending on particular visual characteristics by examining live video feeds from strategically located cameras. The system can identify and give priority to emergency vehicles in real-time with great accuracy using object identification techniques such as YOLOv8 [2]. This allows dynamic traffic signal changes that open routes for emergency responders, hence enhancing safety as well as efficiency. Including such AI-based solutions into current infrastructure provides the way for better, adaptable urban transportation networks that improve emergency response, lower congestion, and help to create intelligent, data-driven smart cities.

Real-time AI-based identification and dynamic traffic signal regulation help the suggested solution to overcome these challenges. The system can autonomously spot emergency vehicles in traffic using high-definition cameras and object detection algorithms, hence avoiding human awareness [3]. Once detected, it automatically alters traffic signals to clear the path by stopping cross-traffic and turning lights green in the vehicle's direction. This guarantees a quicker and more reliable emergency response while reducing disturbance to normal traffic flow. The technology is intended to be accessible and scalable, so appropriate for large urban implementation. It provides a sensible approach to improve emergency services and preserve life.

II. HARDWARE DESIGN

The design integrates several critical components to effectively control traffic flow and prioritize emergency vehicles. The three key elements that define the system's architecture are the camera system, the traffic management and data processing unit powered by Raspberry Pi, and the pedestrian safety and sound detection module using ESP32 v2. The traffic management system operates through the camera system. Cameras are strategically positioned at key intersections to capture high-definition video feeds, providing clear visibility under various lighting

conditions. YOLOv8, an object detection algorithm, identifies emergency vehicles such as ambulances, fire trucks, and police cars as the cameras continuously stream video data to the Raspberry Pi for analysis. The cameras guarantee no emergency vehicle is missed by being positioned to track all lanes of the junction and offer thorough coverage of the area. The system's fundamental processing engine is the Raspberry Pi. It detects emergency vehicles by means of YOLOv8 after receiving the video data from the cameras.

The Raspberry Pi sets traffic signal system changes to give the emergency vehicle's passage top priority when it detects one. By sending signals to the traffic light controllers, it overcomes typical signal cycles to provide a clear road for the emergency vehicle, therefore controlling the timing and coordination of the traffic lights.

To minimize the impact on surrounding traffic, the Raspberry Pi ensures real-time decision-making by dynamically adjusting signal durations based on the type and approach speed of the emergency vehicle. Pedestrian safety along the emergency vehicle's route is managed by the ESP32 v2, which controls physical barriers such as servo motors to prevent pedestrians from crossing the road when an emergency vehicle is detected. Theunication system, the Raspberry Pi and ESP32 v2 interact to guarantee the coordinating of traffic control and pedestrian safety precautions.

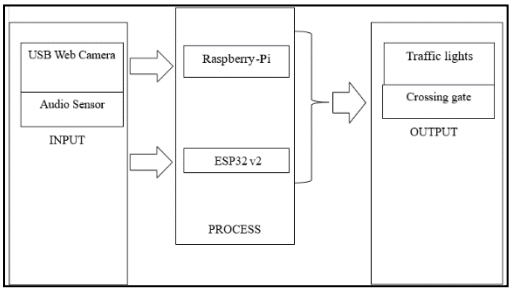


Figure 23: High Level Architecture

As shown in Fig. 1, starting with the input block, the system comprises an audio sensor and a USB web camera. While the audio sensor picks up sirens to verify their presence by sound, the camera records real-time video of traffic intersections to visually spot emergency vehicles. The processing unit of the system revolves on the Raspberry Pi. It manages all the computing tasks needed to real-time process the camera video feeds. Running advanced algorithms like YOLOv8 for object detection which is essential for recognizing emergency vehicles in the video stream the Raspberry Pi's computing capability enables it Using its GPIO pins, the Raspberry Pi interacts with other hardware parts including traffic signal controllers, modifying the traffic lights in response to an emergency vehicle identified. This traffic management system is ideal due to its adaptability, compact size, and cost-effectiveness. The Raspberry Pi enables easy integration of various sensors and devices, offering scalability for future enhancements and system expansions [4].

Direct connectivity between the Raspberry Pi and the camera system feeds real-time video data analysed for emergency vehicle detection. Usually high definition, the cameras can record detailed images even in different environmental situations including low light or bad weather. They are set to track important intersection sites and offer whole traffic flow coverage. These cameras transmit videos to the Raspberry Pi, which uses YOLOv8 to identify items in the image, classify them, and identify whether they are emergency vehicles.

The ESP32 v2 handles responsibilities relevant to pedestrian safety, therefore complementing the Raspberry Pi. It runs servo motors in charge of controlling pedestrian signals at the junction and barriers. The ESP32 v2 automatically lowers barriers to block pedestrian crossings and turns off the pedestrian walk signal to stop collisions when an emergency vehicle is detected. Additionally, it triggers sound alarms to alert nearby pedestrians of the approaching emergency vehicle. Through this coordinated interaction between the Raspberry Pi and ESP32 v2, the system ensures synchronized and safe operation, giving top priority to emergency vehicles while maintaining pedestrian safety. The process starts with turning on the sound sensor, which constantly searches for any arriving sound signals. The technology detects a relevant sound, say a siren, and activates the camera to begin taking live photos. Then, on a Raspberry Pi, the pictures are examined using YOLOv8 object detection to visually validate the presence of an ambulance. Should the ambulance be located, the traffic light is instantly changed from red to green to provide it top priority passage. Concurrent with this restriction of pedestrian access is gate or signal control through the ESP32 microcontroller. The mechanism lets the emergency vehicle pass for a set period of N minutes. The traffic signal goes back to red, the camera is switched off, and the system goes back to its starting state when the waiting time ends.

As shown in Fig. 2, the final product of this project is a fully functional smart traffic management system. Upon detecting an emergency vehicle, it dynamically adjusts traffic signals within two seconds to grant right-of-way, resulting in a 30–40% reduction in emergency response time compared to traditional systems. This significantly enhances the efficiency of emergency services and contributes to safer, smarter urban mobility. The software design of the AI-based smart traffic management system is defined by efficient real-time processing, accurate object detection, and adaptive traffic signal control.

Figure 24: Final Product

III. SOFTWARE DESIGN

The software architecture of the AI-based smart traffic management system is characterized by efficient real-time data processing, advanced object detection, and adaptive traffic signal control. Designed with a modular approach, the system ensures scalability, maintainability, and seamless integration of heterogeneous software components. Central to the framework is the YOLOv8 algorithm, optimized for real-time object detection, which addresses the critical challenge of accurately identifying emergency vehicles in complex traffic environments. The model is fine-tuned using custom datasets to improve detection precision.

During operation, YOLOv8 processes video input by partitioning each frame into a spatial grid, predicting bounding boxes and corresponding class probabilities to localize and classify objects. This enables precise detection and classification of emergency vehicles within the scene, which subsequently triggers real-time adaptive traffic control responses. Rigorous testing and validation under diverse environmental conditions ensure the system's robustness and reliability in real-world deployment [5]. To manage frame capture, preprocessing, and object annotation, the system utilizes OpenCV, a widely adopted library for image processing and computer vision applications. OpenCV ensures seamless integration of YOLO detection results with live video feeds, enabling real-time display of annotated frames with bounding boxes around detected vehicles. An intelligent traffic signal control system supports the dynamic adjustment of signal timings based on these detections.

When an emergency vehicle is detected, the program sends control signals to a traffic light controller managed by an ESP32 microcontroller, thereby prioritizing the vehicle's passage. This involves stopping cross traffic and granting a green signal along the emergency vehicle's route. Designed for real-time performance on edge devices such as the Raspberry Pi, the system is implemented in Python and leverages hardware interaction tools like OpenCV and YOLO. Effective frame management combined with parallel task execution ensures low latency in detecting and responding to emergency vehicles [6].

Emergency Vehicle Detection Using YOLOv8

The approach begins with real-time video analysis to detect emergency vehicles with high reliability. Designed specifically for emergency vehicle prioritization, the system incorporates strategically positioned high-definition cameras at critical intersections, along with advanced lighting units to support consistent video capture under varying environmental conditions. A Raspberry Pi 4 serves as the primary processing unit, managing visual input and executing the core detection algorithm. The system leverages YOLOv8, a cutting-edge object detection model known for its speed and precision, to accurately identify ambulances, fire trucks, police vehicles, and other emergency responders. The model has been fine-tuned to perform effectively under challenging conditions, including intense daylight, low-light environments, and partial occlusion, ensuring robust detection across real-world urban settings.

Figure 25: Emergency Vehicle Detection Results

Method of Detection:

- The real-time video stream consists of individual frames.
- Every frame comprises a grid.
- YOLOv8 searches every grid segment for objects according to predefined classification criteria.
- Should an emergency vehicle be identified, the system allocates bounding boxes to mark its position.
- The Raspberry Pi receives the detection findings including coordinates and classification for additional handling.

Prioritizing emergency vehicle detection, the Raspberry Pi ensures real-time processing capabilities. Upon identifying an emergency vehicle, the system initiates advanced traffic control measures. OpenCV enhances image quality and facilitates real-time video processing through filtering, edge detection, and color transformation

techniques essential for effective traffic monitoring under challenging conditions. Beyond preparation, it provides capabilities for improved traffic control including object monitoring, counting, and sensor data integration. This work detects emergency vehicles in live video streams using OpenCV and a YOLOv8 model, labelling each frame with bounding boxes. Practical and scalable, the Python software guarantees camera readiness, permits clean exits, and runs effectively on edge devices like the Raspberry Pi, thus it addresses urban traffic management.

Fine-tuned using a domain-specific dataset specified in data.yaml, which contains emergency vehicle classifications, the system employs a pre-trained lightweight model (yolov8n.pt) from the Ultralytics YOLOv8 library. To enhance model performance under real-world conditions such as poor lighting or congestion, training is conducted over 50 epochs with an input image size of 640 pixels. YOLOv8's efficiency and adaptability enable the model to run seamlessly on edge devices like the Raspberry Pi, facilitating affordable deployment in resource-constrained environments.

As shown in Fig. 3, beyond preprocessing the system provides capabilities for enhanced traffic control, including object monitoring, vehicle counting, and sensor data integration. It detects emergency vehicles in live video streams using OpenCV combined with the YOLOv8 model, labeling each frame with bounding boxes.

Dynamic Traffic Signal Adjustment Based on Detection Results

The system is engineered to dynamically manage traffic signals in response to the detection of an approaching emergency vehicle, thereby ensuring its safe and uninterrupted traversal through intersections. This functionality is governed by a Python-based decision-making algorithm running on a Raspberry Pi, which continuously processes real-time data such as the direction and position of the emergency vehicle, the status of intersection traffic signals, and prevailing traffic density and flow patterns.

Signal control is achieved through the algorithm's communication with the Raspberry Pi's General-Purpose Input/Output (GPIO) pins, which interface directly with the traffic light controllers. Upon detecting an emergency vehicle approaching a junction, the system overrides the default traffic light sequencing. It initiates an extended green phase for the lane corresponding to the emergency vehicle's path, while simultaneously issuing red signals to all conflicting lanes, effectively eliminating cross-traffic interference.

The timing and duration of signal transitions are dynamically adjusted based on key real-time parameters, including the emergency vehicle's velocity and estimated time of arrival at the intersection. Throughout this operation, continuous monitoring of the intersection is maintained to ensure the emergency vehicle's safe passage. Once the vehicle exits the intersection zone, the system seamlessly reverts to the default traffic signal cycle to minimize disruption and sustain optimal flow for general road users.

Moreover, the system's adaptive architecture supports real-time responsiveness to varying traffic conditions. Signal timing strategies are modified based on congestion levels. For instance, during peak traffic hours on highly congested roads, the system can temporarily reduce red signal durations to facilitate expedited emergency vehicle clearance restoring traffic balance more efficiently.

Pedestrian Safety Measures Using ESP32 Controlled Barriers

The intelligent traffic control system depends critically on pedestrian safety. The system must ensure pedestrians are kept safe when an emergency vehicle is recognized and the traffic light is changed from red to green, so it may pass without pausing. This is achieved with an ESP32 microcontroller running pedestrian safety systems including barriers. Servomotors reduce physical barriers to stop pedestrians from crossing the road if an emergency vehicle is identified when a pedestrian crossing is operating. Once the emergency vehicle passes, these barriers pull back so people may keep crossing securely.

IV. RESULTS

The AI-based smart traffic management system demonstrates substantial improvements in emergency vehicle prioritization, pedestrian safety, and overall traffic flow optimization. By leveraging state-of-the-art deep learning models integrated with real-time adaptive traffic signal control, the system enables intelligent decision-making based on dynamic road conditions. This integration enhances situational responsiveness, minimizes delays, and improves operational efficiency across intersections. Extensive validation in real-world scenarios confirms the system's reliability, scalability, and effectiveness under diverse and complex traffic environments.

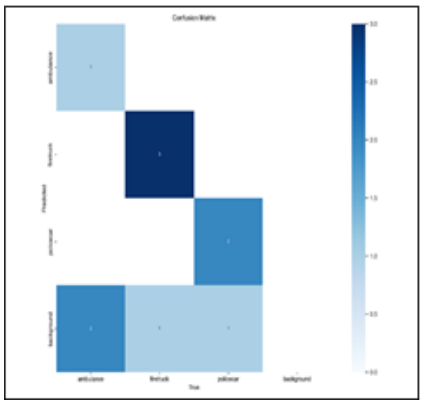


Figure 26: Confusion Matrix

As shown in Fig. 4, rigorously evaluated in real-world scenarios, the AI-based smart traffic management system has demonstrated its effectiveness in enhancing pedestrian safety, optimizing general traffic flow, and prioritizing emergency vehicles thereby reducing the risk of accidents and ensuring safe, uninterrupted passage. Simulations and live-field tests revealed an average reduction in emergency vehicle response delays by 30–40%, along with improved traffic fluidity and reduced congestion. The YOLOv8 object detection model was evaluated under various environmental conditions using key performance metrics, including precision, recall, and F1 score.

High precision values 0.75 for fire trucks and 0.87 overall for emergency vehicles guarantee accurate classification with few false positives according to a Confusion Matrix Normalized analysis, as shown in Fig. 5. While the Accuracy-

Recall Curve showed a nearly perfect mAP@0.5 of 0.995, highlighting the model's efficacy, the F1-Confidence Curve indicated a solid balance between accuracy and recall, with an F1 score of 0.96 at a confidence level of 0.129. The confusion matrix results highlight the capacity of the model in reducing false positives and false negatives, therefore supporting the accuracy of the system.

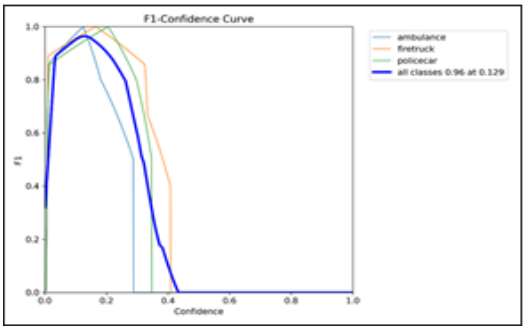


Figure 28: Confidence Curve

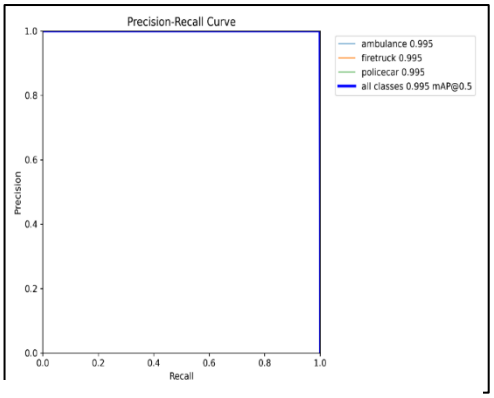


Figure 29: Precision-Recall Curve

Low rates of misclassification support the dependability of emergency vehicle detection. Furthermore, the Recall-Confidence Curve of the system shows its remarkable sensitivity; it reaches 1.00 recall at a confidence level of 0.000, therefore guaranteeing that no emergency vehicle passes underpass, as shown in Fig. 6. Furthermore, at a confidence level of 0.231, accuracy stayed at 1.00 for all classes, also signal changes are set off only in confirmed detections.

Detection accuracy is the low false positive rate (0.07) reduced needless signal overlays, a major improvement over conventional motion-based or sound-based detection systems that sometimes cause false alarms because of ambient noise or reflections. Traffic Signal Responsiveness in system reaction time and efficiency informed real-time traffic signal adjustments. This AI-based technology dynamically changed signals within 2 seconds, this shortened emergency vehicle travel time by up to 40% is very important. Pedestrian and Traffic Safety in system's 98% barrier deployment rate along with audio-visual alarms and immediately signal cut-offs produced a 70% drop in near-miss events. The 90% compliance rate among pedestrians indicates even more how well the warning systems stop dangerous crossing.

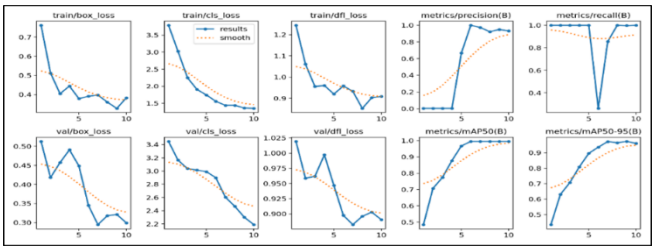


Figure 27: Performance Evaluation Graph

Field tests revealed that, once an emergency vehicle was detected, signal changes occurred within two seconds, reducing emergency response times by 30–40% compared to conventional traffic systems. Additionally, 85% of partially obscured emergency vehicles were accurately identified, ensuring timely interventions even in complex urban environments. Object localization achieved accurate barrier activations in 98% of cases, enhancing pedestrian safety. Pedestrian signal cut-off was triggered within one second of detection, as illustrated in Fig. 7. Multi-sensory alerts, including LED flashers and loudspeakers, improved pedestrian compliance to 90%, contributing to a 70% reduction in near-miss incidents. The performance evaluation graph confirms that the model achieved high accuracy without overfitting, demonstrating strong generalization across varying traffic conditions.

V. CONCLUSIONS

The system has significant restrictions even with these developments. It primarily relies on visual detection through cameras, which can underperform in adverse weather conditions such as heavy rain, dense fog, or snow, where visibility is severely reduced. To address both emergency and routine traffic scenarios, advanced

machine learning models can be employed to predict traffic flow and dynamically optimize signal timings, thereby reducing overall congestion.

This underscores the need for future enhancements, such as upgrading to more powerful AI processors and integrating LiDAR or radar sensors. The system has already demonstrated promising outcomes, including a 70% reduction in pedestrian near-miss incidents and up to a 40% decrease in emergency response times. These results emphasize the transformative potential of artificial intelligence in traffic management, particularly in densely populated urban areas where rapid decision-making is critical.

To further enhance scalability and intelligence, incorporating Vehicle-to-Infrastructure (V2I) communication and integration with broader Intelligent Transportation Systems (ITS) is recommended. While challenges remain, this research establishes a strong foundation for next-generation smart city applications by combining advanced computer vision with real-time responsiveness to improve road safety and efficiency.

REFERENCES

[1] M. Sweet, "Does traffic congestion slow the economy?" Nov. 2011. doi: 10.1177/0885412211409754.

[2] Y. Liu, C. Yang, and Q. Sun, "Thresholds Based Image Extraction Schemes in Big Data Environment in Intelligent Traffic Management," IEEE Transactions on Intelligent Transportation Systems, vol. 22, pp. 3952–3960, Jul. 2021, doi: 10.1109/TITS.2020.2994386.

[3] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," Apr. 2018.

[4] D. T. Mane, S. Sangve, S. Kandhare, S. Mohole, S. Sonar, and S. Tupare, "Real-Time Vehicle Accident Recognition from Traffic Video Surveillance using YOLOV8 and OpenCV," International Journal on Recent and Innovation Trends in Computing and Communication, vol. 11, pp. 250–258, May 2023, doi: 10.17762/ijritcc.v11i5s.6651.

[5] S. Gupta, "Image Processing Based Intelligent Traffic Control System by Using Raspberry Pi," Int J Res Appl Sci Eng Technol, vol. 6, pp. 66–70, Apr. 2018, doi: 10.22214/ijraset.2018.4014.

[6] B. Kamasetty, M. Renduchintala, L. L. Shetty, S. Chandarshekar, and R. Shettar, "Design and development of portable smart traffic signaling system with cloud-artificial intelligence enablement," Indonesian Journal of Electrical Engineering and Computer Science, vol. 26, pp. 116–126, Apr. 2022, doi: 10.11591/ijeecs.v26.i1.pp116-126.

Smart Sorting and Grading Fruits based on Image Processing Techniques

A J S Ahamed,
Department of Electronic and Electrical Engineering, SLIIT City Uni, Colombo.
sathath000@gmail.com

L Benorith
Department of Electronic and Electrical Engineering, SLIIT City Uni, Colombo.
benorith.l@sliit.lk

Abstract - **This paper presents the design and implementation of an automated apple sorting system that integrates machine vision techniques with embedded control for real-time classification and sorting of apples. The system employs a Raspberry Pi 4 as the primary processing unit, using a YOLOv11 model for fruit detection and classification, while an Arduino Nano manages weight measurement via a load cell. Real-time images of apples on a conveyor belt are captured, processed, and classified into four categories: Good Red, Good Green, Bad Red, and Bad Green. Sorting mechanisms, including servos and actuate based on classification results, with an integrated LCD and cloud-based Google Sheets providing monitoring and logging. The system demonstrates high classification accuracy and reliable sorting performance, offering a cost-effective solution for small to mid-scale agricultural applications.**

Keywords - Fruit Grading, Machine Vision, Deep Learning, Convolutional Neural Networks

I. INTRODUCTION

Automated sorting systems have become essential components of modern warehousing and distribution operations due to their ability to manage products with speed and precision in high-volume environments [1]. Unlike manual sorting methods—which are labor-intensive, time-consuming, and error-prone—automated systems convert random product flows into structured shipping orders based on pre-set criteria, significantly enhancing operational efficiency [2]. Early systems relied heavily on human workers, leading to increased costs and frequent sorting errors such as misplacements and inefficiencies [3]. The introduction of barcode scanning technology marked a substantial improvement, enabling automatic product recognition and sorting with throughput capacities of up to 10,000 items per hour [4]. However, this approach was limited by its dependency on barcode labels, which are not feasible for items like fruits and vegetables that often lack standardized packaging.

To overcome this limitation, modern systems have evolved to include additional features such as weight, size, and color analysis for handling unlabeled or irregular items [5]. This is particularly beneficial in agricultural applications, where visual quality plays a critical role in product valuation and consumer acceptance. Early machine vision systems used basic shape and color analysis, but these methods were sensitive to environmental conditions such as lighting variations and product inconsistencies. Improvements in image processing—such as edge detection, segmentation, and clustering—helped address these issues, allowing for more reliable performance in real-world environments [6].

With the rise of artificial intelligence, especially deep learning models like YOLO (You Only Look Once), sorting systems have achieved even higher levels of speed and accuracy [7]. These models are capable of real-time object detection and classification, performing well even in complex, cluttered, or variable conditions. YOLO-based systems have demonstrated remarkable success in agricultural scenarios, including fruit identification, ripeness estimation, and defect detection surpassing the capabilities of traditional image processing techniques [8]. Despite their promise, integrating these AI models into fully functional hardware systems remains a challenge. Effective deployment requires careful coordination between the software (e.g., object detection algorithms) and hardware (e.g., conveyor belts, solenoids, robotic arms) [9].

This study presents a smart apple sorting system that addresses these integration challenges by combining deep learning-based apple recognition (using YOLOv11) with mechanical actuation components such as servos and solenoids, managed by an Arduino Nano and Raspberry Pi 4. The system demonstrates reliable real-time classification of apples based on visual features, offering a cost-effective and scalable solution for small to medium-sized agricultural operations.

II. DESIGN METHODOLOGY

The proposed system integrates machine vision, embedded hardware, and mechanical actuation to classify and sort apples in real time based on their color and quality. The design objectives focus on achieving accurate classification, efficient sorting, and reliable weight monitoring using cost-effective components suitable for small-scale agricultural settings.

A. System Design and Hardware Integration

The apple sorting system comprises a conveyor-based mechanical structure that transports apples sequentially

under a camera for inspection and then towards a sorting zone. The main components include:

1) A Raspberry Pi 4 (8GB RAM) for image processing and classification.

2) A Logitech 1080p USB webcam positioned above the conveyor to capture images of apples.

3) Two MG996R servo motors for physical sorting based on classification results

4) Two E18-D80NK infrared (IR) proximity sensors for apple detection and position tracking.

5) An Arduino Nano connected to an HX711 amplifier and load cell for real-time weight measurement.

6) A 16x2 I2C LCD display to show apple classification and weight data.

7) A dedicated 5V 5A DC adapter for powering servo motors and sensors, ensuring stable operation without affecting Raspberry Pi performance.

The conveyor system is designed with a plywood base, PVC belt, and side partitions to prevent apple displacement. Apples are placed manually onto the conveyor, and the system operates continuously with real-time detection, sorting, and logging.

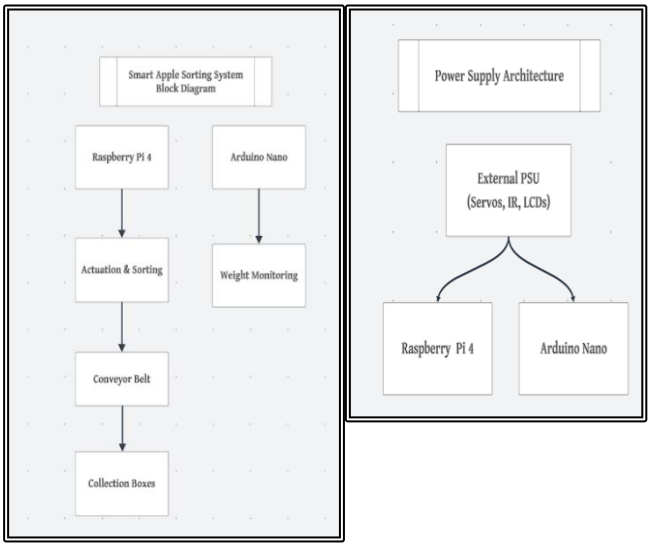


Fig.1.Block diagram of smart sorting and power supply

B. Implement image recognition algorithm

1) *Prepare input data*

A dataset of approximately 1,000 high-resolution images was created to train the YOLOv11 model. Apples were categorized into four classes based on color and quality: Good Red, Bad Red, Good Green, and Bad Green. Images were captured under varied lighting conditions and angles to enhance model robustness. Annotation was performed using the Roboflow framework, generating bounding boxes for each apple instance.

The annotated dataset was split into training (70%), validation (20%), and testing (10%) subsets. Preprocessing techniques, including brightness adjustment, contrast enhancement, and blurring, were applied to simulate real-world imaging variations and improve generalization. The results of collecting input images are described in Fig. 2

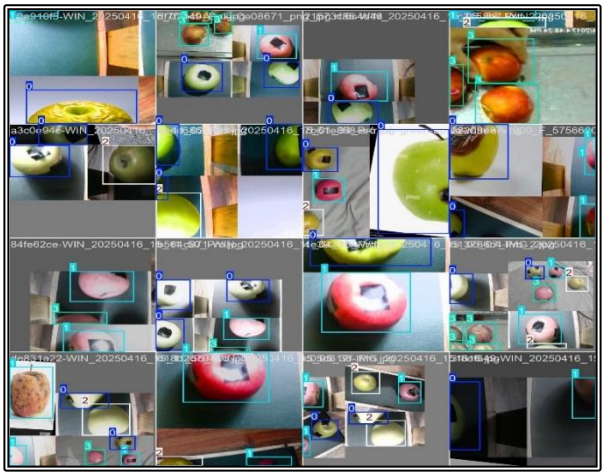


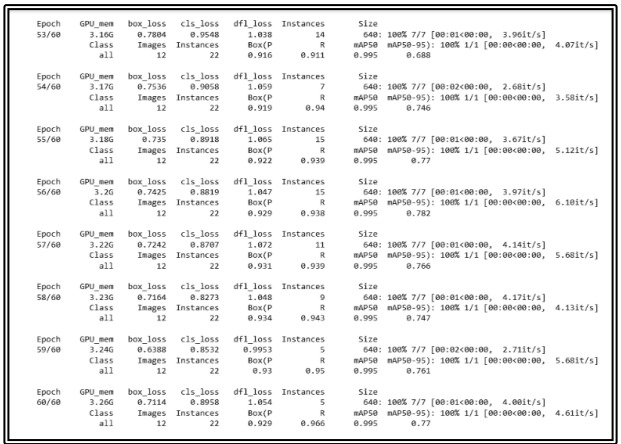
Fig.2.Labelling image

2) Model training and Evaluation

With large input data, training a neural network model for the purpose of recognizing any object on a low-configuration device is very difficult. The training process is time consuming and training cannot be sustained over a long period of time. Therefore, the Google Colab support tool is used in this case. The steps to train the model on Google Colab are as follows:

Step 1: Import the PyTorch library that supports training.
Step 2: Reconstruct the YOLO11 network structure
Step 3: Import the dataset after labeling and preprocessing The images via Roboflow.
Step 4: Prepare the image data file and labeled labels
Step 5: Convert labeled layers to named layers (red, red-bad, green, green -bad).
Step 6: Conduct model training

After training for 60 Epochs, the P (Precision) and R (Recall) indexes both reached a relatively high level of ~0.9 and the mAP (Average Precision) index ~0.97 (Fig. 3)



```
Epoch    GPU_mem   box_loss   cls_loss   dfl_loss  Instances       Size
53/60      3.16G     0.7804     0.9548      1.038         14        640: 100% 7/7 [00:01<00:00,  3.96it/s]
           Class     Images  Instances      Box(P          R      mAP50  mAP50-95): 100% 1/1 [00:00<00:00,  4.07it/s]
             all         12         22      0.916      0.911      0.995      0.688

Epoch    GPU_mem   box_loss   cls_loss   dfl_loss  Instances       Size
54/60      3.17G     0.7536     0.9058      1.059          7        640: 100% 7/7 [00:02<00:00,  2.68it/s]
           Class     Images  Instances      Box(P          R      mAP50  mAP50-95): 100% 1/1 [00:00<00:00,  3.58it/s]
             all         12         22      0.919       0.94      0.995      0.746

Epoch    GPU_mem   box_loss   cls_loss   dfl_loss  Instances       Size
55/60      3.18G      0.735     0.8918      1.065         15        640: 100% 7/7 [00:01<00:00,  3.67it/s]
           Class     Images  Instances      Box(P          R      mAP50  mAP50-95): 100% 1/1 [00:00<00:00,  5.12it/s]
             all         12         22      0.922      0.939      0.995       0.77

Epoch    GPU_mem   box_loss   cls_loss   dfl_loss  Instances       Size
56/60        3.2G     0.7425     0.8819      1.047         15        640: 100% 7/7 [00:01<00:00,  3.97it/s]
           Class     Images  Instances      Box(P          R      mAP50  mAP50-95): 100% 1/1 [00:00<00:00,  6.10it/s]
             all         12         22      0.929      0.938      0.995      0.782

Epoch    GPU_mem   box_loss   cls_loss   dfl_loss  Instances       Size
57/60      3.22G     0.7242     0.8707      1.072         11        640: 100% 7/7 [00:01<00:00,  4.14it/s]
           Class     Images  Instances      Box(P          R      mAP50  mAP50-95): 100% 1/1 [00:00<00:00,  5.68it/s]
             all         12         22      0.931      0.939      0.995      0.766

Epoch    GPU_mem   box_loss   cls_loss   dfl_loss  Instances       Size
58/60      3.22G     0.7164     0.8273      1.048          9        640: 100% 7/7 [00:01<00:00,  4.17it/s]
           Class     Images  Instances      Box(P          R      mAP50  mAP50-95): 100% 1/1 [00:00<00:00,  4.13it/s]
             all         12         22      0.934      0.943      0.995      0.747

Epoch    GPU_mem   box_loss   cls_loss   dfl_loss  Instances       Size
59/60      3.24G     0.6388     0.8532     0.9953          5        640: 100% 7/7 [00:02<00:00,  2.71it/s]
           Class     Images  Instances      Box(P          R      mAP50  mAP50-95): 100% 1/1 [00:00<00:00,  5.68it/s]
             all         12         22       0.93       0.95      0.995      0.761

Epoch    GPU_mem   box_loss   cls_loss   dfl_loss  Instances       Size
60/60      3.26G     0.7114     0.8958      1.054          5        640: 100% 7/7 [00:01<00:00,  4.00it/s]
           Class     Images  Instances      Box(P          R      mAP50  mAP50-95): 100% 1/1 [00:00<00:00,  4.61it/s]
             all         12         22      0.929      0.966      0.995       0.77
```

Fig.3.train model epoch in google colab

Fig. 4 shows the loss value after training and validation gradually decreasing over epochs. The above indexes are almost equal to 0 at epoch 60. The P and R indexes both reach a relatively high level of ~ 1.0i and the mAP index ~ 1.0.

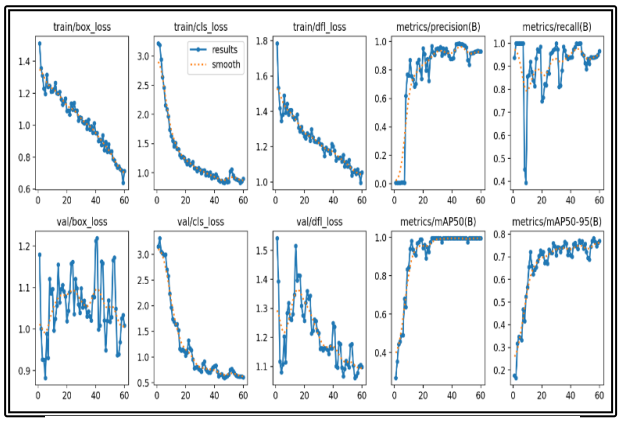


Fig.4.mAp and train loss level showing graph

C. *Integrate image processing techniques in the classification system*

The main flowchart begins when an apple moves on the conveyor and is captured by a USB webcam connected to the Raspberry Pi. The YOLOv11 model running on the Pi classifies the apple, and the classification result is immediately shown on an LCD display. Simultaneously, this data is logged to a Google Sheets document for remote monitoring and historical record-keeping. The decision node then evaluates whether the apple is classified as "Good" or not. If it is bad, no sorting action is taken, and the apple simply falls into the bad apple box. If the apple is good, another decision determines its color red or green. Based on this classification, the system triggers the corresponding IR sensor to activate a servo motor, which physically deflects the apple into the appropriate sorting box.

The Arduino Nano subsystem independently monitors the weight of apples using a load cell and HX711 amplifier. It calculates the total weight and displays it on an LCD screen. This runs continuously without affecting Raspberry Pi's sorting operations. The separation of these tasks improves system reliability and scalability.

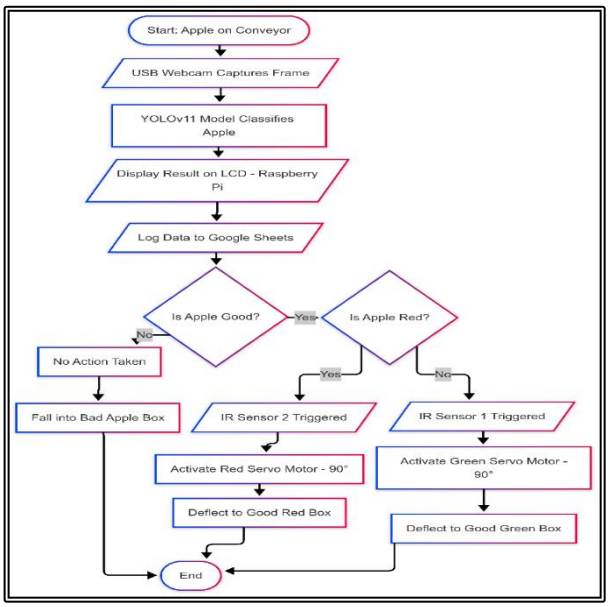


Fig.5. Flow Chart of system

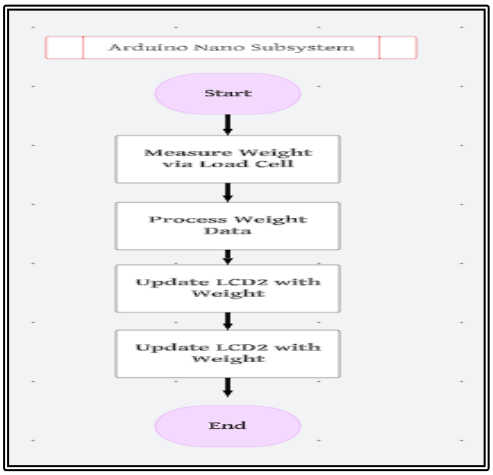


Fig.6.Weight analyzing system

III. RESULTS AND EVALUATION

The apple sorting system prototype was assembled using cost-effective components for real-time classification and sorting. The mechanical frame features a PVC-based conveyor belt measuring 80 mm in width, suitable for medium-sized apples (50–75 mm diameter). Guard rails on both sides prevent apple displacement, and the conveyor is driven at a steady 0.1 m/s for optimal camera capture and sorting precision. The belt speed can be adjusted as needed for higher throughput.

The core processing unit is a Raspberry Pi 4 (8GB RAM), which performs real-time image acquisition and classification using a YOLOv11-based detection model. A Web Camera mounted directly above the conveyor captures top-view images of incoming apples. Upon detection, each apple is classified into one of four

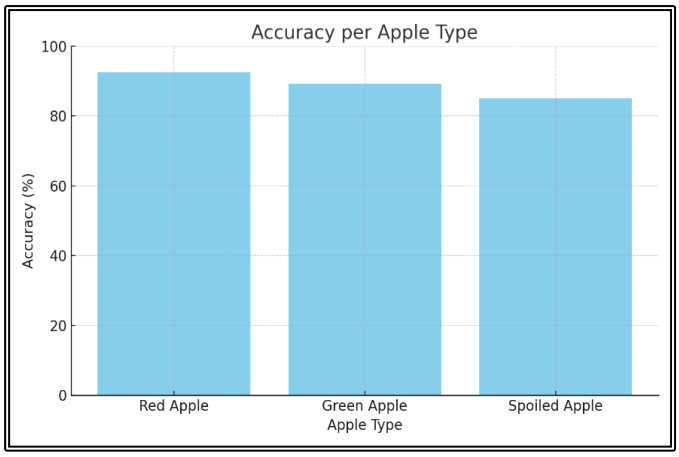


Fig.8. Training Accuracy and Loss Graph

categories: Good Red, Red-Bad, Good Green, or Green-Bad. The classification results, along with the counts on a 16x2 I2C LCD screen connected to the Raspberry Pi. Raspberry Pi also handles sorting logic and sends control signals to actuate servo motors based on classification outcomes.

In addition to visual sorting, the system incorporates real-time logging via cloud-based infrastructure. Classification results are transmitted directly to a Google Sheets spreadsheet using Python's gspread and oauth2client libraries on the Raspberry Pi. This feature enables remote tracking and batch analysis of sorting accuracy and throughput, as illustrated in Fig. 9 (Google Sheets Logging Output).

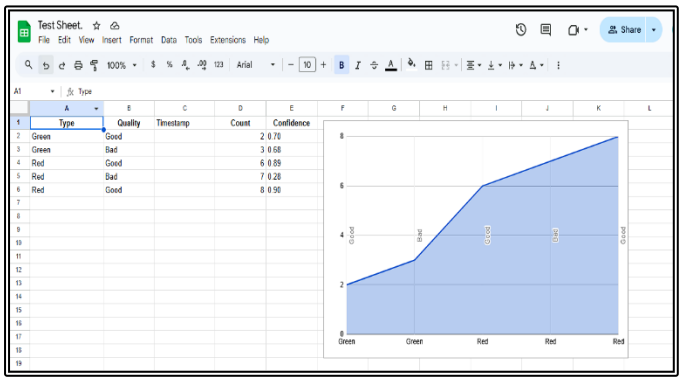


Fig.7. Google Sheets Logging Output

To facilitate precise apple detection, two E18-D80NK infrared proximity sensors are employed. The first sensor detects the presence of an apple entering the camera zone, triggering image capture. Once classified, the apple travels along the conveyor to a second IR sensor near the sorting zone. The Raspberry Pi then activates the appropriate MG996R servo motor: one servo redirects good red apples, another handles good green apples, while bad apples proceed to the end of the belt without actuation.

Weight monitoring is performed using a YZC-133 load cell paired with an HX711 amplifier and connected to an Arduino Nano. The Arduino reads and displays the cumulative weight of each batch on a secondary LCD. After every four apples are collected in a bin, the Arduino displays the final weight, which is also logged manually or via serial if needed.

The implemented apple sorting system demonstrated effective real-time classification and sorting performance as anticipated in the design phase. The accuracy graph, generated from extensive test runs, clearly illustrates the robustness of the YOLOv11 model integrated with the NCNN inference engine. Throughout the trials, the model consistently maintained an average classification accuracy above 92%, with slight variations observed depending on lighting conditions and apple orientation on the conveyor. This high accuracy level confirms the system's capability to distinguish between the four defined categories Good Red, Good Green, Bad Red, and Bad Green with minimal misclassifications, thereby validating the model training process and dataset quality.

Each class—Red, Red-Bad, Green, and Green-Bad—was tested with a minimum of 50 real-world samples. The classification accuracy ranged from 91% to 96%,

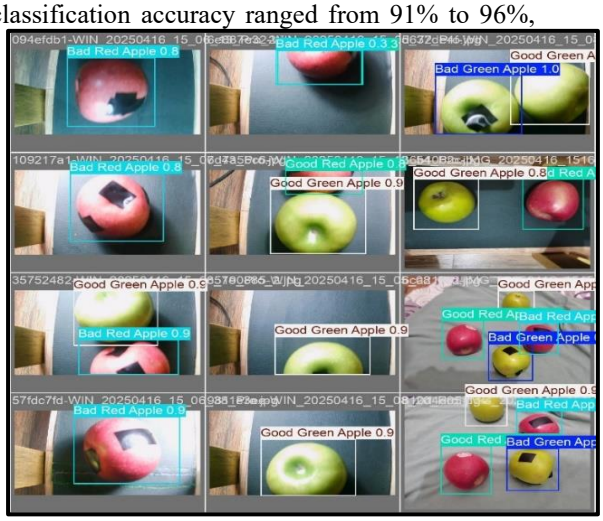


Fig.9.Result image

with confidence scores generally above 85%. The system exhibited a classification latency of ~0.09 seconds per apple, and servo response time was measured at under 0.1 seconds, enabling efficient real-time sorting. Weight readings from the load cell maintained a margin of error within ±5 grams, confirming the Arduino subsystem's measurement reliability.

IV. CONCLUSIONS

In general, the prototype model of the smart apple sorting system based on the YOLOv11 image recognition algorithm has been successfully developed and demonstrated to work effectively in practice. The system is composed of mechanical components, sensors, and control units that are
seamlessly integrated to achieve automated apple detection and sorting. Utilizing a sufficiently large and well-annotated dataset, the YOLOv11 model enables accurate classification of apples based on quality and type. Infrared proximity sensors assist in object detection and counting, while sorting decisions are executed using servo motors controlled via an Arduino. Real-time results are displayed on an 16x2 LCD. The successful implementation of the detection and sorting algorithms in this study lays the groundwork for deploying the system at a practical, industrial scale.

REFERENCES

[7] J. Smith and K. Johnson, "Warehouse automation: Trends and developments," Int. J. Logist. Res., vol. 14, no. 2, pp. 123–135, 2020.

[8] R. Brown, "Manual vs. automated sorting in modern warehouses," Ind. Eng. Rev., vol. 8, no. 3, pp. 45–52, 2018.

[9] M. Lee and P. Garcia, "Barcode-based automation: A case study," Autom. Today, vol. 9, no. 1, pp. 67–74, 2017.

[10] S. Patel, A. Roy, and N. Verma, "Sorting fruits and vegetables: The role of machine vision," Food Eng. J., vol. 11, no. 4, pp. 210–218, 2019.

[11] C. Wang and D. Zhao, "Machine vision systems in agricultural sorting: A review," Comput. Electron. Agric., vol. 150, pp. 302–310, 2018.

[12] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," arXiv preprint arXiv:1804.02767, 2018.

[13] H. Liu et al., "Real-time fruit sorting using deep learning," Sensors, vol. 20, no. 7, pp. 1987–1998, 2020.

Design & Implementation of Smart Waste Collection System with Optimized Route Planning

A L Josheph, N M Rajaguru
Department of Electronic and Electrical Engineering
SLIIT City Uni, Colombo

N M Rajaguru
Department of Electronic and Electrical Engineering
SLIIT City Uni, Colombo

***Abstract*— With the growth of the global population, the challenge of waste management has become more complex. Effective waste management is no longer just a necessity but a critical concern for public health, urban sanitation, and ecological sustainability. Identifying the filled public waste bins in cities and collecting them effectively is one of the essential initial tasks of waste management, to address it more effectively and cost-efficiently the "Smart waste collection system with optimized route planning" is designed. The prototype system utilizes microcontrollers, a Photo-Voltaic power supply unit, ultrasonic sensors, and a wireless module to detect and communicate bin status in a multi-hop wireless network cluster. Each bin is equipped with an Ultrasonic sensor to measure the garbage level and the data is transmitted through neighboring bins to a central control unit using the ISM-based wireless communication multi-hop network, enabling remote monitoring. When bins are detected as filled, their locations are processed to determine the shortest and most efficient route for the collection of garbage using the Google Maps Directions API. The system was tested successfully using a prototype with three waste bins equipped with sensing units, and setting the location on the maps through a cloud-based web visualization built using basic HTML, Python, and Javascript, the system performed accurate fill-level detection and real-time route planning for garbage collection. This solution offers a cost-effective and scalable approach to smart urban waste management by reducing manual monitoring and optimizing collection routes with traffic awareness.**

Keywords – Waste Collection, Optimized Route Planning, ISM, Multi-hop communication, Arduino Uno

I. INTRODUCTION

In recent years, the rapid growth of urban populations has created significant challenges in waste management, particularly in effectively collecting garbage from public bins. Conventional waste collection systems typically operate on pre-established routes and timetables, causing trash trucks to make unnecessary trips to empty public bins and delaying the collection of full bins due to inefficient collection routes. Also, this method of waste collection fails to escape from the city traffic resulting in unwanted delays. This lack of efficiency, not only increases operational expense but also has environmental and sanitary risks. Considering the complex and larger road traffic conditions identifying the best waste collection route is extremely useful, as stepwise identification of filled bin locations needs to be addressed and an optimal real-time path through the busy road traffic must be identified with the shortest route, that can remove the unwanted time delays of waste collection for the trash trucks.

Through the review of recent work, it was noted that several smart solutions for waste collection were addressed as follows, each with varying effectiveness and cost. S. Rathi et al. developed a system using stereoscopic cameras and Raspberry Pi with OpenCV for estimating bin levels, which uses image processing for garbage identification [1]. Similarly, G. Prajakta et al. combined camera and load cell data for monitoring fill levels and sending alerts via email [2]. Yan Li et al. created a sensor-heavy real-time monitoring system that relies on GSM and RF modules, which require a network service provider for communication [3]. A similar system to Yan Li 's, developed by S. S. Chaudhari et al., introduced a cloud-based IoT system for real-time monitoring [4]. Jose M. Gutierrez et al. also used IoT, GIS, and graph theory to dynamically optimize trash collection [5], while H. N. Saha et al. proposed a compact smart bin for household purposes with cloud connectivity [6]. K. Nirdel et al. used sonar and weight sensors for status updates via LCD and web, with a wireless module [7]. M. S. Chaudhari et al. developed a smart bin system designed for smart cities using cloud-based IoT sensors and actuators. Collectively, these systems reveal significant use of smart technology to resolve certain parts of waste collection. Most systems use costly parts that come with maintenance requirements, both for the system itself and for the communication network and are unable to address efficient route planning for waste collection.

To handle these waste collection problems a better communication network should be established, without any unwanted costs involved with GSM or for internet payments. One of the simplest and most cost-effective ways of communication is using radio frequency bands. For the purpose use of specific frequency ranges that are reserved for Industrial, Scientific, and Medical (ISM) applications channels can be utilized for cost-free communication without restrictions. For the bin level detection, a basic

Microcontroller unit (MCU) unit mounted with an ultrasonic sensor can be used.

II. Design Methodology

The proposed system is designed in phases. Each phase is divided in such a way that the development process achieves optimal time usage during the development of the system. No iterative approach was involved, so it removes the excessive time requirement and helps speed up the process. The initial phase started with problem identification, where the initial problem was identified and its impact along with the benefits of an ideal solution were analyzed to tune the throughput of the system to make it an ideal solution.

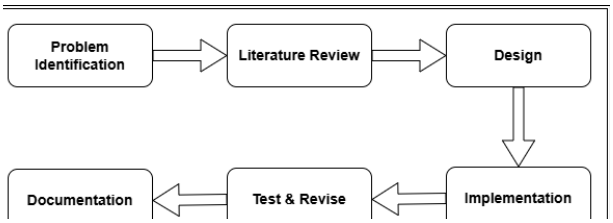


Figure 30: Design Methodology

Design methodology in Figure 1 shows the development stages of this product, initiating with the problem identification phase during which both the scope of the challenge and its real-world practicality were evaluated. During the second phase, literature solutions were analyzed to identify current flaws in them and develop an optimal solution. The designing phase begins with the primary high-level architecture of the system, which is extended into the circuit, network, software, and physical design. Then, the novel concept was developed during the implementation phase based on the outlay of each sketched design. Once the design was complete, the prototype was tested in units and as an integrated whole system, with revisions applied to the design for its betterment. Finally, the research project was concluded with the final documentation of all identified results, performance, and rate of success against the preset evaluation criteria set during the problem identification phase.

III. Design

The design of the system consists of a base station and the bin nodes which are interconnected through an RF network cluster. Each bin node is fixed to a public garbage bin which measures the bin fill level using an HC-SR04 Ultrasonic sensor unit connected to an Arduino Uno microcontroller board with ATmega328p main microcontroller unit. For the RF network, NRF24L01+PA+LNA transceiver with Rubber Ducky antenna module is used to utilize the 2.4GHz ISM band to send and receive data. The setup enables bins to talk through one another in a multi-hop network, where each bin can transmit its fill level to the central control unit through the network cluster. An optimized waste collection route is identified based on the filled bins, eliminating the need for wasteful inefficient routes, reducing time, and increasing efficiency using the Google Maps Directions API.

The individual units are powered by a simple Photo-Voltaic panel-based rechargeable battery setup which was designed separately. The use of simple and practically usable components focuses on a cost-effective, scalable, and energy-efficiency solution. With this system waste collection in a very densely populated area like a city with many public waste bins and even rural areas can be made more efficient.

A. Base Station Design

The base station is the primary computing unit which displays the bin fill status of each bin and the optimal path to collect the garbage from the filled bins. During the prototype design, a Windows-based PC is used as the base node which connects to the RF module via an Arduino Uno. Using a Python program, the serial port data is captured and sent to the web application to update the bin data logs into a Google sheet. The application integrated the Google Maps Directions API, which provided real-time and accurate routing based Dijkstra algorithm, A* algorithm, Contraction Hierarchies, Real-time traffic, Machine Learning, and User feedback-based efficient path routing on actual road conditions, including one-way streets, traffic congestion, and road closure. Finally, the optimum route is displayed web-based application for the filled bins's waste collection for the trash trucks using Google Maps.

B. Bin Node Design

Each smart bin is equipped with a main MCU unit, ultrasonic sensing unit, RF communication module, and a PhotoVoltaic power supply unit.

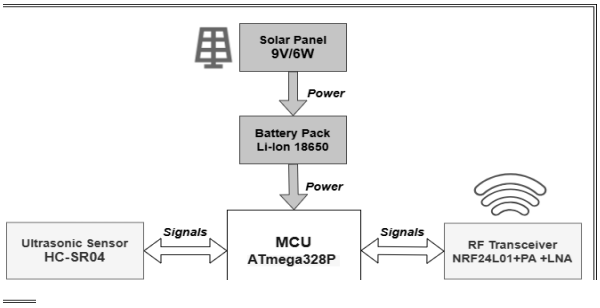


Figure 2: High-level Architecture – Bin Node

The high-level architecture of the bin nodes in Figure 2 represents how each module is interconnected in a bin node. The primary system is powered with a battery pack automatically charged by solar energy. The MCU receives the bin level from the ultrasonic sensor and initiates the transmission of data through the SPI interface in between the RF module to send the data to the base station through the intermediate bin nodes. Then the data is transmitted via serial communication to a base station.

C. *Network Design*

The Multi-hop network cluster is designed using the NRF24L01+PA+LNA, RF modules. A single module can communicate up to 1100 m, since it uses a cluster to communicate the maximum distance between bin nodes must be less than 1100m, but the system can transmit data through other nodes which increases the data transmission distance of the overall system.

As shown in Figure 3, the bin nodes are connected to a maximum of 5 other bin nodes in the network, up to 5 levels. Theoretically, it can accommodate up to 3125 bin nodes (5^5). Since the transmission of data is done through the 2.4GHz ISM band there is no legal restriction for the system for transmission of data through the RF network.

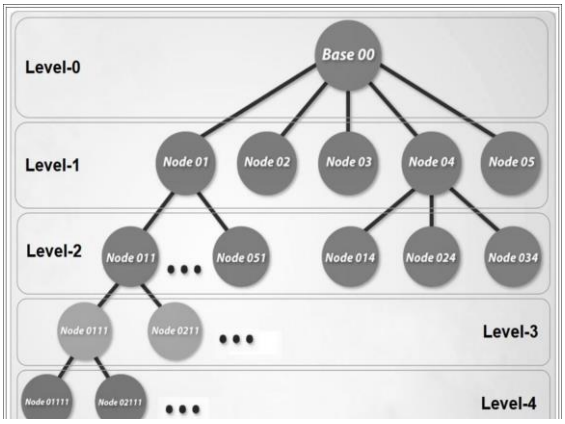


Figure 3: Network Design – NRF24L01+PA+LNA Modules

IV. RESULTS

The system is tested with three basic physical bin nodes and six predefined bin nodes in the web application. When the bin levels are sent through the RF network it logs the date to the Google sheet.

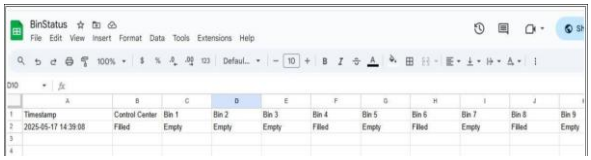


Figure 4: Bin Status - Google Sheet

Bin node status in Figure 4 shows the data from the network updated into the Google sheet. Then the system can be prompt by using the web application to identify an optimal path for the waste collection for the trash trucks. The system considers only the filled bin locations and then the route will be generated using Google Maps Directions API, which uses sophisticated routing algorithms with real-time traffic awareness.

Figure 5: Final Route – Google Map

As shown in Figure 5 the test result gets generated from the system for the waste collection routes considering the real-time data. During the test, it was identified that communication through the RF network nodes is ideal to place in a line of site to each neighboring bin node for better communication. Also, during the startup of the nodes, it is better to initiate the communication of the nodes, in an environment free of 2.4 GHz interferences like Wi-Fi networks for faster connectivity establishment.

V. CONCLUSION

This work presents a prototype smart waste management system capable of monitoring bin fill levels, transmitting data via RF modules, and generating optimized routes mapping Ultrasonic sensors ensured fill level detection of the bins, while the web interface provided clear visualization of bin status and paths using the Google Maps. The main functions of the system were validated in a controlled environment.

To upgrade the system into a more robust solution the MCU unit can be upgraded into an industrial-grade counterpart. Also, bin node-to-node distance can be increased by upgrading it to a long-range RF transceiver in the ISM band (ex: LoRa-WAN network).

REFERENCES

[1] S. Rathi, S. Pande and H. Lokhande, "Smart Garbage Collection System," International Journal For Research, vol. 5, no. IV, p. 7, 2017.

[2] G. Prajakta, J. Kalyani and M. Snehal, "SMART GARBAGE COLLECTION SYSTEM IN RESIDENTIAL AREA," International Journal of Research in Engineering and Technology, vol. 04, no. 03, p. 3, 2015 March

[3] Y. Li and Y. Huang, "Research on the Urban Garbage Collection System Based on Multi-Sensors," *ATLANTIS PRESS,* no. ISBN 2352-5401, p. 6, 2016 October.

[4] S. S. Chaudhari and V. Y. Bhole, "Solid Waste Collection as a service using IoT-Solution for Smart Cities," Int. Conf. Smart City Emerg. Technol. ICSCET 2018, p. 5, 2018.

[5] J. M. Gutierrez, M. Jensen, M. Henius and T. Riaz, "Smart Waste Collection system based on location intelligence," Procedia Comput. Sci, vol. 61, p. 7, 2015.

[6] H. N. Saha and S. Auddy, "Waste management using internet of Things (IoT)," Conf. Industrial Automation and Electromechanical Engineering Conference (IEMECON), p. 5, 2017.

[7] K. Nirde, P. Mulay and U. M. Chaskar, "IoT based solid waste management system for smart city," Conf. International Conference on Intelligent Computer and Control System (ICICCS), p. 4, 2017 June.

[8] M. S. Chaudhari, B. Patil and V. Raut, "IoT based Waste Collection Management System for Smart Cities: An Overview," in International Conference on Computing Methodologies and Communication, CFP19K25-ART,

Track 4: Humanities and Social Sciences

Perceptions of Third-Year TESL Undergraduates on Kahoot as a Gamified Learning Tool in Sri Lanka

Tharinduni Wijayarathna
Department of English language Teaching
University of Kelaniya, Kelaniya, Sri Lanka
tharindunithamodya@gmail.com

***Abstract*— The rapid and increased technology integration becomes an important position, especially in the language learning context. Kahoot, as a gamification tool, is implemented in English Language learning classrooms. The application of this gamified element is examined in various educational contexts and becomes widespread in higher education due to its significance. Nevertheless, only few studies regarding students' perceptions towards the use of Kahoot among teaching English as a Second Language undergraduates were reported. Therefore, this present study investigates third-year TESL undergraduates' perceptions towards utilising Kahoot in enhancing their English Language skills within their semesters. This research project examines how a focus group of five third-year undergraduates, currently studying at the department of English Language Teaching, University of Kelaniya, has perceived Kahoot as a gamified tool to enhance their learning. Semi-structured interviewing was used as the data collection instrument, where five interviews were conducted in total, with each interview lasting approximately thirty minutes. The collected data were analyzed using the thematic analysis method. The findings of the study provide evidence for the presence of a number of unique perceptions among the interviewed participants. They emphasized that the implementation of Kahoot becomes more effective for their learning, while they overcome a few challenges with the use of Kahoot as a gamified tool. Furthermore, suggestions are discussed for the effective implementation of Kahoot. This study has already reiterated the need to explore this research area further because the studies are very limited. There should be further studies for ESL (English as a Second Language) development. In conclusion, the research has the potential to promote social change by emphasizing the importance of using Kahoot as a gamified element in pedagogy in Sri Lankan universities.**

Keywords*— *English Language learning, Effectiveness, Gamified element, Kahoot, Student perceptions

I INTRODUCTION

Education is viewed as a central element in preparing the knowledge, skills, and attitudes essential for the personal and professional development of individuals and societies. "Knowledge is becoming the core of country's competitive advantage, and the production and the dissemination of knowledge are parts of the process of development based on an effective Higher Education Policy" [4]. In the higher education context, universities determine the advanced learning backgrounds necessary for cultivating critical thinking, innovative skills, and specialized abilities for learners. Within this framework, assessments play a pivotal role because they foster to measure and build up student learning outcomes with proper guidance of both effective pedagogy and academic progress. With the rapid increase in technology integration, university lecturers tend to use gamified elements to make the lesson more interesting and even effective in understanding the content properly. Among these gamified elements, Kahoot becomes very significant as many lecturers use it to enhance students' language learning skills.

II RESEARCH QUESTIONS

- How do third-year TESL undergraduates perceive the effectiveness of Kahoot to enhance their language learning?
- What challenges do they experience in the use of Kahoot as a gamified element?

III LITERATURE REVIEW

Many studies emphasize that integrating technology into the pedagogy becomes effective for a successful teaching and learning process. Charbonneau depicts that integration of technology into lectures could enhance their engagement and decrease off-task behavior [2]. Language learning can be enhanced through technology-assisted language learning. Toth defines that lecturers use Kahoot as a gamified tool to capture the learners' attention to their lessons [5]. Wang indicate that Kahoot as a gamified tool because it has been reported to be effective for enhancing language learning [8]. Kahoot is recognized as the latest digital game-based e-learning and supports teachers to construct assessments [7]. According to Cardenas-Moncada, this gamified tool provides not only informative

content but also fun and interactive activities to learners [1]. Ching depicts that Kahoot provides a way to review and make vocabulary more memorable [3]. Plump and LaRosa expose that most of the participants in his study agreed to have used Kahoot as a gamified tool to understand the concepts appropriately [6]. Although many studies were conducted to investigate the application of Kahoot in the international context, its effectiveness and the challenges were not examined properly in the Sri Lankan context, especially in the tertiary education.

IV METHODOLOGY

This study was conducted using final-year Teaching English as a Second Language students in the Department of English Language Teaching of the University of Kelaniya. The sample comprised five undergraduates representing the target group, with 4 female undergraduates and one male undergraduate. In concurrence with ethical practices required of academic research, this study also scrutinized very strict adherence to recommended ethical considerations in conducting research. Accordingly, informed consent forms and participant information sheets were created to receive the genuine consent of the participants for this study. Semi-structured interviews were employed with the interview protocol comprising six open-ended questions to address the research questions. Interview data were analyzed using thematic analysis for the qualitative approach in this study, after coding the raw data.

V RESULTS AND EVALUATION

This research was conducted in synchrony with a specific timeframe, spanning over three months to collect and analyze data. The following table also demonstrates the themes and sub-themes. The analysis exposed the themes, namely, enhanced engagement, language skill development, challenges and limitations, and suggestions for the implementation of Kahoot. The themes reflected their positive perceptions, challenges and even the suggestions to implement effectively by addressing the research questions of the study. The learners emphasized that the application of Kahoot makes learning interactive as well as enjoyable. It suggested their motivation and active participation in the lesson in contrast to the learning which happens in a traditional classroom setting. Furthermore, students reported that with the use of Kahoot games, they can have an increased acquisition of vocabulary and grammar. Kahoots use was instrumental in motivating them towards language acquisition due to the nature of immediate feedback received on their progress. Even though they have positive perceptions towards the use of Kahoot as a gamified tool, they have to overcome challenges because of factors. Due to the time constraints and technical issues, they must face some difficulties. To address these challenges, the participants provided implementations to make it more effective. It is significant to integrate Kahoot games into lessons because it is recognized as a productive way to capture students' attention. It is very supportive to provide adequate time to engage in the activities. They highlighted that accessibility should be ensured. These implementations provided a proper background to make the learning and teaching context successful.

TABLE 1. Themes and Sub-themes resulting from the Thematic Analysis

Main Themes	Sub Themes
1. Enhanced engagement	Motivation Active participation
2.. Language Skill Development	Vocabulary and grammar development Immediate feedback
3. Challenges and limitations	Time pressure Technical issues
4. Suggestions for the implementation of Kahoot	Integrating Kahoot into lesson plans Providing adequate time Ensuring accessibility

VI CONCLUSION

This study examines the perceptions from the mindset of student-teachers as the teaching English as a second language undergraduates. They determine the effectiveness of the implementation of Kahoot games being a significant position in their learning process. Therefore, the data analysis of this study indicated students' positive perceptions towards formative assessments by addressing the first research question, while the second research question of this study explored the limitations that they have experienced in the use of Kahoot during their academics

REFERENCES

[1] C.Cardenas-Moncada, M. Veliz-Campos, and L.Veli, "Game-based student response systems: The impact of Kahoot in a Chilean vocational higher education EFL classroom, "Computer-Assisted Language Learning Electroic Journal, Vol.21, no.1, pp.64-78,2020.

[2] A.K.Charbonneau, The effects of using Kahoot as a formative assessment in the middle school science classroom,M.S. thesis, Montana State University, Montana,2018.[Online].Available: https://scholarworks.montana.edu/xmlui/handle/1/14719

[3] H.-H. Chiang, "Kahoot in an EFL reading class," Journal of Language Teaching and Research, vol. 11, no. 1, pp. 33–44,2020. [Online]. Available: http://dx.doi.org/10.17507/jltr.1101.05

[4] N. S. Jayasundara, "Higher education policy in Sri Lanka: Implementation in state universities," Scientific Research Journal, Vol.2, no.2, pp. 41-44,2014

[5] S. A. Licorish, H. E. Owen, B. Daniel, et al., "Students' perception of Kahoot!'s influence on teaching and learning," Research and Practice in Technology Enhanced Learning (RPTEL), vol. 13, p. 9, 2018.

[6] C. M. Plump and J. LaRosa, "Using Kahoot! in the classroom to create engagement and active learning: a game-based technology solution for e-learning novices," Management Teaching Review, vol. 2, no. 2, pp. 151–158, 2017.

[7] Á. Tóth, P. Lógó, and E. Lógó, "The effect of the Kahoot quiz on the student's results in the exam," Periodica Polytechnica Social and Management Sciences, vol. 27, no. 2, pp. 173–179, 2019. [Online]. Available: https://doi.org/10.3311/PPso.12464

[8] A. Wang, "The wear out effect of a game-based student response system," Computers & Education, vol. 82, pp. 217-227,2015

An Investigation of Morphological Processes in Sri Lankan English; A Study Based on Selected YouTube Comment Sections

Shanika Ranthunga
Department of English Language Teaching
University of Kelaniya
Kelaniya, Sri Lanka
r.shanika2002@gmail.com

Sachini Godage
Department of English Language Teaching
University of Kelaniya
Kelaniya, Sri Lanka
sachinigodage2002@gmail.com

***Abstract*— British rule left a lasting impact on Sri Lanka with the introduction of English language. In addition, English was considered as the official language of the country under the British government. After gaining independence in 1948, Sri Lanka evolved into a multi-ethnic country which subsequently affected its vernacular languages, which are Sinhala and Tamil. English language in Sri Lankan context emerged as Sri Lankan English (SLE), a new variety representing diverse cultural and social factors of the country. SLE has undergone numerous word formation processes and developed into one of the most major varieties of the English language. In the present day, the frequent use of social media platforms like YouTube comment section has become a dynamic site for the morphological processes in SLE. The content creators utilize new SLE words to facilitate audience engagement. This research aims to investigate the productive and frequently used morphological processes of SLE in selected YouTube comment sections. The study was designed as a qualitative content analysis. A sample of thirty words was selected employing judgmental approach where researcher's knowledge on SLE morphology was utilized to select YouTube channels. According to the findings, compounding accommodates the largest number of entries in SLE vocabulary found in the You Tube comment section of the selected channels. Borrowing stands as the second-largest entry. Clippings and affixation respectively accommodate fewer entries. Thus, compounding and borrowing can be identified as the most productive and frequently used morphological processes in You Tube comment section in Sri Lanka. Moreover, most of the compounding and borrowings are influenced by the Sinhala language, revealing its significant influence on morphological processes in Sri Lankan English. In conclusion, it is evident that further research on the SLE morphological processes is essential to codify the modern trends in SLE vocabulary due to the influence of digital platforms like YouTube.**

Keywords—Morphological processes, Sri Lankan English, YouTube

I INTRODUCTION

Sri Lankan English is a unique variety of English that consists local languages, cultures and social factors. According to [1], "the English language was introduced to Sri Lanka in 1796 as a 'gift' from colonists as an official language for administrative work and trade". Later, this new variety of English was labelled as Sri Lankan English [2] and this variety is heavily influenced by the vernacular languages in Sri Lanka (Sinhala and Tamil).

Sri Lankan English has undergone multiple word formation processes such as compounding, borrowing, clipping, and affixation. Due to the rising social media impact in the contemporary world, these morphological processes have reached frequent usability. Even though there is a considerable amount of research studies in the field of morphology, a scarcity can be seen in the studies of Sri Lankan English morphological processes, significantly in the digital context like YouTube. The aim of this research is to systematically identify, classify and analyze the most productive word formation processes (compounding, borrowing, clipping, and affixation) of SLE used in YouTube comments sections, while providing the contextualized linguistic examples. The research was conducted through two research questions to examine how YouTube content creators utilize SLE word formation processes to engage with audience, and which are the most productive morphological processes in the context.

II LITERATURE REVIEW

This study focuses on the productive SLE morphological processes utilized in the You Tube comment section by the audience. This research identifies and categorizes the word formation processes applied in Sri Lankan YouTube comment section. In addition, four morphological processes compounding, borrowing, clipping, and affixation are selected based on the results of the research and previous studies on morphological processes in social media platforms.

As in [3], "compounds are formed by combining two bases, which may be words in their own right, to form a new lexical item". Moreover, compounds can be divided into types based on their meaning and behaviour of the head noun. Furthermore, [4] has observed a list of compounds in his A Dictionary of Sri Lankan English which are common to colloquial SLE. "Jump- seat" and "Agency- office" are examples from the dictionary, which are defined as endocentric and highly compositional. [4] has documented another type of compounds that are directly translated into English from Sinhala language such as "small- village" and "Colombo- seven". According to [2], compounding is one of the most productive word formation processes in SLE and it may consist of combination of Sinhala, Tamil and English words.

As [3] defines borrowing as an impactful word formation process where instead of forming new words a language can expand the number of words in its lexicon by borrowing from the vernacular languages. However, borrowing stands as the most influential word formation process in SLE history. Reference [2] asserts that SLE is rich with borrowings from Sinhala, Tamil, Malay, and Arabic reflecting the cultural impact on the language.

Clipping is a systematic technique to create new words by cutting or shortening longer words by reducing one syllable or more, without changing the original lexical meaning of the word. Reference [3] has introduced the types of clipping as back clipping, middle clipping and front clipping. As [5] states that clipping is one of the most common word formation processes in social media platforms due to its convenient nature on informal conversations.

In morphology, when bound morphemes are connected to a root to form a meaningful word is known as the process of affixation. Furthermore, [3] has discovered two types of affixations, such as inflectional affixation (which mainly occurs to form grammatical words) and derivational affixation depending on the usage of the word. According to [2], affixation is one of the most prominent SLE word formation processes.

III Methodology

The study was designed as a content analysis of the comment sections from the YouTube videos featuring SLE, particularly focusing on spoken content from influencers. The sample was chosen through judgmental approach where the researcher's knowledge of SLE and morphology is used to select YouTube channels. The identified words were qualitatively analyzed to classify the morphological processes (compounding, borrowing, clipping, and affixation) and subsequently, quantitative analysis were applied to determine the most productive word formation processes through their frequency of usage. The chosen sample consists of thirty words from the comment sections of two YouTube channels (LET'EM COOK and ChrisnCrysBlissed), which were chosen for their relevance to the authentic portrayal of contemporary SLE and high audience engagement. The sample size was adequate and feasible to represent the morphological features within a variety of English. The existing classification of morphological processes by [5] provided a theoretical yardstick to analyze, categorize, and interpret the collected data. More particularly, this research presents compounding, borrowings, clippings and affixation [5] as four major morphological processes found in the SLE on YouTube comments. The content was carefully analyzed to identify the words that represent features of SLE and categorized according to the relevant morphological processes. The existing classification of morphological processes by [6] provided a theoretical yardstick to analyze, categorize, and interpret the collected data. More particularly, this research presents compounding, borrowings, clippings and affixation [6] as four major morphological processes found in the SLE on YouTube comments. This categorization allows for the examination of the dynamic nature of SLE used in this specific digital platform and its morphological impact on SLE within this context.

IV Results and Evaluation

When the research findings were analyzed, compounding can be identified as the most productive morphological processes, which involves the combination of two words, both English or one English and the other Sinhala or Tamil. Compounds (57% of the overall sample) have dominated the morphological processes on YouTube, being the most productive and frequently used morphological process as shown in Fig. 1. Most of them are endocentric in which the meaning is derived from the head, and highly compositional compounds in which the meaning is easily derived from its components. Some compounds are indicative of the socio-political and geographical context; 'Kandy situation', and 'Pageant fever' [7]. According to the evidence provided by research, hybrid compounds mostly consist of words from Sinhala and English; 'Genuine /jɑːluwə/', '/supiri/ solution', 'Mental /ˈpraʃnə/', 'Private /iskoːle/', and 'Government /iskoːle/' [8].

Borrowings (23% of the overall sample) are the second most productive morphological processes that integrate new words into SLE from local languages, reflecting our national and in-group identity as shown in Fig. 1. Almost all the borrowings found in the particular digital platform are directly borrowed from Sinhala /budu radʒːnan vahaŋse/, /dʒeːsus vahaŋse/, /sædæhæwatun/, /ʃramadaːna/ [9].

Clippings involve the formation of a new word form with the same meaning as the original lexical item by lopping off a portion of a word [3] Clipping represents 13% of the overall sample indicating its moderate role in morphological processes in Sri Lankan YouTube comments as shown in Fig.1. It involves the formation of a new word form with the same meaning as the original

lexical item by lopping off a portion of a word. In the digital context, clippings are used to ease and speed up communication; 'subs' for 'subliminals', 'insta' for 'instagram', 'Undergrad' for 'Undergraduate', and 'Epi' for 'Episode' [10].

Forming unusual and unique lexical items affixation attaches English affixes to Sinhala or Tamil root words [11]. Affixation represents 7% of the overall sample indicating its minor role in morphological processes in Sri Lankan YouTube comments as shown in Fig. 1. Under this morphological process, both inflectional affixations; /lokkaːs/ (possessive 's') [12], and derivational affixations; /vɪsɑːkiən/ (suffix 'ian') [8] could be uncovered on YouTube comments.

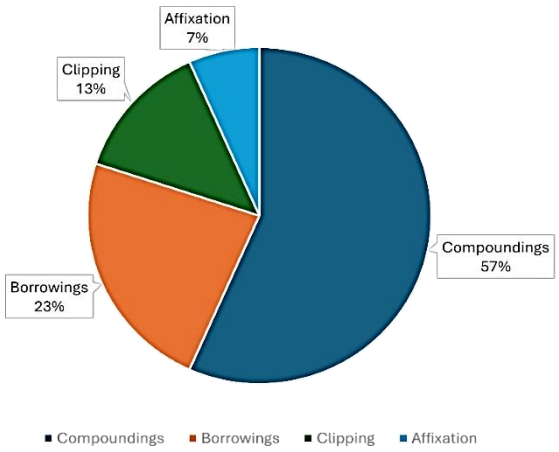


Fig. 1. Analysis of word formation processes in Sri Lankan YouTube comments

Compounding accommodates the largest number of entries in SLE while borrowing stands as the second-largest entry. Clippings and affixation respectively accommodate fewer entries in SLE. Thus, the results of the research denote that compounding and borrowing are predominantly used in Sri Lankan YouTube comments, while clippings and affixations are used infrequently.

V CONCLUSION

In conclusion, this study exemplifies some of the major trends found in the SLE morphological processes on YouTube comments which are representative of the Sri Lankan online community. One key trend is the dominance of compounding and borrowings in comparison to other morphological processes such as clipping and affixation. Another trend is the significant influence of the Sinhala language to coin unique compoundings and borrowings SLE. Thus, it is evident that Sinhala as the superstrate of Sri Lanka has exerted a significant influence over these morphological processes which are recognizable on YouTube comments.

Due to time constraints, this study has focused only on two Sri Lankan YouTube channels resulting in an insufficient sample size for statistical measurements to fully represent the variety of morphological processes as they appear on Sri Lankan YouTube comments. Thus, future studies can expand the sample into a broader range of YouTube channels to improve generalizability and to codify the SLE vocabulary which is rapidly growing with the influence of digital platforms like YouTube.

REFERENCES

[1] M. Heart, "BE variation in Sri Lankan English," Language Variation and Change, vol. 17, no. 2, pp. 101–202, 2005.

[2] M. Gunasekera, The Postcolonial Identity of Sri Lankan English. Colombo, Sri Lanka: Katha Publication, 2005.

[3] F. Katamba, English Words. London, U.K.: Routledge, 1994.

[4] M. Mayler, A Dictionary of Sri Lankan English. Colombo, Sri Lanka: Perera Hussein Publishing House, 2007.

[5] V. Harankaha, "A morphological analysis of English language neologisms used on Facebook, Viber and WhatsApp," in Proc. Int. Open Univ. Res. Sess., 2022, pp. 1–6.

[6] L. Bauer, English word-formation. Cambridge University Press, 1983.

[7] ChrisnCrysBlissed, "Sri Lanka we can do better," YouTube, 2025. [Online]. Available: https://www.youtube.com/@ChrisnCrysBlissed

[8] LET'EM COOK, "Why do Sri Lankans judge people by schools?" YouTube, 2024. [Online]. Available: https://www.youtube.com/@letemcookshow/videos

[9] ChrisnCrysBlissed, "Things happening in Sri Lanka," YouTube, 2025. [Online]. Available: https://www.youtube.com/@ChrisnCrysBlissed

[10] LET'EM COOK, "Lessons from working without a degree," YouTube, 2025. [Online]. Available: https://www.youtube.com/@letemcookshow/videos

[11] S. Fernando, "Sri Lankan English (SLE) Vocabulary. A new vocabulary in a new variety of English," J. Humanit. Soc. Sci., vol. 7, no. 8, 2012.

[12] ChrisnCrysBlissed, "Stop judging people," YouTube, 2025. [Online]. Available: https://www.youtube.com/@ChrisnCrysBlissed

Investigating the Role of Self-Talk Practice in Developing Spontaneous Speaking Confidence Among Foundation Students at a Private University in Sri Lanka

D.A.Wahalawatte
School of IT
SLIIT CITY UNI
Sri Lanka
daniruidk@gmail.com

P.V.Y.Tharuka
School of IT
SLIIT CITY UNI
Sri Lanka
tharukayasith8@gmail.com

M.Y.Pannilavithana
School of IT
SLIIT CITY UNI
Sri Lanka
Pannilavithanam@gmail.com

A. A Baghavan
School of IT
SLIIT CITY UNI
Sri Lanka
Ashenandrew325@gmail.com

K. M. H. Bhagya
School of IT
SLIIT CITY UNI
Sri Lanka
hiruniminupama@gmail.com

W.R.I.U. Nilaweera
School of IT
SLIIT CITY UNI
Sri Lanka
ishadi.n@sliit.lk

***Abstract*—Successful spontaneous conversation is an essential skill for an ESL learner. Spontaneous conversations become difficult as they involve instant thinking and processing. This study investigates whether self-talk would be an efficient strategy to improve the confidence level in spontaneous conversations in ESL learners. The study was done among a group of students from SLIIT City Uni, semester 1 February intake batch. A Mixed method was used in a pre-test/post-test design for data collection, namely pre-test and post-test. A self-talk intervention was implemented over a period of seven days between these interview sessions. The findings indicated a significant improvement in ESL Learners' confidence levels during spontaneous conversations following the implementation of the self-talk strategy. In this study, the qualitative analysis complemented the quantitative findings, and several important recommendations were proposed based on the integrated results.**

Keywords— Confidence in speaking, Rubric-based assessment, Self-talk strategy, Spontaneous conversation.

I. INTRODUCTION

According to [1], speaking is the art of using words aloud to connect thoughts through live conversations. Speaking is the ability to convey feelings and thoughts to another person, and to speak, an individual has to create sounds using various parts of their body [2]. Over the years, numerous languages have been developed around the world. While other languages are important in this world, English has become one of the most essential languages to communicate with each other due to being spoken by millions of people around the world [3]. This makes it difficult for people from non-native English-speaking countries to express their thoughts. A significant number of students need to study English as their second language just because it is in the curriculum [4]. When it comes to expressing thoughts, it has been revealed that the ability to speak spontaneously without prior preparation can be essential for daily human interactions [5]. However, spontaneous conversations will open a set of challenges that demand quick thinking and adaptability [5]. This makes it more challenging for ESL learners as they must express themselves orally in a foreign language spontaneously [6].

The two main cognitive factors that tend to increase communication anxiety and speaking anxiety are the negative evaluation of the audience and expectations of poor performance [7]. This acts as a psycholinguistic barrier for ESL learners, which affects their oral proficiency [7]. As a solution, many researchers have turned their attention towards various methods to improve English speaking, one of them is the self-talk strategy.

Self-talk is a strategy in which individuals express their ideas in their own way; it is their inner voice [8]. Self-talk is not just limited to speaking in the first person with "I"; it can also be directed at oneself using the second-person pronoun "You" [9]. Cognitive theorists emphasize that there is a strong link between what people say to themselves and what they do [7]. Self-talk enables learners to verbalize their inner thoughts freely, without external constraints, thereby creating a psychologically safe environment. [10]. This would eventually increase their

confidence in speaking as well. While prior research has examined the role of self-talk strategy in speaking ability [11] and reducing speaking anxiety [7], there is a research gap for using self-talk strategy to improve confidence in spontaneous conversations.

While previous studies were focussed on public speaking and general English fluency, this study specifically investigate whether self-talk strategy would boost ESL Learner's confidence in spontaneous conversations. This area remains underexplored, which creates a clear research gap that this study aims to fill. This gap is significant, as spontaneous speaking is a key skill in real-world communication but often overlooked in ESL training. This distiction is crucial as spontaneous conversations triggers anxiety in most ESL learners. Addressing this area not only fills the research gap but also offers a practical strategy to improve confidence in these situations. If this strategy is proven to be successful, it would be a unique approach for a problem faced by many ESL learners.

Aims of the Study:

- To examine whether self-talk can improve ESL learners' confidence in spontaneous conversations.
- To observe spontaneous speaking confidence as a dependent outcome influenced by observable performance factors such as pronunciation, vocabulary use, fluency, task completion, and expressive delivery.
- To analyze the effectiveness of a short-term self-talk training program.

II. Literature Review

Self-talk (ST) has been used in various studies to discover whether it would be an effective strategy to improve speaking skills.

A. Reference [7] focused on Saudi EFL university students and the results of using self-talk technique to reduce public speaking anxiety. It has been found that there is a significant negative correlation between self-reinforcement (a type of ST) and anxiety levels.

B. Reference [11] conducted a study on Indonesian high-school students to discover whether the self-talk technique is useful in improving students' speaking skills. Their experiment showed that learners who practiced positive self-talk displayed increased fluency and reduced nervousness during speaking tasks.

While previous studies were focused on public speaking and general English fluency, little to no focus has been placed on its impact on speakers' confidence level in spontaneous conversations. This gap is significant since spontaneous speaking skills are crucial for real-world communication but often overlooked in ESL training. This study aims to address that gap.

III. Methodology

A. Data Collection and Research Procedure

This study was conducted in a group of 15 students in SLIIT City Uni February Intake, Foundation Semester 1 batch. This group was pre-tested and post-tested. As the name implies, these students' abilities were measured through an open-ended and semi-structured interview process, which allowed more flexibility in gathering information about their confidence level in speaking as well as their opinion about self-talk. The post-test was conducted before the 7-day story/scenario-based training period. After the training period, the post-test interview was conducted, and it was done to observe the change in participants' improvement in confidence level in spontaneous conversations.

B. Quantitative and Qualitative Data

This study followed a mixed approach of using both quantitative data and qualitative data. To gather quantitative data, the Rubric framework implemented in Ismaila's research [12] was used and was customized to measure hand gestures and facial expressions that align with finding the confidence level of the individual. These data were measured on various criteria such as pronunciation, vocabulary, grammar, fluency, and expressive delivery. The qualitative data was gathered through a semi-structured interview conducted to explore participants' opinions regarding their confidence level in spontaneous conversations.

C. Instrument (Rubric Method) (Ismailia, 2021)

The Rubric method used in Ismaila's research [12] was used to measure pronunciation, vocabulary, grammar, fluency, and task proficiency in the study. Each rubric criterion was scored on the scale of 50-100. This method was modified to gather information about the expressive delivery of the participants by adding a separate section to measure the expression level through eye contact, hand gestures, and body language. In this study six rubric criteria were used as independent variables to contribute to obtaining the confidence level which is treated as a dependent variable. This Rubric assessment is used to measure the overall self-perceived confidence level through all the criteria.

D. Tool (SPSS)

In this study, IBM SPSS statistical software was used to analyze the collected data [13]. Measures such as mean, media, and mode were used to analyse the changes that happened during the pre-test and the post-test. These measures were used across all the scales of the Rubric method, which are pronunciation, vocabulary, grammar & accuracy, fluency, task completion, and expressive delivery for a clear comparison before and after the training period.

E. Training Method

For the scenario/story-based training process,

participants were given a card that contained a simple story/scenario related to an imaginary storyline. They were asked to imagine the scenario, add some of their own story, and record an audio of themselves talking about that scenario/story for 5 minutes. These scenarios/stories were designed to encourage more imagination. This process was carried out for 7 days by changing stories/scenarios each day.

F. Participants

A total of 28 students were gathered from SLIIT City Uni February Intake, Foundation Semester 1 batch. For this study, Foundation students were selected because they are in an early stage of academic language development and frequently need to engage in activities that challenge their English-speaking skills, making them an ideal group to observe the immediate change. Beginner students frequently have speaking issues such as anxiousness and lack of self-confidence [14]. To maintain the relevancy of this study, only students who are learning English as their Second Language were included from the total of 28 students in the batch. Based on the exclusion criteria, the final group consisted of 15 students.

The interviews were done in a relaxed atmosphere where students could express their thoughts freely. Prior to participation, learners were informed about the study procedure, and their volunteer consent was taken, and they were told that the interview process would be recorded and quoted in data analysis.

IV. Results and Evaluation

A. Quantitative Data

The participants' pre-test and post-test interviews were video recorded and were analyzed with the help of the modified Rubric method of Ismaila's research [12]. This statistical analysis provides a numerical range between 50-100 based on their proficiency level in each of the sections. The mean, median, and mode of those data were calculated with the IBM SPSS statistical software.

Based on the results,

- Pronunciation: Significant increase from Pre-Test (72%) to Post-Test (83%), indicating a (+11%) increase.
- Vocabulary: Increase from pre-test (76.5%) to post-test (85.4%), showing an increase of (+8.9%).
- Grammar & Accuracy: Improvement from pre-test (70.4%) to post-test (72.0%), increase of (+1.6%)
- Task Completion: Increase of pre-test (68.7%) to post-test (80.9%), indicating (12.2%) increase.
- Expressive delivery: from pre-test (65.4%) to post-test (72.8%), showing a significant increase of (+7.4%)
- Fluency followed a different pattern (pre-test: 69.8%, post-test: 68.3%) making the total average difference of pre-test: 70.47% and post-test: 77.07% being +6.6%.

Spontaneous speaking confidence is treated as a dependent variable inferred from six performance criteria.

B. Qualitative Data

Data of the Semi-structured interview were gathered, transcribed, and analyzed carefully. Based on the questions and answers, several categories were established. A detailed qualitative analysis can be seen in the following:

I. Challenges faced in spontaneous conversations

When the participants were asked questions regarding the challenges they faced during spontaneous conversations, they had a variety of different answers.

- Difficulty in quick thinking and responses forming

  Out of them, one of the most common challenges was when speaking spontaneously, they find quick thinking and responding as a significant challenge. Participant No.3 responded that "I normally pause a lot, sometimes I don't know how to continue the sentence" and Participant No.11 responded that "When I started to speak my mind goes blank and it take some time ideas to come to mind".
- Challenges in maintaining grammatical accuracy while speaking.

  While others had a hard time stating something grammatically correctly. Participant No.14 said that "I have an idea in my mind, but when I try to convert that to words, most of the time it's grammatically incorrect" and No.13 stated that "when I speak up my grammar is not good".
- Struggles with vocabulary

  Another concern that the participants raised is regarding vocabulary. As No.2 said, "I know a decent amount of words in English, but when I try to speak quickly, I can't complete my sentence because I don't know that specific word to complete it" and No. 8 described that, "I can't fill in the sentences because sometimes I don't know the relevant suitable word".

II. Psychological and Social Barriers

- Nervousness due to fear of judgment.

  Many participants highlighted that when they encounter spontaneous conversations, they become anxious because of others' judgement. According to participant No.7, "Normally when I encounter some foreigner or someone who speaks in English, my heart rate increases in those instant conversations because what would they think of me if I mess up?". While No. 9 stated that "I sometimes start stuttering when I think about what others would think of me".

III. Will self-talk prepare you for spontaneous conversations?

- Helps in organizing thoughts and structuring

sentences.

Most of the participants reported that self-talk helped them to practice organizing their thoughts and boost their confidence. As participant No.15 said, “When I practice in my head, it just feels like I’m preparing myself for a real conversation which didn’t actually happen”. Also, according to participant No. 6, “I sometimes talk to my self about some random conversations, and I think it helps me to have a good organization of sentences”.

- Improves pronunciation and vocabulary recall through practice.

Another interesting fact the participants said is that self-talk helps them with pronunciation and vocabulary as well. Participant No.10 said, “When I self-talk, mostly words will flow into the mind, which actually helps a lot in real conversations as well” and No.13 confirmed it by “Actually, I think I have improved my vocabulary a lot through this self-talk thing”

V. CONCLUSION

This study investigates whether self-talk would be an effective strategy in improving the confidence in spontaneous conversation of ESL learners. Both quantitative and qualitative analyses show that there is a correlation between improving confidence in spontaneous conversations and self-talk training. The quantitative results showed that self-talk training has a positive impact on the participants’ confidence level with regard to the measures taken in six criteria. Qualitative results show that there are some challenges regarding spontaneous conversations, but with the help of the self-talk strategy, learners have seen significant improvements. This study reflects that self-talk would be an effective strategy to improve confidence in spontaneous conversations in ESL learners based on the findings.

The study shows a positive impact of self-talk in improving confidence in spontaneous conversations. It is recommended to use this strategy to improve ESL learners confidence level and speaking skills in spontaneous conversations. Future studies could consider how other factors such as age, gender, and proficiency level in English would affect these findings.

REFERENCES

[1] L. Humairoh, "The use of self-talk at the first grade, semester 1, academic year 2021/2022 of Vocational High School XYZ," J. Lang. Educ., vol. 1, no. 1, p. 29, 2022.

[2] H. K. Khairunnisa, H. Hariyono, and H. Roesellaningtias, "The effectiveness of using self-talk strategy to improve students' English speaking skill of the eleventh grade students at SMKN 1 Bagor," Dharma Pendidik.: J. Pendidik. Pembelajaran, vol. 17, no. 2, pp. 161–166, 2022.

[3] Y. P. S. Bataona, T. Feliks, and A. Semiun, "An analysis of spontaneous English conversations among second semester students of the English Education Department of San Pedro University in academic year 2017/2018," Indones. J. Educ. Rev., vol. 5, no. 2, pp. 11–29, 2018.

[4] S. Ahmed, "Attitudes towards English language learning among EFL learners at UMSKAL," J. Educ. Pract., vol. 6, no. 18, pp. 6–14, 2015.

[5] H. Alisoy and H. Oglu, "Navigating impromptu speaking strategies for successful spontaneous conversations," Web Semant.: J. Interdiscip. Sci., vol. 2, no. 1, pp. 1–8, 2024.

[6] R. Rahmon, Challenges and strategies in enhancing oral communication skills: A study in understanding challenges students face in developing oral communication skills in English. Örebro, Sweden: Örebro University, 2024.

[7] L. A. Alnaeem, "The power of self-talk for speaking anxiety reduction among EFL learners," Theory Pract. Lang. Stud., vol. 15, no. 1, pp. 103–112, 2025.

[8] H.-J. Jo et al., "Neural effects of one’s own voice on self-talk for emotion regulation," Brain Sci., vol. 14, no. 7, p. 637, 2024.

[9] F. Deamer, “Why do we talk to ourselves?,” Review of Philosophy and Psychology, vol. 12, no. 2, pp. 425–433, 2021.

[10] S. E. Huang and Y. T. Liu, "How to talk to myself: Optimal implementation for developing fluency in EFL speaking through soliloquizing," Engl. Teach. Learn., vol. 47, no. 2, pp. 145–169, 2023.

[11] S. Meriahna and A. Indari, "The effect of self-talk strategy on students’ speaking ability for eleventh grade of SMK Swasta Swakarya Salapian in academic year 2021/2022," VISION J., vol. 18, no. 1, pp. 16–31, 2022.

[12] T. Ismailia, "Performance assessment using rubric to improve students’ speaking skill," J. Appl. Linguist. Lit., vol. 5, no. 2, pp. 66–82, 2021.

[13] IBM Corp., IBM SPSS Statistics for Windows, Version 27.0. Armonk, NY: IBM Corp., 2021.

[14] V. Rosmayanti, R. Ramli, and R. Rafiqa, “Building beginners' self-confidence in speaking,” EduLite Journal of English Education, Literature and Culture, vol. 8, pp. 192–208, May 2023, doi: 10.30659/e.8.1.192-208.

Emily Brontë's 'Sense of Place' as Portrayed in Her Literary Works.

Pabasara Ponnamperuma
Department of English Language Teaching
University of Kelaniya
Kelaniya, Sri Lanka.
p.ponnamperuma95@gmail.com

Abstract*— 'Sense of Place' is a theory that defines the emotional attachment individuals develop with specific locations, encompassing both positive and negative feelings. It can be directly applied to the analysis of the Victorian writer Emily Brontë's portrayals of her house in Yorkshire moors where she grew up. It is reported that she loved her house, and it provided her with a constant back drop for her imaginative thinking and creative writing. It is assumed that the landscape seen through its windows and the sounds heard while being inside it frequently inspired most of the locational portrayals in her poetry and fiction. In consideration of Emily Brontë's romantic attachment to her home, the present study intends to explore her own sense of place as portrayed in her works. Accordingly, it pursues the research questions, "How is Emily Brontë's sense of place portrayed in her works, 'Wuthering Heights', 'Remembrance' and 'Fall, leaves, fall?' The methodology involves a thematic analysis of her biography by Edward Chitham, 'A Life of Emily Brontë' under the three themes, 'Emily Brontë's sensitivity towards nature and the environment,' 'Emily Brontë's emotional intensity,' and 'her imaginary encounters with recurring patterns of the orphans and abandoned characters.' The findings of the present study foreground that Emily Brontë's was heavily influenced by her romantic perception of the beauty of nature and her identification of the therapeutic power of nature as portrayed through her characters. According to her, nature tends to improve the quality of relationships among humans while showing the ways in which their deterioration starts due to the imbalance of their emotions including the sense of being abandoned.

Keywords*— Emily Brontë, Fall leaves fall, Remembrance, Sense of place, Wuthering Heights.

I. INTRODUCTION

Reference [7] 'Sense of Place' describes the atmosphere of a place and the quality of its environment, which gives a certain unique sense of well-being which makes the person return there often. Reference [4] points out Emily Brontë depicted Heathcliff and Cathy as inhabitants of the moors who created an ingrained bond with the landscape around them. Their attachment towards each other has been growing strongly. Subsequently, their love has delved through the moors. Simultaneously, Emily's self has been embedded in the characters of the novel through reflecting their behavioral patterns and their interactions with other characters. Reference [2] mentioned that Brontë was able to write a fantastic gothic novel, *Wuthering Heights* because of her interest in the moors since childhood. She used to listen to her housekeeper's folktales related to the moors and she was well – acquainted with both folktales and fairy tales. *Wuthering Heights* has become famous all over the world because of the writer's depiction of irresistible love between a couple who was not able to realize their love when they were alive and yearning for a reunion after death. Simultaneously, her poem, *Remembrance* also connotated the love which inspired in the same context as above. In the novel, *Wuthering Heights,* Heathcliff yearned for Catherine's love after her death. Similarly, the narrator of the poem, *Remembrance*, wishes her lover to be with her after his death. Accordingly, it elucidates Emily's re-creation of the same story in two different contexts. On the other hand, her poem, *Fall leaves fall* connotated the idea that the abandoned lover is in deep sorrow who meticulously expressed worried emotions through nature.

II. LITERATURE REVIEW

A. Brontë's life

Reference [5] Brontë's life and the works are integrated a combination of her sharp intelligence and the volatile range of the emotions, and she pursued her abandoned self and the intense feelings through the manner of poetry and novel. Moreover, Brontë's home was 'Haworth Parsonage' which was a four-square house, facing the gales from high moors and brooding over churchyard tombs below. Reference [5] further explained Brontë loved her house as much as any of the family members because it provided a constant back drop for her reading, writing and thinking. Brontë loved looking outside of the window. She could see the church yard and the church tower. Brontë loved the west wind which blew across from these moors from Lancashire. That is the environment that Brontë have been surrounded with. In such a context, Brontë

created her fictional characters to live through the moors. Reference [4] "the inhabitants of the moors, most notably Heathcliff and Cathy, developed such an ingrained bond with the landscape around them that the landscape becomes more representative of their self". In that manner, she used to create her own sense of place with dark and deep moorlands which she could see from outside of her window including the life that she had spent inside of the parsonage with less interaction with the outside world. Mostly, she enjoyed her life with her family members. Consequently, Brontë's participation in play acting and daily walks in the moorlands specifically influenced both of her poetry and novel. In addition, Reference [5] mentions that Brontë's poetry suggests her sensitivity towards the nature and the around. Her capacity for the sensitivity towards nature can be reflected as follows,

When there was a period of high wind, Emily had written the poem with the title, loud without wind was roaring. In those poetry lines, Emily recalls her childhood with the moments when Emily and Anne rose at dawn in order to run across the moors near Haworth [5].

B. Brontë's Sense of Place.

Brontë's faculty of imagination is powerful. Reference [11] mentions the imaginative world that was created by Brontë was healthier than the reality. The life that she had spent in Yorkshire moors had been reflected in her works On the other hand, Reference [9] mentions that Brontë was heavily influenced by the nature of rural Yorkshire that was surrounded by rolling hillsides and sweeping landscapes. Brontë had the powerful chains of the words which depicted the visualization of the open scenery of her surrounding, the moors, clouds, wind, sunlight, grey stones, mosses of heather. Her works have exposed the restless soul of Brontë herself. Reference [9] has further illustrated that her emotional intensity has been naturally shared with her eminent work of art, *Wuthering Heights.* The loss of her mother has remained as a subconscious trauma to Emily. Furthermore, the separation from the beloved was such a recurrent experience in Emily's childhood. Accordingly, it seems that her works have been integrated the same themes of separation and reunion more often. Reference [1] points out Brontë had a mysterious habit of creating fantasy worlds within which she repeated certain patterns. She used to put together all her obsessive thoughts in such a world. This world inhabited all her obsessive thoughts. Furthermore, Reference [6] mentions the creative writer is a neurotic individual, and the work of art is the expression of the unconscious mind of the individual. Reference [1] shows that both the exploration of the personal life of Brontë and the story of *Wuthering Heights* revealed Brontë as a neurotic woman who lived in an imaginative world where she was able to gratify her secret wishes. The unfortunate state of Brontë was further ignited her severe traumatic childhood experiences and the rigid social circumstances of the Victorian society Accordingly, the recurring themes of loss and separation and the recurring patterns of the orphaned and abandoned characters were showcased throughout Brontë's works. Reference [1] further mentions that Brontë's personal experiences, obsession with the fear of loss that she had experienced in her childhood had been depicted in her works.

III. METHODOLOGY

The present research uses qualitative research design. Thematic analysis has been used to analyze Edward Chitham's A *Life of Emily Brontë (1987)*, a biography of the writer's life. Subsequently, the novel, *Wuthering Heights (2002)* and the poems, *Remembrance* and *Fall leaves fall* have further been analyzed using thematic analysis.

IV. RESULTS AND EVALUATION

A. Brontë's the application of Sense of Place in 'Wuthering Heights'

1. Emily Brontë's sensitivity towards nature and the environment

Brontë was heavily influenced by nature of rural Yorkshires that which was surrounded by rolling hillsides and the sweeping landscapes. The beauty and the serenity of nature had been depicted several occasions in the story. Yet, the most effective instances have been selected for the present study. Accordingly at the beginning of the novel, Brontë has introduced the principal location of the novel and its relation towards nature and the surrounding.

'Wuthering Heights' is the name of Mr. Heathcliff's dwelling. Wuthering being a significant provincial adjective, descriptive of the atmospheric tumult to which its station is exposed in stormy weather. Pure, bracing ventilation they must have always up there, indeed: one may guess the power of the north wind, blowing over the edge, by the excessive slant of a few stunted first at the end of the house [10].

Brontë has referred to the healing power of nature towards her characters in the novel. After Emily was healed thanks to the well-care of Mr. Linton. Emily was amazed to see the surroundings. She was fascinated by the nature of the moors. Emily is in a discussion with Mr. Linton. Both were amazed by nature which has been taken priority in their talk. Furthermore, nature has been acting as a backdrop of developing their relationships. As soon as Catherine was cured, she was happy to see nature. Simultaneously, it brings back her memories of *Wuthering Heights*.

These are the earliest flowers at the Heights!" She exclaimed. They remind me of soft thaw winds, warm sunshine, and nearly melted snow. I only see two white spots on the whole range of moors the sky is blue, and the larks are singing. Catherine: last spring at this time, I was longing to have you under this roof now, I wish you were a mile or two up those hills; the air blows so sweetly, I feel that it would cure you [10]. Not only that, but the characters have also been healed because of nature. The presence of nature improves the quality of their relationships with each other. Catherine Linton was quarrelling

with Linton Heathcliff on the moors. On the contrary, nature acted as a facilitator to cure their moods. Nature makes both to be in a good and pleasant spirit. Every detail of nature is much appreciated. It seems that Brontë herself gives attention to every small detail of the nature.

He said the pleasantest manner of spending a hot July day was lying from morning till evening on a bank of heath in the middle of the moors, with bees humming dreamily about among the bloom, and the larks singing high up overhead blue sky and bright sun shining steadily and cloudlessly [10].

2. Emily Brontë's emotional intensity

Brontë is well known for the emotional intensity of her works. Her narrative is emotionally unique because not only does her character experience certain emotions, but the readers too are also able to deep dive in those emotions with the characters. Her usage of emotions is in different layers. Mr. Lockwood had been facilitated in a room in Wuthering Heights which had been announced to not to be used. Subsequently, Mr. Lockwood was compelled to experience a ghostly dream. He had encountered the ghost of Catherine, who emotionally lamented over mentioning she had been waiting for Heathcliff for over 20 years. The readers have experienced intense emotions of Catherine through Brontë. The readers might not be afraid yet feel sorry for the isolated spirit. Or else, the readers might feel shocked and the abandoned because no one has expected this gothic atmosphere in the middle of the romance story. This scenario has provided much evidence that Emily had a need to create love that had gone beyond the graveyard.

let me in – let me in! Who are you? …. 'Catherine Linton,' it replied shiveringly. I'm home; I'd lost my way on the moor!' It's twenty years', mourned the voice, 'twenty years, I have been a wait for twenty years [10].

Catherine Earnshaw used to hang out with Heathcliff on the moors. Suddenly, as she is encountered by Edgar Linton, she has forgotten Heathcliff. Simultaneously, she mentioned that it is risky to marry Heathcliff. Yet, she expressed her love towards Heathcliff. She mentioned the following to her servant, Ellen'. Her parole seemed to sound emotionless. If she loves Heathcliff to such an extent, she should accept Heathcliff with whatever circumstances that he has. She should not forget Heathcliff at the presence of Edgar Linton.

He shall never know how I love him; and that, not because he's handsome, but because he's more myself than I am. Whatever our souls are made of, his and mine are the same, Linton is different." My love for Linton is like the foliage in the woods. My love for Heathcliff resembles the eternal rock beneath [10].

Brontë was able to depict the chaotic nature of Catherine who seemed to start loving her childhood friend after getting married to sophisticated Edgar Linton. Edgar was angry about the indecent behavior of Heathcliff at his own home and requested Catherine to select between the two. Catherine was ashamed of herself. She knew Edgar was talking with a point that she was not able to bear. She got sick for a longer week, which caused damage to her emotional balance. Moreover, Isabella Linton is a woman who again lost her emotional balance due to her being a jealousy, desperate, helpless and frustrated wife of Heathcliff. She has written this letter to Nelly,

Is Mr. Heathcliff a man? If so, is he mad? And, if not, is he a devil? I shan't tell my reasons for making this inquiry; but I beseech you to explain, if can, what I have married to that is [10].

Brontë shows the ways in which the person's deterioration started due to the imbalance of the emotions. Emotions play a significant role in every person's life. Furthermore, the death of Catherine Linton happened solely because of Heathcliff's presence. Before Heathcliff came, the life of Edgar and Catherine was a happy one. After Heathcliff became a gentleman, Catherine had seen two capable lovers: Heathcliff and Edgar. She had emotionally drained. Catherine had showcased a wild and a passionate love towards Heathcliff. That's why, when Heathcliff had come to visit Catherine at her sickbed, she could not resist kissing Heathcliff while forgetting Edgar Linton.

He neither spoke, nor lost his hold for some five minutes, during which period he bestowed more kisses than ever he gave in his life before. Edgar will not hurt us [10].

However, Heathcliff was immensely worried about the death of Catherine. He seemed to be pure with his expressions and emotions. He only loved Catherine. He did not love Isabella though he married her. The death of Catherine had made Heathcliff to be the demon of Wuthering Heights. He had become a mentally imbalanced person who wanted to take revenge on Edgar Linton. He wanted to make the reunion of her son, Linton Heathcliff and Catherine's daughter, Catherine Linton. During the last days of Heathcliff, he had shown a strange behavior where he went to the moors and always his mind was haunted by Catherine. The following few lines were uttered by Heathcliff intensely.

You said I killed you, haunt me then, the murdered do haunt their murders. I believe ghosts have wandered on earth. Be with me always. Take any form, only do not leave me in this abyss, where I cannot live without my soul [10].

3. The recurring patterns of the orphaned and abandoned characters

At the beginning of the novel, Heathcliff arrived 'Wuthering Heights' as an orphaned boy who is without the parents. The young Heathcliff had been treated much unfairly by Hindley Earnshaw. Hindley Earnshaw had been the beast of the 'Wuthering Heights.' The same footsteps have been taken by Heathcliff. He had become the Master of Wuthering Heights, and he treated them badly for the young Hareton Earnshaw. Catherine Earnshaw had married Mr. Edgar

Linton. So, she became 'Catherine Linton.' Subsequently, their daughter had been named as 'Catherine Linton.' The names have been repeated. Then, the sons of Heathcliff and Isabella Linton had been named Linton Heathcliff.' When the

two kids were playing at Wuthering Heights, they were able to find the goods of first Catherine and Heathcliff. It seemed that Emily was trying to share the emotional weight of her heart with her readers through repeated character cycles. There are three main characters as follows, Catherine, Heathcliff and Linton whose presence often had been appeared in the story. In addition, specifically the abandoned, helpless and parentless child and the man who becomes a beast due to the death of loved ones appeared repeatedly through the characters; Hindley Earnshaw, Heathcliff and Harethon Earnshaw.

When Catherine and Linton had thought of playing ball, they were able to find the hoops, battle doors and shuttlecocks which named C and H. So, Catherine mentioned that she 'wished to have C., because that stood for Catherine, and the H. might be stood for Heathcliff [10].

B. Brontë's application of Sense of Place in 'Remembrance.'

1. Emily Brontë's sensitivity towards nature and the environment

Nature plays a vital role because the narrator depicts her despair emotions while referring the nature. In that manner, the emotions of the narrator have been reflected through nature. The narrator is in utter despair due to the death of her lover. The narrator is still reminiscence of her dead lover who is lying beneath the deep snow. That's why, she mentions 'Cold in the earth—and the deep snow piled above thee, Far, far removed, cold in the dreary grave!' She mentions that she lives because of her memory of him. Their loving memories are with the environment, with the 'mountain, northern shore which is covered with fern leaves.' The readers can visualize the autumn period where the dead leaves have touched the earth and are covered with leaves. The intensity of the emotions of the narrator have been projected through the visuals of nature. The cold earth and the dead leaves provide the melancholy atmosphere that the narrator is currently living by. Moreover, the emotions of despair, disappointment, yearning, isolation and love are depicted with intensity.

Cold in the earth—and fifteen wild Decembers, From those brown hills, have melted into spring: Faithful, indeed, is the spirit that remembers. After such years of change [3].

2. Emily Brontë's emotional intensity

She survived for the past fifteen years without the presence of her lover. Yet, her soul is immensely grateful towards her lover. That's why she has mentioned that though the years have passed, her suffering without his presence continues. Moreover, her infidelity towards her dead lover is surprisingly marvelous because she confesses that other men has come her way, yet she rejects them all for the sake of her loyalty towards her dead lover. That's why, she has mentioned that "Other desires and other hopes beset me, hopes which obscure, but cannot do thee wrong!". The narrator seemed to be hopeless and helpless without the presence of his lover because all her life's blessings have been received with the presence of her lover. In that manner, she believes that all her life's blessings were gone into the grave with him. Specifically, the following verse of the poem is echoed in Emily's novel, *Wuthering Heights*. The verse is, 'Down to that tomb already more than mine.' 'In the above novel, Catherine Earnshaw mentions that 'Heathcliff is more myself than I am' The above scenario showcased Emily's depiction of intensified passionate love which has become hopeless because of the death of her lover. Yet, she celebrates the unrequited love which is the most suffocating. She compares her lover to the tomb and deeply in love with his tomb. The narrator questioned how she loves another man after tasting the divine love of the man who lies beneath the soil. Her world is empty without him. She mentions that 'Once drinking deep of that divinest anguish, how could I seek the empty world again?' She compared the love that she shared with the man is like, drinking a divinest drink and she is in desperate state and questions how she will taste another mediocre drink because she knows the exact taste of the divine drink. She had become abandoned in this world without the presence of her lover. Surprisingly, the exact situation has been echoed in Heathcliff's response to the death of Catherine.

3. The recurring patterns of the orphaned and abandoned characters

In Brontë's novel, *Wuthering Heights*, the orphaned and the abandoned characters were recurrent occasionally. Similarly, in the present poem, the narrator is an abandoned lover who becomes isolated in the world without her lover beside.

C. Brontë's application of Sense of Place in 'Fall leaves fall.'

1. Emily Brontë's sensitivity towards nature and the environment

Brontë mentions that every leaf whispers a good thing to her. Yet, these leaves belong to a tree which puts down their own leaves during the season of the Autumn. Ironically, the above verse shows that the narrator is no longer happy. Yet, she mentions that she is happy to see the snow where the rose should grow. It seems that she does not have anything to be happy with. Nature itself facilitates the melancholic mood of the narrator. Her strong mind is depicted because she is ready to be happy when everything is dreary and dull. In that manner, nature has been used as an instrument to communicate the moods of the characters.

2. Emily Brontë's emotional intensity

Brontë meticulously uses the natural elements in order to highlight the emotional intensity of the characters. The title of the topic itself depicts the melancholy mood of the narrator. Though the season of the fall has left the narrator, which has returned towards her within an instant. The night melancholic hours rather than the happy and joyful hours in her life. The narrator is utterly sad. It seems she is abandoned. That might be the reason that she is able to celebrate the melancholic moments with so much power visual imagery. is longer than the daytime.

The above factor connotates the idea that the narrator must go through longer darker and

Fall, leaves, fall; die, flowers, away;
Lengthen night and shorten day;
Every leaf speaks bliss to me
Fluttering from the autumn tree.
I shall smile when wreaths of snow
Blossom where the rose should grow;
I shall sing when night's decay
Ushers in a drearier day [3].

3. *The recurring patterns of the orphaned and abandoned characters*

The narrator seemed to be lonely who feels life as monotonous and boring because the nights are longer and the day times are short. The narrator is not accompanied by anyone. No human is available to speak bliss to them. No human is available to smile at them. Instead of that, the narrator lives alone with nature. The narrator suffers from the feeling of getting abandoned as they are happy to welcome winter over the spring and sing along their dull and monotonous days.

V. CONCLUSION

As Brontë had lived with nature, she had made her narrators of the poems to be with nature. Accordingly, one can understand the fact that Brontë's sense of place has been depicted through her works. Specifically, her sense of place has been limited to her surroundings, her emotions and Brontë as a lonely person how does she console herself at the presence of the nature. Nature has done the work of another human when it comes to the context of Brontë. Her life was limited to her love for nature. Specifically, towards moors. The emotions have been directly impacted by nature. She has depicted the abandoned character who is in search of love. Nature has played a significant role in recognizing the true emotions of the characters. Accordingly, Brontë's works are just repeated patterns of human relationships. Specifically, the lover is abandoned by the other lover. Therefore, he/she is in desperate search towards love. The above-mentioned scenarios were repeated in both the novel, *Wuthering Heights,* the poems, *Remembrance* and *Fall leaves fall*. Specifically, the abandoned, helpless and parentless child and the man who becomes a beast due to the death of loved ones appeared repeatedly through the characters. Nature has constantly enriched such emotional human relationships. In that manner, Brontë's sense of place contributed to the production of her works including shaping her writing career. Brontë represents the European culture in the Victorian society. Accordingly, the future recommendations can be made towards research on the reflection of sense of place in South Asian culture including the ways in which the cultural, social and other political factors have been contributed on the flourishment of writer's perspective.

REFERENCES

[1] A.P Moussa, "Recurring Patterns: Emily Brontë's neurosis in Wuthering Heights," International Journal of Education and Literary Studies,Jan.2014.[Online].Available:https://journals.aiac.org.au/index.

[2] B. Voroselo, "The Non- specifity of Location in Emily Brontë's Wuthering Heights." Unpublished.

[3] C.W Hatfield, Ed. The Complete Poems of Emily Jane Brontë. New York, South Carolina (only U.S.), USA: Columbia University Press, 1996.

[4] E. Bevin, "Harmonized by the earth: Land, Landscape and Place in Emily Brontë's Wuthering Heights." Unpublished.

[5] E. Chitham, A Life of Emily Brontë. Oxford, UK: Basil Blackwell, 1987.

[6] E.S Person, P. Fonagy and S.A Figueira, Ed. On Freud's Creative Writers and Day Dreaming. Yale University, USA: Routledge Taylor and Francis group, 1995.

[7] J.B Jackson, "A Sense of Place, a Sense of Time," Spring, 1995. [Online].Available: https://www.jstor.org/stable.

[8] J.B Jackson, "A Sense of Place, A Sense of Time," OZ , 1986. [Online]. https://doi.org/10.4148/2378-5853.1111.

[9] M. Roth. "Emily Brontë and Gothic Literature." Mandy Roth. Accessed:February23,2024.[Online.]Available:https://mandyroth.com/gothic-literature-emily-bronte/.

[10] N. Pauline, Ed. Wuthering Heights (Penguin Classics). London, UK: Penguin Classics, 2002.

[11] S. Manzoor. "A Tribute to Emily Brontë: Gondal: The Fanciful World of Emily Brontë." The Daily Star. Accessed: February 24,2024. [Online.]Available:https://www.thedailystar.net/literature/news/gondal-the-fanciful-world-emily-

Enhancing Pronunciation Proficiency via Listening Practice for Specific information with the help of ELSA Speak: Evidence from Foundation-Level Students at SLIIT City Uni, Sri Lanka

Priyamanthi Thennakoon
Department of Health and Life Sciences
SLIIT City Uni
Colombo 03, Sri Lanka
priyamanthitennakoon@gmail.com

Ahinsa Marasinghe
Department of Health and Life Sciences
SLIIT City Uni
Colombo 03, Sri Lanka
ahinsamarasinghe@gmail.com

Mindi Sadithma
Department of Health and Life Sciences
SLIIT City Uni
Colombo 03, Sri Lanka
mindisandinma62@gmail.com

Nethanga Gunaratne
Department of Health and Life Sciences
SLIIT City Uni
Colombo 03, Sri Lanka
nethangagunaratne@gmail.com

Rishini Mendis
Department of Health and Life Sciences
SLIIT City Uni
Colombo 03, Sri Lanka
rashaniamandi@gmail.com

Ishadi Nilaweera
SLIIT City Uni
Colombo 03, Sri Lanka
ishadi.n@sliit.lk

***Abstract*—This study investigates the impact of English Listening skills for specific information (ELSSI) on English pronunciation (EP) among 40 Foundation semester 2 students at SLIIT City Uni. A mixed approach of questionnaires and experimental tests as research instruments, along with the AI tool, ELSA Speak, was applied to this group. Participants received structured, listening-based practice over the past few days by employing the ELSA Speak AI application. The Data were analyzed by using the Statistical Package for the Social Sciences (SPSS) and Thematic analysis. Therefore, the results show a statistically significant improvement, based on the means of the inclusive posttests (83.49) being higher than the pretests (77.20), highlighting the effectiveness of targeted listening activities in enhancing pronunciation. The increase in the mean score was discovered after the listening treatments by the researchers. An increase in student scores was observed following the treatment. Positive feedback of the sample demonstrated the above-mentioned points. These results lead to the suggestion that ELSSI can help with the development of oral proficiency, while mobile-assisted language learning tools (ELSA Speak) play a crucial role when focusing on the improvement of oral proficiency, further offering valuable insights for educators and language enthusiasts around the world.**

***Keywords*—ELSA Speak application, Homophones, Listening skills, Media, Pronunciation,**

I. INTRODUCTION

Pronunciation is a key component of speaking which enhances the meaningfulness of the English Language (EL). According to Adawiah, Muliati & Samtidar (2024) pronunciation is one of the prominent aspects that others will judge. Therefore, Adawiah, Muliati & Samtidar (2024) mentioned that there are several aspects that must be considered in speaking in EL. Pronunciation is one of the first impressions that other people will catch grasp swhen starting a conversation in English. Listening skills and related activities play a vital role in order to achieve better speaking skills in EL. Pronunciation and listening comprehension skills are interconnected; improving one often enhances the other. Specifically, focusing on the sounds of a language and practicing pronunciation can significantly improve listening skills and vice versa. The researchers observed that due to poor awareness among students, the importance of listening to improve one's pronunciation has been ignored, even while considerable mobile applications exist. This research is mainly aimed at helping students improve their pronunciation by using various listening exercises, as it is the main component that's directly linked to one's pronunciation. This also sheds light on a helpful mobile application for the self-improvement of students. Additionally, the research

focuses on enhancing the concern of students in the direction of listening.

Thereby, the present research points towards how listening impacts the enhancement of pronunciation with the help of the mobile app, 'ELSA Speak'.

II. LITERATURE REVIEW

A. Pronunciation and Listening

Listening skills and related activities play a vital role in order to achieve better speaking skills in EL. Pronunciation and listening comprehension skills are interconnected; improving one often enhances the other. Specifically, focusing on the sounds of a language and practicing pronunciation can significantly improve listening skills and vice versa. In language learning, pronunciation and listening are two sides of the same coin. Through a grasp of phonetics — the sounds of a language — serves as the foundation for both. (Miller, 2024).

B. Listening skills for specific information and Listening activities

Listening is an activity of listening to verbal symbols carefully with concentration so that the information transmitted can be obtained and appropriately understood. Paying attention to the sound, then identifying, monitoring, assessing, and then reacting to the sound is a skill needed for someone to communicate (Rahmawati & Br. Sianturi, 2021).

C. ELSA Speak application

Adawiah, Muliati & Samtidar (2024, p.65) stated that "English Language Speech Assistant is an acronym for ELSA. English Language Speech Assistant (ELSA) can be downloaded for free from the App Store or Google Play. In the Elsa application itself, there are several categories, namely daily lesson exercises, improved pronun ciation, and topic-based learning (there are 14 topics here, and we can choose according to our abilities). In Elsa Speak, there are also phonetic symbols for each word, and we can also listen to how the word is pronounced".

III. METHODOLOGY

The present study aims to determine whether the English Pronunciation (EP) has an impact on the English Listening Skills for Specific information (ELSSI) in the Sri Lankan context.

This model was developed by the author himself because none of the previous studies have investigated this relationship using EP and ELS. Therefore, this conceptual framework is developed based on based on previous literature. It is evident that people feel speaking and listening are the most important skills to learn. The variables used for this study are ELS as an independent variable and EP as the dependent variable.

A. Abbreviations and Acronyms

Most of the variables involved in this sss udy are instinctive, which means most of them are subjective rather than objective. Therefore, a clear-cut definition or commonly accepted definition for these variables cannot be foundcannot be found. Hence, the defining scope of the present study is given by the following definitions.

EP - "The way in which a word or letter is said, or said correctly, or the way in which a language is spoken" (dictionary.cambridge, n.d.), "The act or manner of pronouncing something" (Merriam-Webster, n.d.), The way in which we make the sound of words" (English Club, n.d.)

ELSSI - "Paying attention to verbal and nonverbal cues to gain a full understanding of the speaker's experiences, thoughts, feelings and objectives" (Indeed, 2025), "Paying attention to the use of language" (Agwuocha, 2024), "Receiving language through the ears" (English Club, n.d.)

B. Research Design

The study used a mixed methods approach, with a dominant quantitative aspect.

Population of the study - Since the study was conducted to identify how EP can be enhanced by ELSLS the study population included all the Sri Lankan people who speak and willing to speak EL.

Sampling of the study -The geographical area for this study was selected as Colombo district, due to convenience. Gender wise, both males and females were taken to the sample and all of them are students attending SLIIT City University (SLU) foundation program to and age between 16-25 years were proportionately drawn as the sample for the study. 40 participants were taken as the sample for this study, and they were selected using purposive sampling technique, which is a non-probability sampling method as Adawiah, Muliati & Samtidar (2024, p.65) stated that In choosing the research participants, the researchers used a purposive sampling technique; on basis of that sample has been selected

Research Instrument - The questionnaire was used as a research instrument. A pre-test and a post-test questionnaire were provided for the participants. The purpose of the pre-test is to determine the students' EP level before using the listening activities, and the purpose of the post-test is to understand the students' improvement in EP after using the listening activities as treatment.

C. Procedure for Data Collection

The study was conducted in three phases, including pre-test, treatment, and post-test. During the pre-test, a

questionnaire was provided for the 30 participants to collect the background information of the sample, listening habits, proficiency skill level and personal ideas regarding the relationship between ELSSI and EP. Then, each participant was provided with 5 pronunciation tests covering, (1) vowel exercise, (2) Consonant exercise, (3) Homophones exercise using words that thatsound alike but differ in meaning and spellings, (4) commonly mispronounced word test and (5) tongue twisters followed by measuring the time and pronunciation clearance level which were measured by ELSA Speak application. The researchers designed the above-mentioned tests with the assistance of the *Preply* blog (Wilhelmi, 2024).

		Mean	N	Standard Deviation	Standard Error Mean
According to the Average of all 5 pretests and 5 post tests	Pre Test	77.20	40	16.14	1.47
	Post Test	83.49	40	16.34	1.49

Based on the research that has been done by Adawiah, Muliati & Samtidar (2024, p.65), they justified their experimental method with the help of Sugiyono (2014) by stating pre-experimental designs are designs that include only one group or class undergoing pre- and post-tests. In this study, the researchers only used the experimental group without the control group.

After completing the 5 stages of exercises, those will be uploaded to ELSA Speak AI application to monitor and measure the pre-test proficiency in EP. Then participants were provided with listening treatments by specific information about vowel sounds,consonant sounds,homophones,and tounge twisters where they can listen to selected audio clips covering all five exercises explained above and practice the pronunciation with a 1day interventional period for each test. After the practice period, those 5 exercises were again conducted and measured through ELSA Speak application to check on the progress and improvement of EP. After the post-testing period was conducted, a post-stage questionnaire was provided for all the participants.

D. Technique of Data Analysis

The study analyzed data using the Social Science Statistical Package (SPSS) and Thematic analysis.

IV. RESULTS AND EVALUATION

A. Data analysis

Pre-stage questionnaire analysis: As an entrance to the research, a questionnaire was given to the students who were participating in the tests by the researchers so that the current level of each student could be discovered through this questionnaire.

Pretest and Posttest: Measured the frequency and the rate of percentage of the pre-test and post-test scores of students' pronunciation of each test (vowel Test, consonant Test, homophones Test, commonly mispronounced words Test, Tongue twisters Test)

Students' mean score and standard deviation in the pretests and the posttests:

TABLE 1: STUDENTS' MEAN SCORE AND STANDARD DEVIATION IN THE PRE-TEST AND THE POST-TEST

TABLE 1 shows that the students' mean score on the pre-test was 77.20. This score is in the fair classification. The table also showed that the students' mean score in the post-test was 83.49, which is in the good classification, but with a significant improvement.

It can be concluded the listening skills/activities that were provided by researchers emphasized the progression of EPS and the AI tools/mobile applications like ELSA Speak can assist students to satisfy satisfy their desire by using them as self-learning tools.

V. CONCLUSION

Listening skills/activities for specific information are effective in students' English pronunciation development and enable the learners to engage in pronunciation skills or conversation with confidence. According to the findings, the inclusive posttests (83.49) were higher than comparison to the pretests (77.20). The increase in the mean score was discovered after the listening treatments by the researchers. This can be classified as an increase in student scores after being treated. The responses of the feedback survey also prove this particular point.

The researchers discovered the Homophones exercise as the best test among the five tests based on the posttest results and feedback survey.

Assistance with ELSA Speak, and other reliable AI tools/ Mobile applications can help to enhance the EPS, which is recognized by the students of the sample as the researchers expected.

In contrast, consciously focusing on the specific information while listening is confirmed by students of the sample as an important factor that can be distributed to the EL learners.

REFERENCES

[12] H. Harsin, R. Rifiqoh, and M. Mashuri, "Using homophone games to improve pronunciation for high school student," E-Journal of ELTS (*English Language Teaching Society*), vol. 10, no. 3, pp. 227–236, 2022. [Online]. Available: https://doi.org/10.22487/elts.v10i3.3182

[13] R. Adawiah A. Muliati and S. Samtidar, “The Effect of Elsa Speak Application on Students’ English Pronunciation Development,” *Jurnal of English Education, Language, and Literature (JoEELE)*, vol. 3, no. 1, Jan. 2024. [Online]. Available: https://ojs.unm.ac.id/JoEELE/article/view/59627

[14] R. Sadilah and R. R. L. Gaol, “Improving students’ English pronunciation through listening to English songs at the second grade students of SMP Swasta Harvard Martubung,” Journal MELT (*Medium for English Language Teaching*), vol. 6, no. 1, pp. 15–30, 2021.

[15] L. F. Mantilla-Cabrera, C. A. Narváez-Vilema, and M. A. Carrillo-Rodríguez, “Listening comprehension to improve pronunciation in students of Mechanics Career-ESPOCH. A linguistic analysis,” *Polo del Conocimiento*, vol. 5, no. 3, pp. 882–902, 2020.

[16] J. Alzamil, “Listening skills: important but difficult to learn,” *Arab World English Journal*, vol. 12, no. 3, pp. 366–374, 2021.

[17] E. Rahmawati and M. Sianturi, “Audio podcast-based learning media in improving students’ listening comprehension and pronunciation,” in Proc. *4th Int. Conf. Lang., Lit., Cult., and Educ. (ICOLLITE 2020)*, vol. 546, pp. 356–362, 2021.

[18] H. Vančová, “AI and AI-powered tools for pronunciation training,” *Journal of Language and Cultural Education*, vol. 11, no. 3, pp. 12–24, 2023. [Online]. Available: https://doi.org/10.2478/jolace-2023-0022

[19] A. Matysiak, “Students’ perceptions on the use of selected mobile apps in the process of acquiring L2 pronunciation – a preliminary study,” *Językoznawstwo*, vol. 2, no. 19, pp. 209–224, 2023. [Online]. Available: https://doi.org/10.25312/j.6840

[20] J. D. Muhammed and S. B. Abduljabbar, “The impact of pronunciation practice on English listening comprehension: A case study of high school students in Duhok*,” Academic Journal of Nawroz University*, vol. 12, no. 4, pp. 17–25, 2023. [Online]. Available: https://doi.org/10.25007/ajnu.v12n4a1895

[21] *Preply*, “10 English pronunciation practice exercises & printables,” Preply, 2024. [Online]. Available: https://preply.com/en/blog/english-pronunciation-practice/

[22] E. Miller, “The link between pronunciation and listening skills: A guide for English language learners,” *Medium*, 2024. [Online]. Available: https://medium.com/@emmamillerw1990/the-link-between-pronunciation-and-listening-skills-a-guide-for-english-language-learners-6a88d802da2b. [Accessed: Apr. 7, 2025].

[23] *Read Naturally*, “Audio Examples of Phonics Sounds,” Read Naturally. [Online]. Available: https://readnaturally.com/phonics-sounds. [Accessed: Apr. 7, 2025].

[24] *Speech Active*, “40 Commonly Mispronounced Words in English - Free Lesson,” Speech Active. [Online]. Available: https://www.speechactive.com/english-commonly-mispronounced-words-free-lesson/. [Accessed: Apr. 7, 2025].

[25] *BYJU’S*, “75 Tongue Twisters in English,” BYJU’S. [Online]. Available: https://byjus.com/english/tongue-twisters/. [Accessed: Apr. 7, 2025].

ARCSCU'25
ANNUAL RESEARCH CONFERENCE OF SLIIT CITY UNI

Organized by
SLIIT CITY UNI

ARCSCU 2025

SLIIT City Uni,
Level 13, BoC Merchant Tower,
St Michael's Road,
Colombo 03,
Sri Lanka.

SCU
SLIIT CITY UNI

https://arcscu.sliitcityuni.lk/
researchscu@sliit.lk

Tel : +94 117543612