

Fluttering into Fluent English: Building an Interactive Voice-Based AI Learning App for Language Acquisition

Akhila¹, Anagha P Nair², Jyothsana K N³, Sreepriya S⁴, Akhila E⁵

¹⁻⁴Department of CSE, JCET, APJ Abdul Kalam Technological University, Palakkad, India

⁵Assistant Professor, Department of CSE, Jawaharlal College of Engineering and Technology, Palakkad, India

Email: ¹mailme.akhilakumar@gmail.com, ²anaghapnair23@gmail.com, ³knjyothsana5@gmail.com, ⁴sreepriyashobhanan@gmail.com, ⁵mailtoakhilasathish@gmail.com

Abstract

The rapid progress in technology has transformed the way people learn languages, offering students innovative tools to improve their skills. SpeakBuddy is built using the Flutter framework and utilizes Azure Speech AI for speech functionalities. Speak Buddy incorporates various features, such as interactive conversation practice, detailed lessons covering vocabulary, grammar, and pronunciation, fun games, and AI-powered voice assistance. The app's creation involved several design cycles where feedback from language experts and user testing was used to ensure it was easy to use and effective. Continuous assessments were conducted to evaluate the app's usability, performance, and impact on language learning outcomes. The results showed high levels of satisfaction among users, who reported feeling more confident and proficient in spoken English after using the app. A key feature of Speak Buddy is its ability to provide instant feedback on pronunciation, helping students improve their speaking skills.

Keywords: Pronunciation Skills, Flutter, Self-Learning, Artificial Intelligence, Spoken English

1. Introduction

In today's world, knowing English well is extremely important because it helps people communicate and collaborate, and opens up many opportunities worldwide [1]. To help students get better at speaking English, we've created an innovative solution called Speak Buddy: the English Speaking App [3]. We've made sure that Speak Buddy focuses on what users need most. It's not just about memorizing words or phrases; it's about actually speaking and getting better over time. SpeakBuddy is designed to address common challenges faced by language pupils, such as limited opportunities for conversational practice, lack of personalized feedback, and difficulty in accessing quality learning resources [4]. During the development of the SpeakBuddy app, iterative design cycles were employed to refine and improve the user experience based on feedback from language experts and user testing sessions. The design process involved multiple iterations, each incorporating feedback and insights gathered from various stakeholders [5-8]. Language experts helped us with SpeakBuddy. They gave us tips on how to teach languages, what lessons to include, and how to make learning fun. We listened to them to make sure our app had the best learning stuff and was easy to use[9-12]. This ensured that the app's educational content was accurate, engaging, and aligned with best practices in language education. We also asked real people who were learning English to try out the app and tell us what they thought. We had all kinds of learners, from beginners to advanced, try it out. They did different tasks on the app and told us what they liked and didn't like[13].

After hearing from both the experts and the learners, we made changes to the app to make it better. We fixed lessons, improved how the app listens to pronunciation, and made it easier to use. This helped us make SpeakBuddy more interactive and useful for everyone. Key features include interactive conversation practice, comprehensive lessons covering vocabulary and grammar, pronunciation training with speech recognition technology, gamified learning, and AI-powered voice assistance [14]. In addition to its rich feature set, the SpeakBuddy prioritizes accessibility, inclusivity, and affordability, ensuring that SpeakBuddy is available to all, regardless of geographical location or socioeconomic status [15-16].

1.1 Problem Statement and Significance

Most of the existing language learning methods suffer from several limitations, including limited opportunities for interactive conversation practice, lack of personalized

feedback, and difficulty in accessing quality learning resources. Traditional methods mostly just give out information without much interaction. This might not keep learners interested or help them with their own way of learning. Additionally, existing systems doesn't use new technology like artificial intelligence (AI) and speech recognition to provide real-time feedback and personalized learning experiences. Because of these issues, learning can be tough, and you might not get as good as you want.

Speak Buddy is a new way to learn English that's way more fun and helpful than the old methods. It's like having a smart friend who helps you get better at speaking English. With Speak Buddy, you can practice talking and get feedback right away. It uses latest technologies like AI and speech recognition to make learning feel like a game. This makes learning English easier and more enjoyable. The goal is to make learning English easier for everyone, so they feel more confident when they talk.

2. Related Work

In the study called "Unlocking English Proficiency: Assessing the Influence of AI-Powered Vernacular Acquisition Apps on Young Learners' Vernacular Acquisition," by researchers Vadivel, Balachandran & Shaban, Ahmad & Ahmed, Zeravan & Saravanan, looked at how special language apps affect young people learning English [2].

They found out that these apps, like the ones they studied, have cool features that help you learn better. These include talking to the app, playing games, and having lessons that are just right for you. The researchers also discovered that these apps can make learning English more fun and help you get better at it. They saw that when kids used these apps, they felt more confident in speaking English and enjoyed learning more.

Overall, the study showed that using these special apps can be a great way for kids to learn English. They can learn at their own pace, have fun while doing it, and feel more confident in their language skills. SpeakBuddy represents a practical implementation of innovative language learning features, the research paper likely offers insights into the broader impact and effectiveness of AI-powered vernacular acquisition apps on young learners' language acquisition processes and outcomes.

2.1 Background Work

This section includes apps like Duolingo, Rosetta Stone, and Babbel, which also help people learn languages like English. Duolingo is popular and has interactive lessons and games, just like Speak Buddy. But it doesn't have enough practice for speaking, which is a downside. Rosetta Stone is famous for making learning feel like real life. It has lessons where you learn in different situations, just like you would in real life. But it can be expensive. Babbel is another app made by language experts. It has lessons for different levels of learning, focusing on everyday language skills. But it might not feel as real as other apps. Table 1 shows the comparison between the existing apps and SpeakBuddy.

These related works offer valuable insights into various approaches and provide inspiration for the design and implementation of SpeakBuddy's features and functionalities. By drawing upon the strengths and innovations of existing platforms and overcoming their weaknesses, SpeakBuddy aims to offer a unique and effective learning experience.

Table 1. Comparison of Existing Learning Apps with SpeakBuddy

App	Usability	Review	Feature	Performance
Speak Buddy	Intuitive interface	Positive feedback	Focus on speaking	Efficient learning
Duolingo	Gamified approach	Mostly positive	Ads and interruption	Effective basics
Rosetta Stone	Immersive learning	Mixed user reviews	High cost	Quality content
Babbel	Clear layout	Generally Positive	Subscription model	Conversational skills

3. Proposed Work

SpeakBuddy is a comprehensive English learning app designed to facilitate language learning through interactive features, personalized learning pathways, and innovative technology. The system integrates various modules and functionalities to offer clients an

engaging learning experience, including support for native inbuilt languages like Malayalam. Figure 1 shows the architecture diagram of SpeakBuddy. Components of the Proposed System:

User Interface (UI): The User Interface of Speak Buddy provides easy navigation, supports multiple languages like English and native languages like Malayalam. The interface of SpeakBuddy is designed to be straightforward and clear, it's like having a roadmap that guides you seamlessly from one task to another, making your learning experience smooth and hassle-free.

Lesson Modules: Speak Buddy offers diverse learning modules covering vocabulary, grammar, pronunciation, and speaking practice. Each module includes interactive exercises

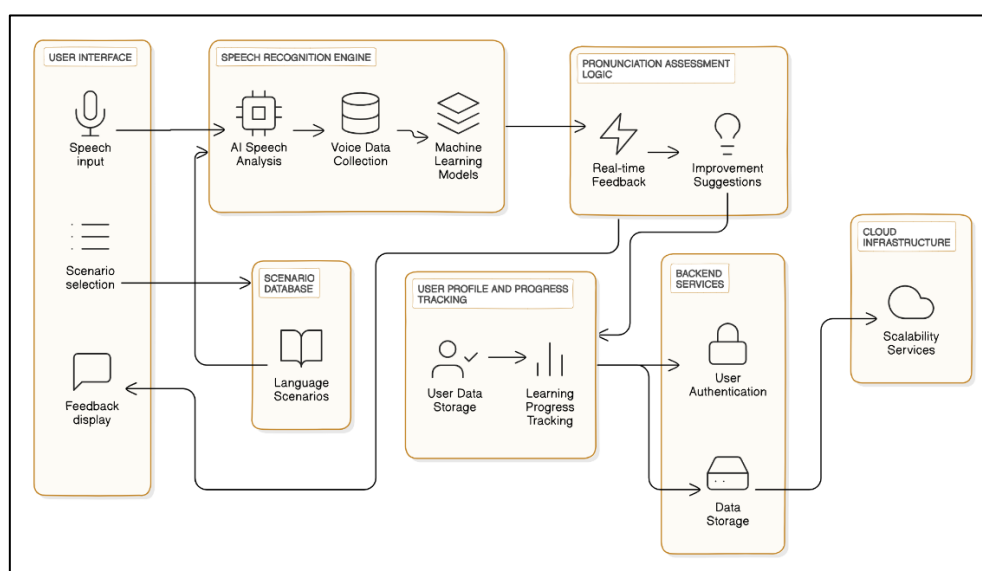


Figure 1. Architecture Diagram

and adaptive learning algorithms to cater to individual learning needs. Support for native languages like Malayalam helps clients in better understanding. **Interactive Conversation Practice:** Clients can engage in pre-trained conversation scenarios which help the clients to practice speaking and listening skills. **Native language support** facilitates meaningful interactions among clients from diverse backgrounds. **Pronunciation Training:** SpeakBuddy utilizes Azure speech recognition technology to assess pronunciation accuracy and provide real-time feedback. **AI-Powered Voicebot:** The AI-powered voicebot serves as a virtual language assistant, offering conversational practice and feedback on language usage. Clients engage in interactive conversations to practice speaking and listening skills naturally.

Database: SpeakBuddy utilizes a dual-database system, employing SQLite for storing lessons and learning content, while Firebase manages user authentication, profile details, and progress tracking. SQLite efficiently organizes lessons, vocabulary lists, and multimedia resources, facilitating fast access and ensuring data integrity. Meanwhile, Firebase handles user registration, login, and profile management, enabling personalized learning experiences and real-time progress tracking across devices

Frontend: SpeakBuddy's frontend components include Flutter framework for cross-platform development, Dart programming language, Material Design for UI/UX consistency, and interactive widgets for user interaction. Some of the actions supported by the Frontend are user registration and login, navigation through app features, selection of conversation topics, lesson categories and game options, recording and playback of user's spoken responses, display of lesson content, pronunciation assessments, and game interfaces, interaction with AI-powered voice assistant, and access to user profile settings and logout functionality.

Backend: SpeakBuddy's backend utilizes Firebase Authentication for user management, Dio HTTP client for API interactions, Azure Speech Recognition service for pronunciation training, and SQLite database for lesson storage and content management. Actions supported by the Backend are authentication of user credentials during login and registration processes, retrieval and storage of user progress data, including lesson completion status and game scores, integration with Azure Speech Recognition service for real-time pronunciation assessment, retrieval of lesson content and user-specific data from the SQLite database. SQLite is employed as the database management tool for organizing lessons, vocabulary lists, and multimedia resources efficiently, ensuring fast access and data integrity. Firebase complements this by managing user authentication, profile details, and progress tracking in real-time.

3.1 Development and Testing

Step 1: Frontend Development: Using the Flutter framework and Dart programming language, the frontend of the SpeakBuddy app was developed. This involved implementing the user interface design, interactive widgets, and navigation features. The frontend components were designed to provide seamless access to various features, including interactive

conversation practice, comprehensive lessons, engaging games, and AI-powered voice assistance.

Step 2: Backend Development: The backend of the Speak Buddy app was developed using Firebase Authentication for user authentication and management. Dio HTTP client was utilized for making HTTP requests to backend APIs, while Azure speech recognition service was integrated for pronunciation assessment and speech-to-text conversion. SQLite database was employed for storing lesson content, user progress data, and app settings locally on the device.

Step 3: Implementation of Features: Each feature of the SpeakBuddy app was implemented according to its specific requirements. Interactive conversation practice was implemented using pre-trained conversation scenarios, providing real-time feedback on pronunciation, fluency, and grammar. We created lessons in SpeakBuddy to cover everything: words, grammar, and vocabulary. We made sure they're fun and simple to grasp. Engaging games were designed to make learning enjoyable. AI-powered voice assistance was incorporated to offer personalized guidance, tips, and suggestions.

Step 4: Testing and Quality Assurance: We tested Speak Buddy a lot to make sure it does what it's supposed to do and is easy for you to use. We fixed all the problems we found and made changes based on what users told us about how they used it. We wanted to make sure SpeakBuddy helps you learn English in the best way possible.

3.2 Use Case Diagram

The SpeakBuddy app's use case diagram illustrates how users interact with the system's functionalities, including accessing learning modules, participating in conversation practice, receiving pronunciation training, and engaging with the AI-powered voicebot. Figure 2 shows the use case diagram for the SpeakBuddy application.

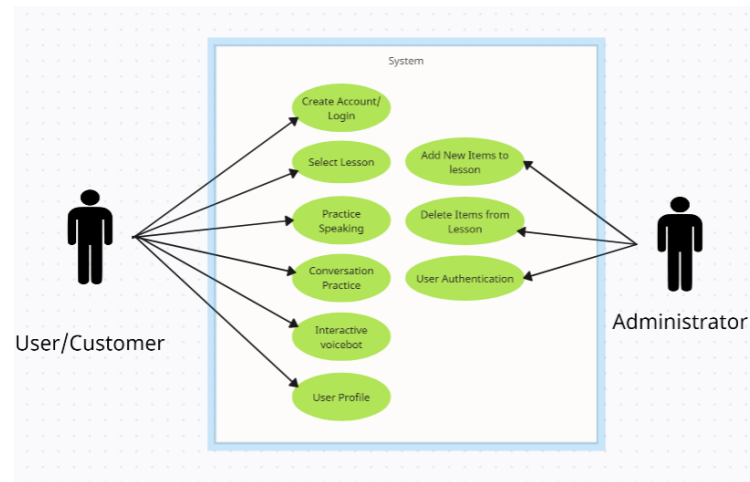


Figure 2. Use Case Diagram

4. Results and Discussion

Speak Buddy is a cool new way to learn English. It helps you get better at speaking, listening, and understanding English in a fun and personalized way. With interactive features and personalized learning, it makes learning English easy and enjoyable.

Users of SpeakBuddy really like the interactive lessons because they're interesting and easy to remember. The app also adjusts to your skill level, so you always feel like you're learning something useful. Additionally, it tells you how well you're doing with your pronunciation, which helps you get better at speaking English and when you practice speaking, it feels like you're having a real conversation, which boosts your confidence. With Speak Buddy, you can track your progress and see how far you've come, which keeps you motivated to keep learning. Overall, Speak Buddy is a great tool for learning English in a fun and easy way

Scenario 1: Interactive Conversation Practice: Users select a conversation topic from the app's menu and engage in a pre-trained conversation. The app guides you through the conversation, offering real-time feedback on pronunciation, fluency, and grammar. People like this feature because they get instant feedback on how they're speaking, which helps them get better. Figure 3 shows the interactive conversation practice in SpeakBuddy app.

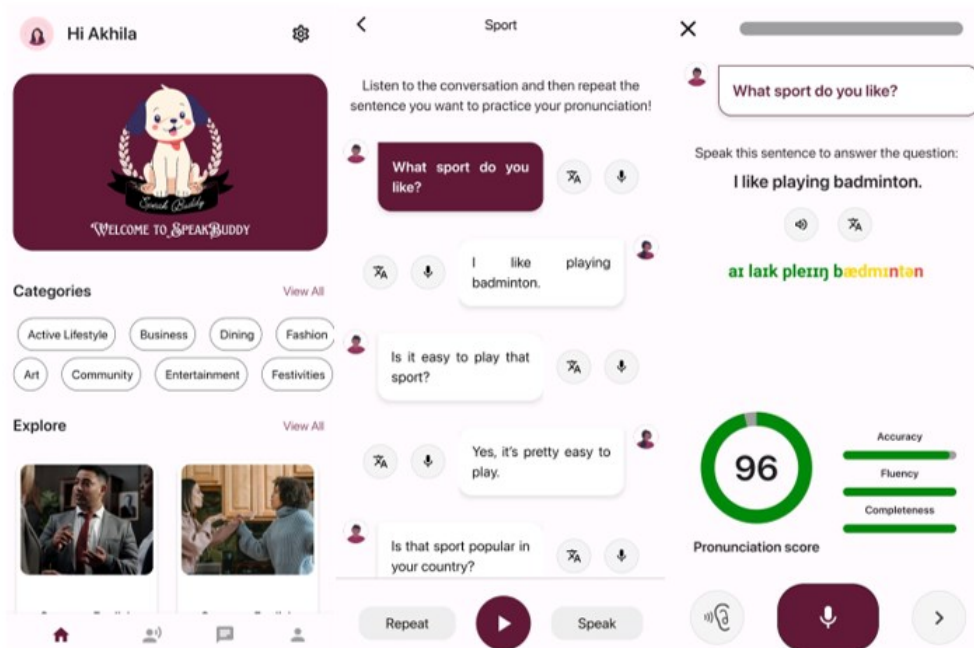


Figure 3. SpeakBuddy App

Scenario 2: Pronunciation Training- Users can access pronunciation practice part of the app and select a word or phrase to practice. The app utilizes Azure speech recognition to analyze user's pronunciation and provides feedback. Users can listen to model pronunciations and compare their own recordings. Users find this part helpful for improving how they say words because they get feedback on what they're doing wrong and can practice as much as they need. Figure 3 shows pronunciation assessment in Speak Buddy app.

Scenario 3: Comprehensive Lessons- Users navigate to the "Lessons" section and select a topic such as "Phrasal Verbs." Lessons are organized neatly with clear explanations, examples, and interactive exercises. Users can track their progress and revisit previous lessons. People like how the lessons are set up because they're easy to follow and the interactive parts help them understand and remember better. Figure 4 shows the comprehensive lessons in Speak Buddy app.

Scenario 4: AI-Powered Voice Assistance- Users interact with the AI-powered voice assistant to ask questions or request language tips. The voice assistant responds intelligently to user queries, providing personalized guidance and suggestions based on individual learning goals. Users find the voice assistant

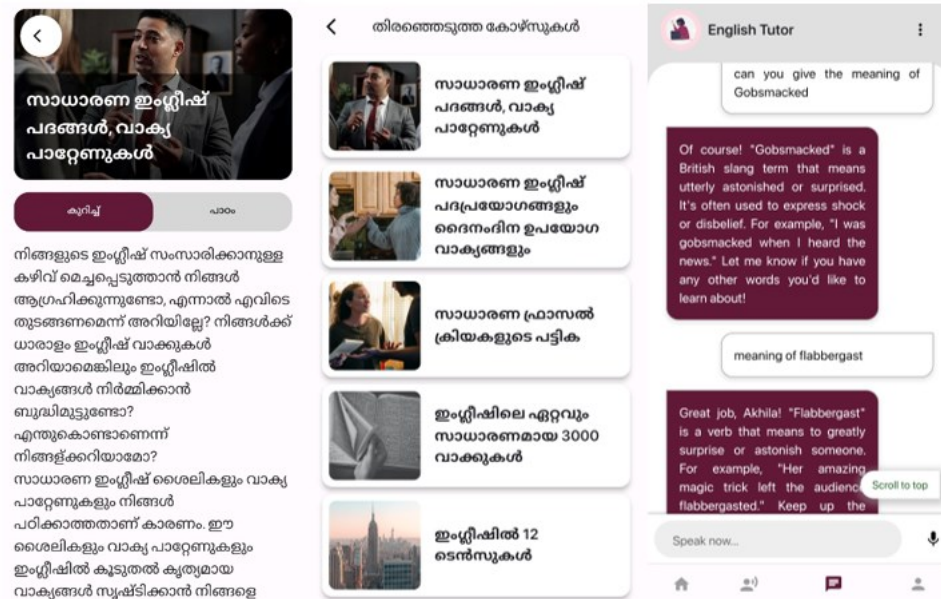


Figure 4. SpeakBuddy App- AI Powered Voice Assistant

helpful for clarifying language concepts and receiving additional support. Figure 4 shows the AI- powered voice assistant in Speak Buddy app.

In comparison to other apps, SpeakBuddy stands out for its user-friendly interface, personalized learning approach, and advanced speech recognition technology. The software offers a seamless and engaging learning experience, with interactive lessons and real-time feedback.

5. Conclusion

The proposed system, SpeakBuddy, offers a comprehensive English talking app. It has interactive features, personalized learning pathways, and uses cutting-edge technology. Advantages of Speak Buddy include an interactive user interface, diverse learning modules covering vocabulary, grammar, pronunciation, and conversational skills, as well as personalized learning pathways tailored to individual proficiency levels. The system also uses azure speech recognition technology to provide real-time feedback on pronunciation and offers virtual conversational practice scenarios, enhancing users' confidence in English communication. Additionally, SpeakBuddy incorporates progress tracking mechanisms to monitor users' learning achievements and provide insights for continuous improvement.

However, there are some limitations. One potential disadvantage is the dependency on internet connectivity for certain features like Azure speech recognition, which may hinder accessibility in areas with limited internet access. Future work could focus on optimizing the app for offline functionality or exploring alternative solutions to mitigate this limitation. Additionally, while Speak Buddy offers a dual-database architecture utilizing SQLite for lesson and content storage and Firebase for user Authentication and progress tracking, further enhancements could be made to streamline data management and improve scalability.

Overall, SpeakBuddy represents a significant advancement in language education technology, offering users a powerful tool for developing their language skills with confidence and proficiency. By continuously making SpeakBuddy better and coming up with new ideas, it can become even more popular for helping people learn languages.

References

- [1] Rukiati, Enik & Wicaksono, Julien & Taufan, Gullit & Suharsono, Degita. (2023). AI on Learning English: Application, Benefit, and Threat. *Journal of Language, Communication, and Tourism*. 1. 32-40. 10.25047/jlct.v1i2.3967.
- [2] Vadivel, Balachandran, Ahmad Abdulkareem Shaban, Zeravan Ayoub Ahmed, and B. Saravanan. "Unlocking English Proficiency: Assessing the Influence of AI-Powered Language Learning Apps on Young Learners' Language Acquisition." (2023).
- [3] Mukhallafi, T. R. A. "Using artificial intelligence for developing English language teaching/learning: an analytical study from university students' perspective." *International Journal of English Linguistics* 10, no. 6 (2020): 40.
- [4] Abimanto, Dhanan & Sumarsono, Wasi. (2024). Improving English Pronunciation with AI Speech-Recognition Technology. *Acitya: Journal of Teaching and Education*. 6. 146-156. 10.30650/ajte.v6i1.3810.
- [5] Umezawa, Katsuyuki, Makoto Nakazawa, Michiko Nakano, and Shigeichi Hirasawa. "Utilizing Automatic Speech Recognition for English Pronunciation Practice and Analyzing its Impact." In *2023 IEEE 12th International Conference on Engineering Education (ICEED)*, pp. 69-74. IEEE, 2023.

- [6] Belda-Medina, Jose, and José Ramón Calvo-Ferrer. "Using chatbots as AI conversational partners in language learning." *Applied Sciences* 12, no. 17 (2022): 8427.
- [7] Kukulska-Hulme, Agnes, and Lesley Shield. "An overview of mobile assisted language learning: From content delivery to supported collaboration and interaction." *ReCALL* 20, no. 3 (2008): 271-289.
- [8] Amin, Eman. (2024). EFL Students' Perception of Using AI Text-to-Speech Apps in Learning Pronunciation. *Migration Letters*. 21. 887-903. 10.2139/ssrn.4746800.
- [9] Zhang, S., Shan, C., Lee, J.S.Y. et al. Effect of chatbot-assisted vernacular acquisition: A meta-analysis. *Educ Inf Technol* 28, 15223–15243 (2023).
- [10] Uchihara, Takumi, and Kazuya Saito. "Exploring the relationship between productive vocabulary knowledge and second language oral ability." *The language learning Journal* 47, no. 1 (2019): 64-75.
- [11] Noga, Tomasz. (2023). The use of chatbots and voicebots by public institutions in the communication process with clients. *Scientific Journal of Silesian University of Technology Series Transport*. 2023. 69-79.
- [12] Kaledio, Potter, Abill Robert, and Louis Frank. "The Impact of Artificial Intelligence on Students' Learning Experience." Available at SSRN 4716747 (2024).
- [13] Ngo, Hui Kiang, and Md Yunus. "Flipped classroom in English language teaching and learning: A systematic literature review." *International Journal of Academic Research in Business and Social Sciences* 11, no. 3 (2021): 185-196.
- [14] Eksi, Gonca Yangin, and Sabahattin Yesilcinar. "An Investigation of the Effectiveness of Online Text-to-Speech Tools in Improving EFL Teacher Trainees' Pronunciation." *English language teaching* 9, no. 2 (2016): 205-214.
- [15] Lee, Yunhyun. (2021). Effectiveness of Mobile Assisted Pronunciation Training in the Acquisition of English Vowels, /o/ and /ɔ:/. *Foreign Languages Education*. 28. 53-76.
- [16] Chennuri, Varun & Rodda, Vamshi Prashanth. (2023). Development of AI-based voice assistants using Large Language Models. 10.13140/RG.2.2.20195.12321.

Author's biography



Akhila is currently enrolled in a Bachelor of Technology program majoring in Computer Science and Engineering at APJ Abdul Kalam Technological University, Kerala, India. With a year dedicated to this research, her primary focus revolves around the integration of Artificial Intelligence and Flutter development.



Anagha P Nair is currently enrolled in a Bachelor of Technology program majoring in Computer Science and Engineering at APJ Abdul Kalam Technological University, Kerala, India. With a year dedicated to this research, her primary focus revolves around the integration of Artificial Intelligence and machine learning.



Jyothsana K N is currently enrolled in a Bachelor of Technology program majoring in Computer Science and Engineering at APJ Abdul Kalam Technological University, Kerala, India. With a year dedicated to this research, her primary focus revolves around the integration of machine learning, deep learning and AI.



Sreepriya S is currently enrolled in a Bachelor of Technology program majoring in Computer Science and Engineering at APJ Abdul Kalam Technological University, Kerala, India. With a year dedicated to this research, her primary focus revolves around the integration of machine learning and artificial intelligence.



Akhila E is an Assistant Professor at JCET, Kerala, India. With a year dedicated to this research, her primary focus revolves around the integration of Artificial Intelligence and machine learning.