A MOBILE APPLICATION TO IMPROVE ONLINE LEARNABILITY TO VISUALLY IMPAIRED ELEMENTARY SCHOOL CHILDREN

Project Id: TMP-23-310

Project Proposal Report

Sathsarani.B.A.D.A

BSc (Hons) in Information Technology Specializing in Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology Sri Lanka

March 19th

A MOBILE APPLICATION TO IMPROVE ONLINE LEARNABILITY TO VISUALLY IMPAIRED ELEMENTARY SCHOOL CHILDREN

Project Id: TMP-23-310

Project Proposal Report

BSc (Hons) in Information Technology Specializing in Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology Sri Lanka

March 19th

Declaration of the Candidate and the Supervisor

I declare that this is my own work, and this dissertation does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any other university or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgment is made in the text.

Name	Student ID	Signature
Sathsarani B.A.D.A	IT20160098	

The supervisor/s should certify the proposal report with the following declaration.
The above candidates are carrying out research for the undergraduate Dissertation under my supervision.

02/05/2023

…….. ……

Signature of the supervisor: Date:

Abstract

Blindness is a condition in which a person loses the capacity to see fully or partially. It is also reported to be one of the most common disabilities worldwide. Most of the time, finding a complete cure for this ailment is unattainable. As a result, people who are blind adjust to living with their blindness as a part of them. This work intends to construct a chatbot that uses artificial intelligence to predict the most relevant answers to the users' requests in order to assist visually impaired persons. This bot is constructed in such a way that it allows the user to give voice notes rather than textual, as do traditional chatbots. The chatbot is built using a specific architecture. This architecture converts user-sent text to text first. The core architecture of the chatbot is comprised of deep neural networks and APIs used for speech-to-text and text-to-speech conversion. This text is subsequently utilized to identify the best response for the user, which is ultimately delivered to the user. The inner workings of the architecture include an AI technique known as deep neural networks, speech-to-text and text-to-speech APIs and etc.As part of my contribution, I will create a supportive voice bot to assist visually impaired students with their studies. I plan to create a voice bot in three languages: Sinhala, English, and Tamil, as well as subject-based questions and answers. Further, I will create an accessibility mode and incorporate it into the user interface design. A supportive voice bot can help reduce stress and anxiety for visually impaired students. Because they may be overwhelmed by the amount of information they must process, the bot can guide them through the process with a calm and supportive voice.

Keywords— Artificial Intelligence, Chatbot, Preprocessing, Deep Neural Network, Speech Recognition

Table of Contents

List of figures

List of Tables

1. Introduction

1.1 Background Literature

IBM pioneered digital voice recognition technology for computers in the 1990s. This was the first stage of this groundbreaking technology, and it lay the framework for contemporary voice command recognition. Virtual house assistance has evolved greatly for the first time since Apple debuted Siri in 2011. Companies such as Google and Microsoft later entered the voice technology business. Google debuted Google Now in 2012, and Microsoft introduced a new tool called "Cortana" at the "annual Build worldwide developer's conference" in 2013. Later in 2014, Amazon released their innovation, Amazon Alexa, setting the path for a revolution in smart speakers.[5].

In 2015, several technological developments were made, including the debut of the Amazon Echo and the Amazon Alexa Skills Kit in the United States, as well as Microsoft's integration of Windows on desktops and smartphones. Google opted to enter the voice technology market in 2016, unveiling the Google Home in May of that year as a feature of the messaging app Allo. The race to give the finest voice support innovation to the broader public begins when Samsung acquires virtual assistant startup viv the same year. A voice assistant (NLP) is built on technologies such as voice recognition, artificial intelligence (AI), and natural language processing. The combination allows the technology to comprehend human inquiries, respond to them, and carry out tasks as requested.

A study [5] by interpreting the users' orders and acting appropriately, the Zira Voice Assistant helps us better comprehend what it does. Furthermore, the Assistant can perform the most frequently requested tasks, making our lives easier. After receiving voice input, the assistant says "OKAY," and it prints down the user's message and takes appropriate action. Zira may speak normally as part of the PythonTTSX3 implementation. Defined the "Speak" function, which allows the system to speak like a human. A voice input method "speech_recognition.Microphone()" is used in the Speech Recognition module to collect voice input, and a "pyttsx3" function is used to convert text to speech. Voice_recognition is being used.Recognizer.recognize_google() Using English as Zira's primary language.

A study [1] use a text-to-speech algorithm capable of transforming speech to a Japanese accent. It is created and trained a text-to-speech model to transform Japanese text to speech. The model is eventually trained to read the text aloud with an accent. This was accomplished through the use of a set of features known as the BERT. The model is also taught to change the tone of the audio based on the emotion expressed in the text.

Another study [2] intends to develop an end-to-end text-to-voice translation algorithm capable of translating speech to a natural tone. According to the study, the cut copy action frequently affects the effectiveness of the algorithm and is removed by the use of a method known as context-aware mask prediction. When compared to other text-to-speech conversion software, this model can produce better results.

In the proposed solution, will use the same strategy as voice assistants in existing research, but there will be some differences. The research gap between the preceding studies and ours is that no one has developed a helpful voice bot in multilingual languages (English, Sinhala, and Tamil) for visually impaired pupils.

1.2 Research Gap

As discussed above, many current research and systems have been presented but it is still not possible to present a complete system with all the proposed features. But this proposed system is more effective and comprehensive.

In existing studies, there is no supportive voice bot to assist visually impaired students. In the proposed solution, we will create a voice bot in a mobile app, and there is already a desktop-based voice assistant. Furthermore, it only uses English, but in the proposed solution, we will create a supportive voice bot in three languages: English, Sinhala, and Tamil. This table shows summary comparison the features of the proposed system and the existing system approach.

Research	Supportive voice bot for multilingual languages	Accessibility mode in the UI design.	Voice input and response	Mobile application
Research 1 [2]	✗	✗	✓	✓
Research 2 [5]	✗	✗	✓	✗
Research 3 [4]	✗	✗	✓	✓
New Research	✓	✓	✓	✓

Table 1

2. Research Problem

Everyday problems for blind persons range from reading a book to walking down the street. Although many tools are available to help them meet the challenges they face, they are insufficient. Vision is the most important thing a human can have, and it plays a vital role in a person's life whether they can see or not. Even simple daily tasks require the assistance of a vision impaired person. In this research, we explored the difficulties that blind people encounter in their daily lives and attempted to give a satisfying answer for them.

Students with visual impairments tend to be at a disadvantage when compared to students with other disabilities. It promotes the use of assistive technologies to enhance learning by decreasing learning complexity. Although these assistive technologies enable disabled people to access e-learning content, they have yet to be proven to be completely successful for everyone. Because of their impairment, visually impaired pupils confront several barriers to schooling. In typical classroom settings, students with visual impairments frequently rely on the assistance of teachers and peers to obtain course information. However, with the increasing adoption of technology in education, there is an opportunity to provide visually impaired students with new forms of support, such as voice bots.

As a result, we cannot be considered disability-aware e-learning systems. Individuals with disabilities will demand individualized information in specific formats; thus, assistive and adaptive technology must be built to give universal knowledge access. Voice bots are computer programs that interact with users via spoken language using artificial intelligence. They can be programmed to reply to user inputs, answer questions, and offer assistance on a variety of topics. Voice bots have the ability to provide a friendly and inclusive learning environment for visually impaired pupils.

3. Research Objectives

In this section, the objectives of the research are divided into main and sub. The objectives are separately listed below.

3.1 Main Objective

The main objective of a supporting voice bot for visually impaired students is to aid and support them in their educational pursuits. A voice bot of this type can be built to perform a variety of functions, including answering questions, providing explanations, helping students through assignments, and providing general academic support. Further need to create a supportive voice bot which facilitates learning in multilingual language.

3.2 Sub Objective

- To create a supportive voice bot which facilitates learning in multilingual language.
- To facilitate the accessibility mode for improving vision in the system.
- To providing relevant information and feedback on specific topics.

4. Methodology

4.1 Methodology including the system diagram

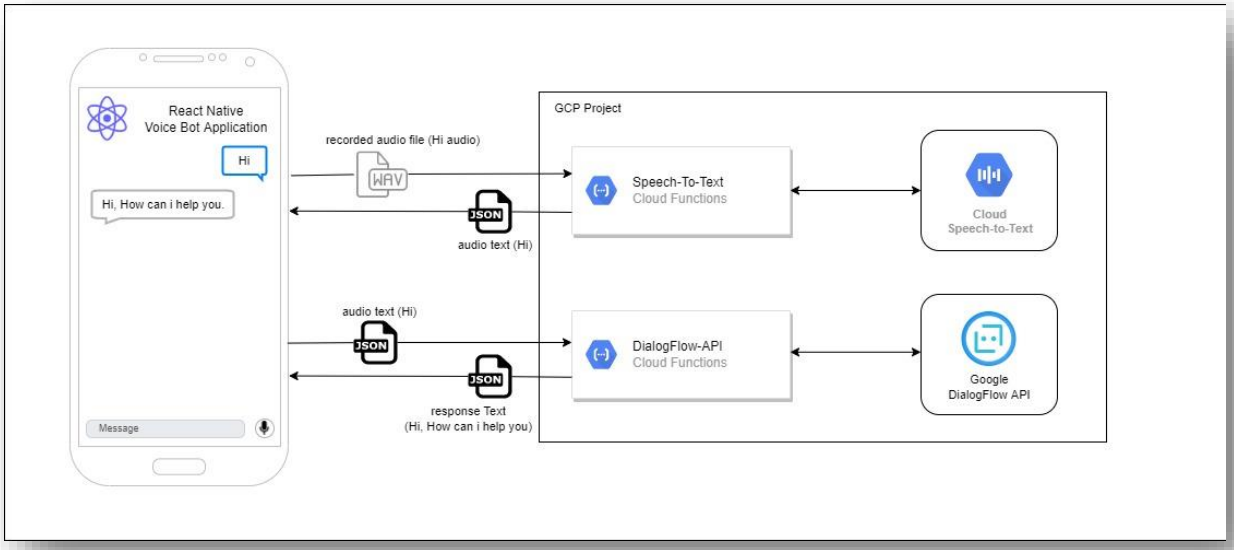

Figure 1

First, create a Google Cloud Platform (GCP) account and project. In project, enable the Cloud Speech-to-Text and Dialogflow APIs. Then create a new Dialogflow agent and establish its parameters, intents, and entities. in addition create an intent for the user's requests or queries, and then populate it with training phrases and responses.Create a Cloud Function that takes audio files, turns them to text using the Speech-to-Text API, and then delivers the text to Dialogflow for processing. Connect the Cloud Function to the webhook fulfillment in Dialogflow. Send an audio file to the Cloud Function or use the Dialogflow simulator to test the voice bot.

4.2 Commercialization of the Product

Elementary schools, special education schools, and online learning platforms that serve visually impaired students place a demand for e-learning systems for elementary students. The primary visually impaired pupils, their teachers, and their parents or guardians should be the target audience.

The e-learning platform should be created with accessibility features like audio explanations, high Contrast Mode, text-to-Speech mode, magnified mode, and other features specific to visually impaired students' requirements.

The e-learning system should not only target primary pupils, but also their teachers and parents/guardians, who are critical to their academic growth.

The target market and audience for an e-learning system for visually impaired elementary pupils would be the visually impaired society as a whole, including primary schools, visually impaired children, educators, parents, guardians, and organizations that serve the visually impaired population.

Commercialization also involves generating revenue from the mobile application. This could involve charging for the application itself, offering premium features for a fee, or monetizing through in-app advertising. Further, use social media to promote.

5. Software Specifications

5.1 Software Solution

Agile project management is an iterative strategy that stresses flexibility, collaboration, and client satisfaction.It was created for software development projects at first, but it has since been adopted by many other industries.

Agile management emphasizes frequent communication, rapid prototyping, and continuous feedback to enable teams to quickly adapt to changing requirements while delivering high-quality products or services. Working in short iterations, known as sprints, and focusing on delivering small, working increments of a project are typical.

The Agile Manifesto outlines the agile management values and principles, which include:

1. Individuals and interactions over processes and tools
2. Working software over comprehensive documentation
3. Customer collaboration over contract negotiation
4. Responding to change over following a plan.

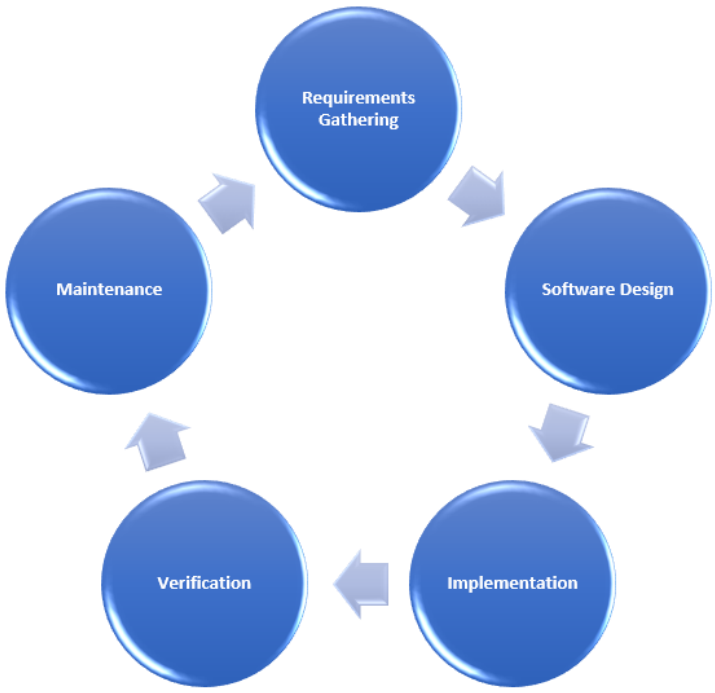

Figure 2

5.2 Tool and Technology

Tools

Dialog Flow

Figure 3

Dialogflow is an NLP (Natural Language Processing) platform that is used to create an application for the company's customers in numerous languages and across multiple platforms. The majority of Google Assistant devices use Dialogflow to build actions.

VS Code

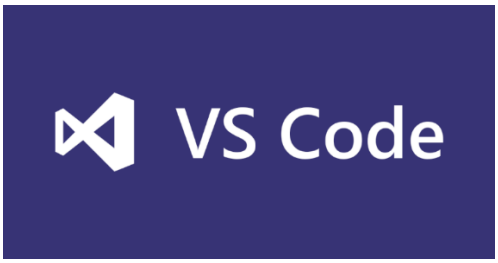

Figure 4

Visual Studio Code, usually known as VS Code, is a source-code editor for Windows, Linux, and macOS that was created by Microsoft utilizing the Electron Framework. Among the features are debugging, syntax highlighting, intelligent code completion, snippets, code refactoring, and embedded Git.

Technology

React Native

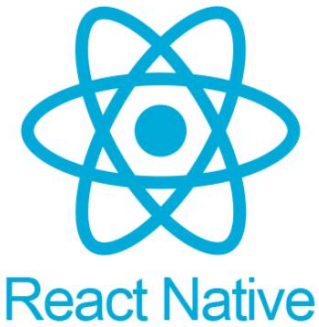

Figure 5

Facebook built React Native, a prominent open-source mobile application development framework. It enables developers to create cross-platform mobile applications for iOS, Android, and the web from a single codebase. React Native is based on the ReactJS library, which is used to create online applications.

Natural Language processing

Figure 6

NLP is used to analyze text and train machines how to speak like humans. Automatic text summarization, sentiment analysis, subject extraction, named entity recognition, parts-of-speech tagging, relationship extraction, stemming, and more real-world applications are available.

5.3 Functional Requirements

- Multilingual support: The voice bot should be able to support multiple languages, to cater to students from diverse backgrounds. So, need to develop a voice bot into 3 languages: Sinhala, English and Tamil.
- Curriculum integration: The voice bot should be able to integrate with the curriculum, providing relevant information and feedback on specific topics.
- Accessibility mode: Add accessibility mode for mobile application, such as High Contrast Mode, Text-to-Speech mode, magnified mode.

5.4 Non -Functional Requirement

1.1.1. Usability

- The application should be able to achieve user's required goals efficiently and effectively.
- The application should have a user-friendly interface that is easy to navigate and understand.

1.1.2. Reliability

- The application output accuracy should be error free, and the code should be bug free.so the translation output could be reliable.
- The voice bot should be reliable and available 24/7 to assist students.

1.1.3. Performance

- The application should be able to respond to user requests quickly and efficiently, without lag or delays.

1.1.4. Security and privacy

- The administrative application should be protected from unauthorized access.
- The database should be protected from attacks and unauthorized access.
- The interface should be protected from attacks.
- All passwords should be stored as a secure hash of the administrator password.

1.1.5. Scalability

- The application should be able to handle a large number of users and requests without crashing or experiencing downtime.

1.1.6. Maintainability

- The application should be designed in a way that makes it easy to maintain, update, and fix bugs over time.

1.1.7. Speed

- The voice bot should respond quickly to student requests to provide timely assistance.

1.1.8. **Portability**

- The application should be able to run on multiple platforms and devices, without requiring major modifications or changes.

5. Conclusions

A supportive voice bot for visually impaired students can be a valuable tool for facilitating their education and improving their overall learning experience. Voice bots can give visually impaired students with a more accessible and interactive approach to learn and interact with course material by utilizing advanced technologies such as natural language processing and speech recognition. Voice bots, which can reply to voice commands and questions, can also serve as virtual assistants, guiding students through their courses and keeping them organized and on schedule. The system's supportive voice bot, which supports Sinhala, English, and Tamil, completes the four components. This function makes sure that students may communicate with the program in their own language, which makes it easier for them to follow and study.

Furthermore, supportive voice bots can be especially beneficial for visually impaired students who face difficulties in traditional classroom settings. Voice bots can help people gain independence, confidence, and self-sufficiency by minimizing their reliance on others. Furthermore, they can help bridge the educational accessibility gap for visually impaired students, allowing them to reach their full potential.

In conclusion, a helpful voice bot for visually impaired students can be a very useful tool in improving their education and enhancing their learning experience. Voice bots are poised to become an increasingly important component of education for visually impaired individuals as technology advances.

References

[1] Y. Y. a. T. Toda, "Investigation of Japanese PnG BERT language model in text-to-speech synthesis for pitch accent language," *in IEEE Journal of Selected Topics in Signal Processing,* 2022.

[2] J. Y. R. F. J. T. a. Z. W. T. Wang, "CampNet: Context -Aware Mask Prediction for End-to-End Text-Based Speech Editing," *in IEEE/ACM Transactions on Audio, Speech, and Language Processing,* 2022.

[3] "Customer Service Automation: Better CX thanks to AI Voice Bots. (2021).," Retrieved 18 October 2021,. [Online]. Available: https://talkie.ai/voice-ai-for-customer-service-automation/.

[4] https://www.researchgate.net/publication/322418348_Next-Generation_of_Virtual_Personal_Assistants_Microsoft_Cortana_Apple_Siri_Amazon_Alexa_and_Google_Home, (2021). Retrieved 5 October 2021, from .

[5] S. V. A. Abhishek Joshi, PERSONALIZED DESKTOP APP BASED INTERACTIVE, Uttaranchal University, Dehradun-248007Uttarakhand, India, 2022.

[6] "3. Bhalla, K., Koundal, D., Sharma, B., Hu, Y.C. and Zaguia, A., 2022. A fuzzy convolutional neural network for enhancing multi-focus image fusion. Journal of Visual Communication and Image Representation, 84, p.103485".

[7] "4. Ramya, G., Priya, G.M., Subrata, C., Kim, D., Tran, D.T., & Le, A.N. (2021). A Review on Various Applications of Reputation Based Trust Management. International Journal of Interactive Mobile Technologies (ijim), 15, 87-102".

6. APPENDICES

6.1 Work Breakdown Structure

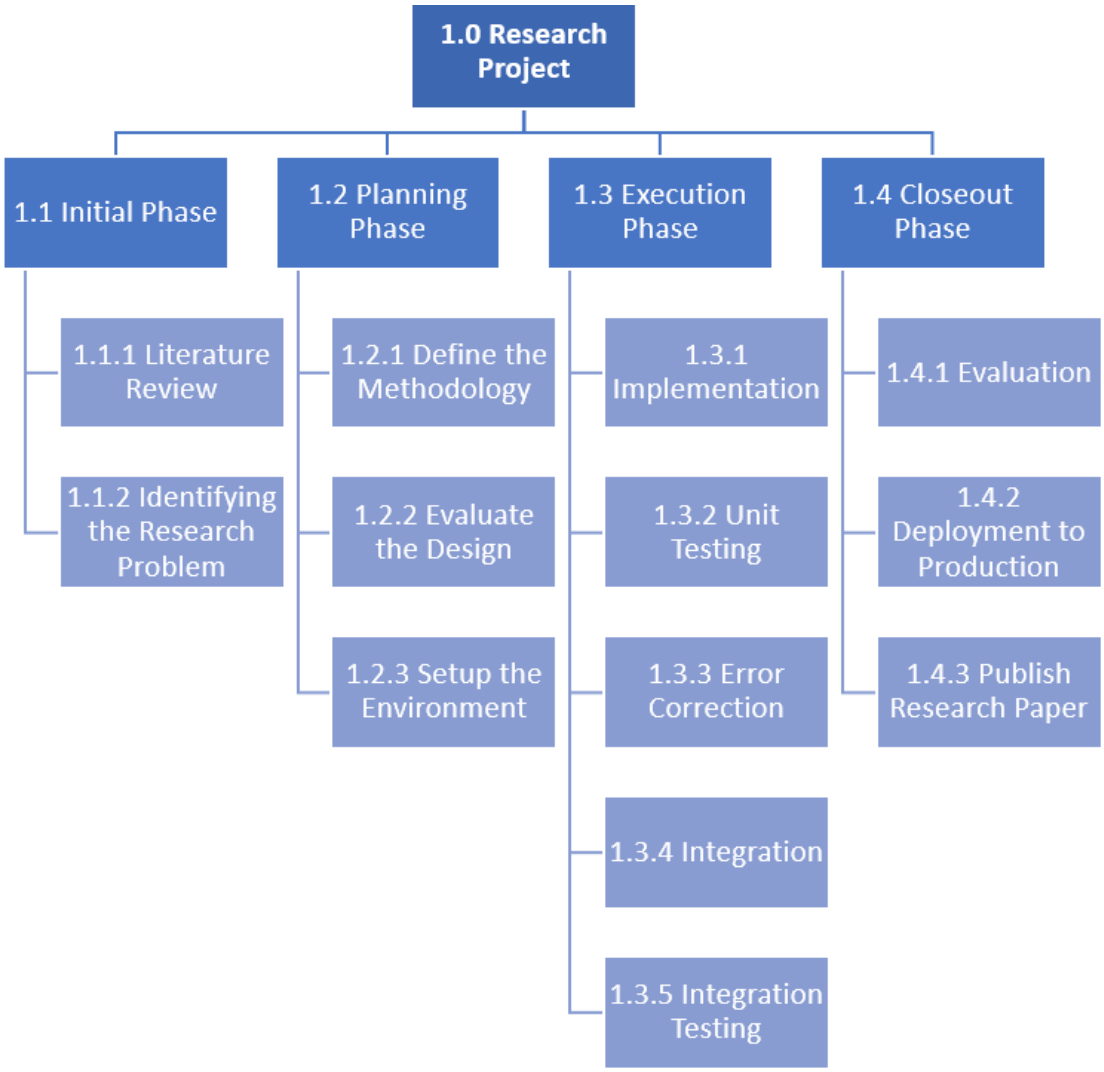

Figure 7

6.2 Gantt Chart

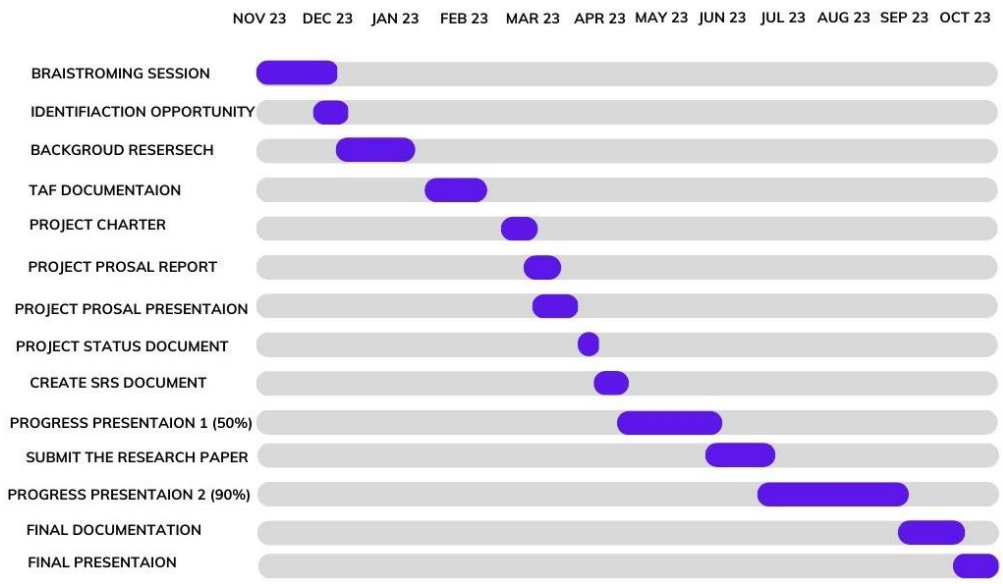

Figure 8

6.3 System Overview

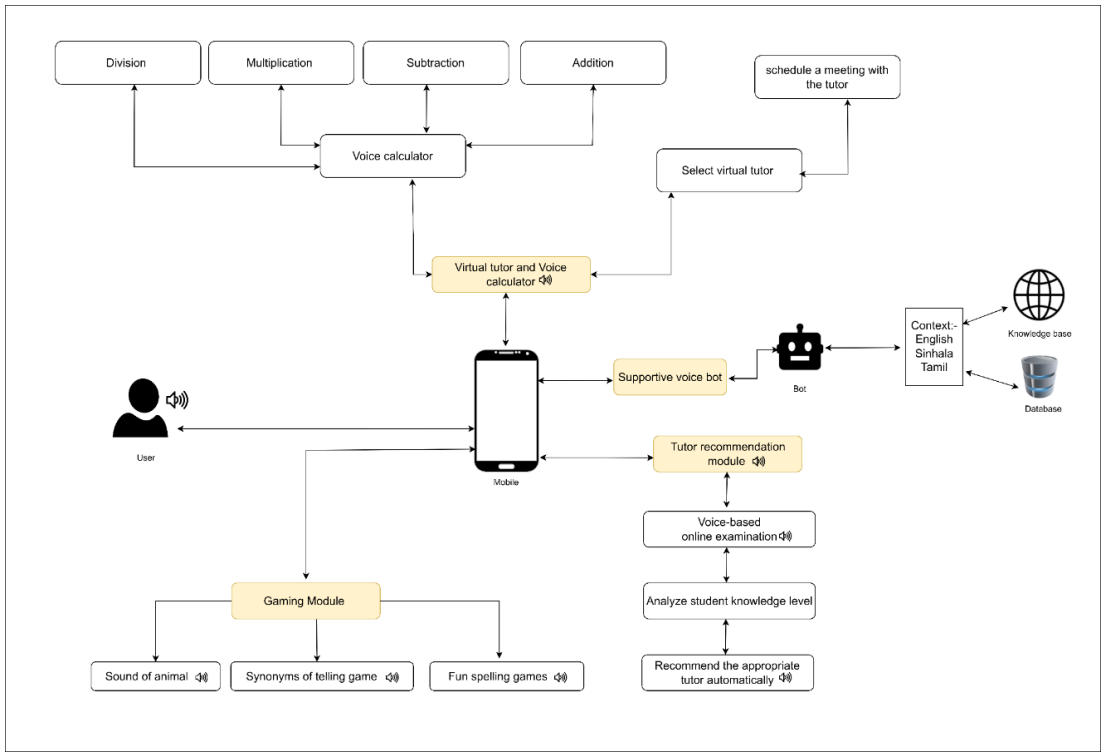

Figure 9

6.4 Blind school images

We went to a Ratmalana blind school to interview instructors and children and learn more about them, such as the challenges they face, the features they are expected to include in mobile application, and so on.

Figure 10


අන්ධ විද්‍යාලය
රත්මලාන
THE SCHOOL FOR THE BLIND
RATMALANA
விழிப்புலனற்றோர் பாடசாலை

Figure 11