Public Speaking Trainer System Using Deep Learning (PST)

by

Arwa Abdullah Aljamea (218022404)

Azizah Ibrahim Habiballah (218012041)

Buthena Ibrahim ALMesthi (218009566)

Nada Ahmad ALEssi (218010977)

Supervised by

Dr. Mona Ali

Committee Member Names Dr.

Hafiz Farouq

October, 2021

ACKNOWLEDGEMENT

UNDERTAKING

This is to declare that the project entitled "Public Speaking Trainer" is an original work done by undersigned, in partial fulfilment of the requirements for the degree "Bachelor of Science in Computer Science" at Computer Science Department, College of Computer Sciences and Information Technology, King Faisal University.

All the analysis, design and system development have been accomplished by the undersigned. Moreover, this project has not been submitted to any other college or university.

Student 1: Arwa Abdullah Aljamea

Student 2: Azizah Ibrahim Habiballah

Student 3: Buthena Ibrahim ALMesthi

Student 4: Nada Ahmad ALEssi

ABSTRACT

In general, the speech process is the first communication language between people in order to understand each other ideas. The presentation helps visual people to understand very well. On other hand, the influence of the presenter's speech helps auditory people to understand the idea of the presenters more. As a result, we would like to create Public Speaking Trainer (PST) web application that satisfies both visual and auditory people. PST web application will help the presenters to appear in their best form by training them to perform in front of the audience. Moreover, PST web application is not limiting in normal people only, we are considering disabled people such as blind people. The aim of our project is to help the presenters to make time management and increase the self-confidence. The system of our application will use Python language and our application will support signal processing techniques while processing the voice signal. In order to let our application more visible and flexible to the presenters, we decided to make it as web application. Eventually, PST web application will support any presenters to get out with the best outcomes.

In this milestone, we will show more details about the proposed PST web application is working and what are the methodologies that will be used. In addition to the detailed prototype for the application. In addition, we will display the expected outcomes for PST web application.

Acronyms and Abbreviations

PST: Public Speaking Trainer

PDF: Portable Document Format

MLSI: Machine Learning for Stuttering Identification

AI: Artificial intelligence.

ML: Machine learning.

DL: Deep learning.

ANNs: Artificial Neural Networks

HMMs: Hidden Markov Model

SVM: and Support Vector Machine

Table of Contents

List of Tables

List of Figures

1 Introduction

Soft skills of body language are the way your reveals the true story behind your words. It combines hand gestures, posture, facial expressions, and movements that tell others what's going on inside your head. The general standards of the soft skills of body language for presenters are eyes contact, stand without over-movement, stand confidently, and proof of personal audacity. These are the soft skills of the presenters. However, our project will not concern about the soft skills. Our project concerned by deducting the rate of stuttering and silence mode in the presenter's voice. Since the stuttering and long silence mode are result of fear and lack of self-confidence which may lead to reduced audience understanding of the message or information that presenters want to be delivered.

This section represents a general background about importance of the speech, the objectives of our PST project, the motivation that motivated us to do this kind of project. Finally, we discuss the problem statement and challenges that will face us.

1.1 Background

Since a long time ago, people have been using their own speaking skills in order to persuade and connect other people to their own ideas. For example, the Prophets used speaking skills to deliver the message. The Prophet of God, Moses, peace be upon him, asked God to make his brother Aaron an aid to him in conveying the message, because he was more eloquent in tongue and this will help him in fulfilling the call of God. Here, the importance of speaking skills becomes clear, as the Prophets used them to reach and achieve their goals.

1.2 Objectives

The aim of this project is to reduce mistakes and technical errors. Since the speakers must prepare well for the presentation with a proper outline and structure, as well as a good vocabulary and practicing pronunciation and annotation. One must be well organized in order to make the audience understand the purpose of the talk and keep them engaged from the beginning to the end of the presentation. The good speaking skills help the audience understand the speakers' ideas better.

1.3 Motivation

The basic motivation in this project is the presenters since the presenters face some difficulties while they are presenting in front of the audience. This motivation inspired us to create this project which will help any presenter to solve speaking problems such as fear and anxiety during presentation. As a result, this will enhance their confidence to face the audience and to overcome their presentation difficulties. We saw the developer who developed similar applications of our application never considered disabled people. As a result, we would like to consider disabled people in our application by supporting some functionalities to guide them while they are using our application.

1.4 Problem Statement

While presenting a presentation the most problems that face the presenters are fear, anxiety and tension from the audience during the presentation. These problems lead to boredom of the audience and lack of focus on the speaker. Furthermore, the speaker failed to deliver the idea or goal of presenting the presentation to the audience because of a lack of self-confidence and poor time management. As a result, we are looking to develop an application that will analysis the presentation and get feedback to the presenters in order to help them to manage time and increase self-confidence. In solving this problem, we might face many challenges one of these challenges is the signal processing techniques that will be used to analyse the presenter's presentation since the university does not cover this subject. As a result, we have to make a self-learning to know more about this field.

1.5 Scope of The Project

The idea of PST is inspired from the presenters who are facing some difficulties while they are presenting in front of the audience. This inspired us to create this application which will help any presenter to solve speaking problems such as fear and anxiety during presentation. The scope of the idea is not just limited to normal people, it has evolved disabled people. Which will allow disabled people to use the application freely. The people who have myopia/hyperopia in seeing, our application can help them to change the font size. For people who cannot see (blind people) our application can read the text for them to let them know what to do.

First of all, we will create a web application. The web application will fit the general standards for any presentation. Furthermore, web applications are more flexible and visible during presentation time. This application will deduct the rate of stuttering and silence mode in the presenter's voice using signal processing technique.

For the PST web application to solve this problem, we will require a dataset of over 28K to train our model on. Based on the model we learn; we will have a solution. Different inputs are possible for trained models, such as live presentations or recorded audio files. As shown in *Figure 1*, this process will be used to use a trained model to analyse the live presentation. The dataset we have found consists of over 28K observations, which makes deep learning a more appropriate technique in this case.
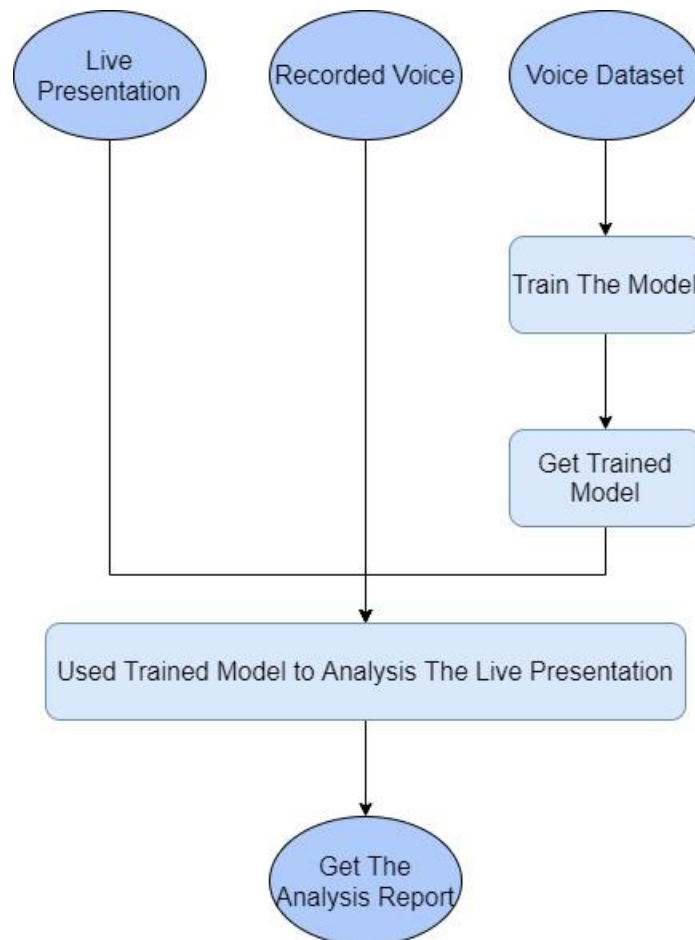
Figure 1: the flow of developing the project

The PST web application consists of two parts; the first part is before the presentation. In this part the users can upload an existing recorded voice file or record a live voice file. The second part is during presentation in this part the users can be able to record a live voice file while they are presenting. In both parts, after the presenters finished, the users will see their performance analysis as voice to hear, table of time, and chart to see the improvement of their performance as shown in *figure 2*. This process will be done by using signal processing technique which will analysis the voice file to provide the performance analysis. While the users record a live voice file, they can upload their own presentation as a PDF file. Furthermore, the record and recorded voice file will be saved as an mp3 file. Before the presentation part in recording a live voice file section the users can have a virtual audience in order to feel the reality of the experience. During the presentation part in recording a live voice file section the users can receive notifications in order to reduce the awkward silence moments of stuttering. The result of these processes is to improve the quality of the presenter's presentation.
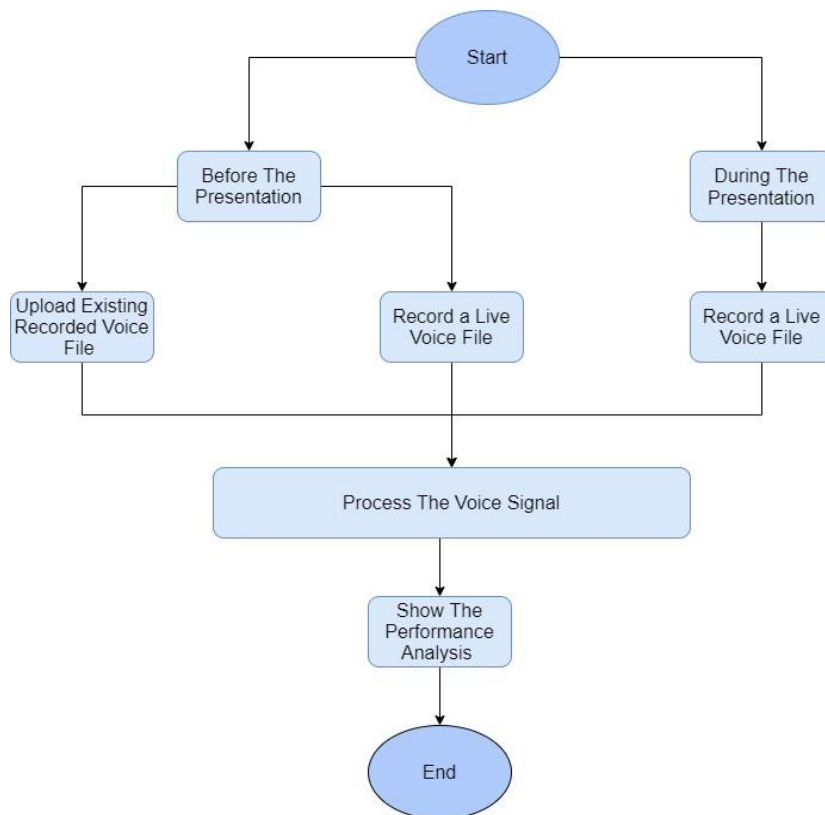
Figure 2: the flow of the project

The application will serve important stakeholders who are the presenters. In addition, it would be applicable for disabled people in which the application will support them by training them the same as normal people.

2 Analysis of related work

In this section we will study some related works that are similar to the PST web application that we want to implement. Then we will compare PST web application with the rest of the related works that have been studied based on some standards.

2.1 LikeSo

This app offers an effective way to train against verbal habits and practice speaking articulately, confidently and without all of those "likes," and "so". LikeSo provides a real time analysis of your speech fitness, including optimal pacing for the fast or slow talker. It provides the average of speech scores over 30 days, the talking time, the fillers, and pace as shown *Figure 3*. It can be installed only on the IOS platform, and it is not free to purchase. [1]
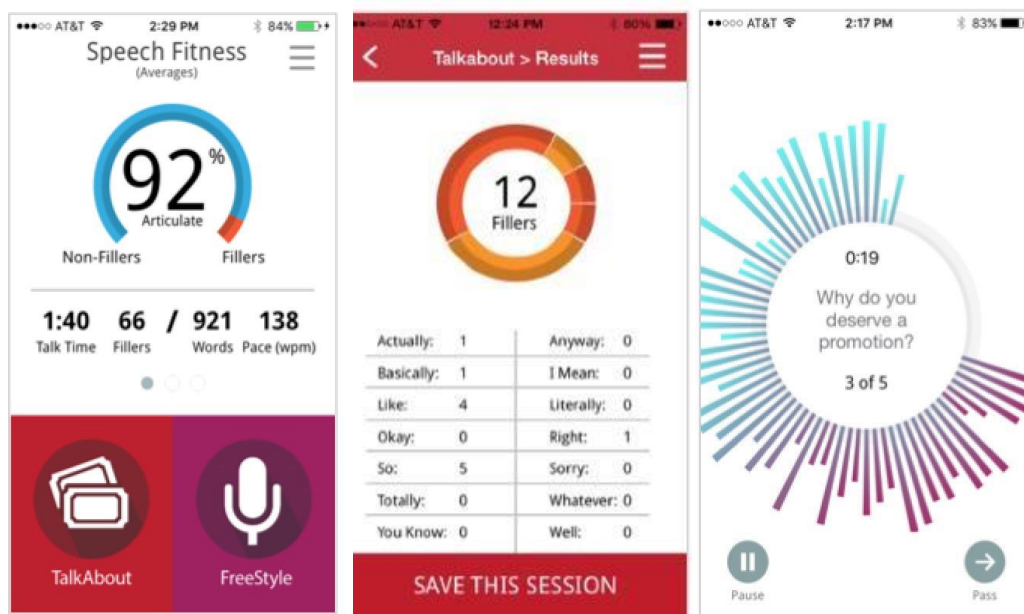
Figure 3: some of the functionalities of LikeSo application

2.2 VirtualSpeech - VR Courses

This VirtualSpeech app helps you practice your public speaking, by providing photo-realistic environments to train in using VR headset, or through the web browser. Features: realistic audience as shown in *Figure 4* and environments, sound distractions, voice analysis, upload your own presentations, interactive courses. They provide online courses and they track the user's performance as shown in *Figure 5*. The application can be installed in IOS and Android. [2]

Figure 5: realistic audience

Figure 4: track of user performance

2.3 Orai

      A mobile app works for IOS and Android users. This app provides a pretest for every user to determine the level of confidence, pace, filler world, pauses, energy. It lasted for 30 seconds Then provided the report with the insight of a specific point of the record. Then prompt for a self-evaluation for the previous categories. After subscribing in the app, the user could get access to five different courses, many blogs, and a customized training plan to improve his/her speaking skills. [3]
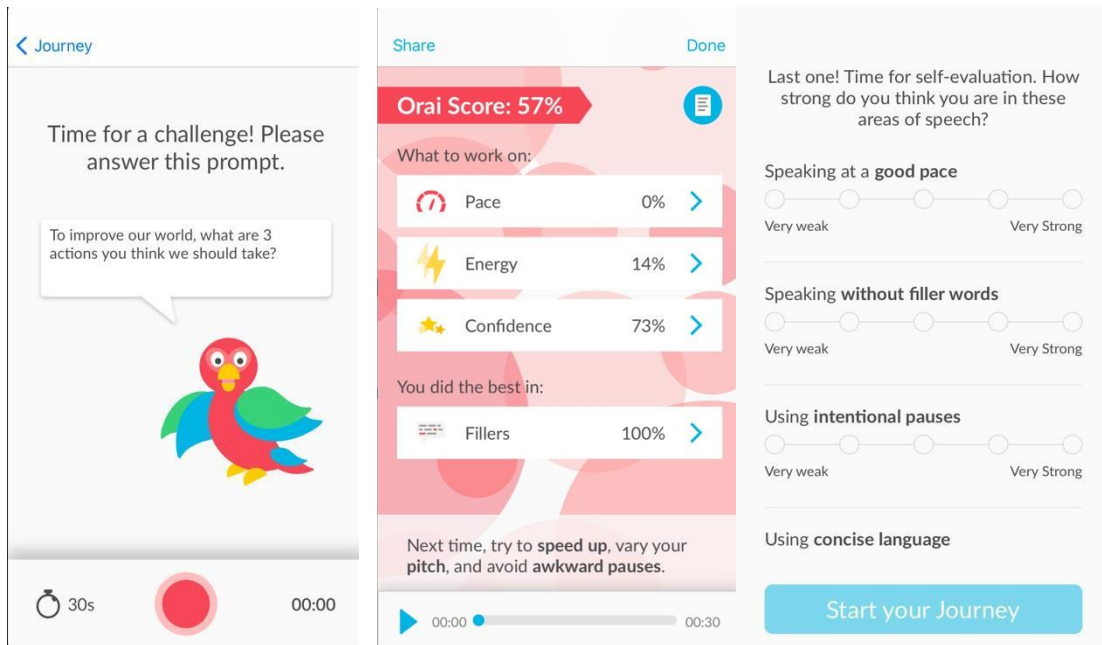
Figure 6: user pre-evaluation test to determine the speaking skills level

2.4 Speeko

A mobile app works for IOS and in future will be for Android users. The app prompts the user to choose the goal of the training (for work, school or personal growth), choose the time for daily training notifications, and choose the level of confidence. After signing up or subscribing, the app will provide numerous options of courses, speaking practices, and warm-up exercises. Also, the user could record a sample audio for (speech, interview, or random topics) and get instant feedback, then see the progress and set of achievements over the time. [4]
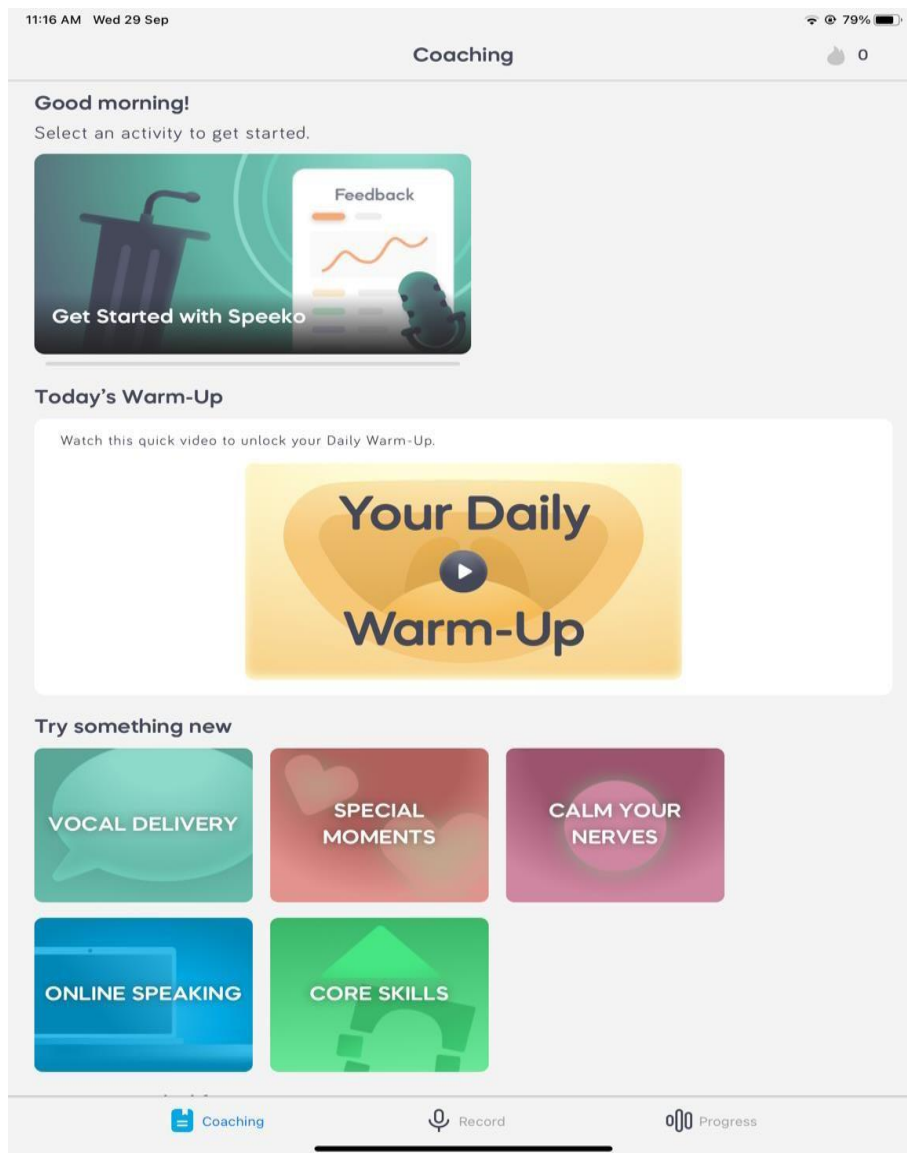
Figure 7: main interface of Speeko.

2.5 Machine Learning for Stuttering Identification (MLSI)

This search paper focused on stuttering, they defined stuttering and talked about the different types of it. They used different datasets for analysing stuttering.[5]

Machine Learning for Stuttering Identification: Review, Challenges & Future Directions
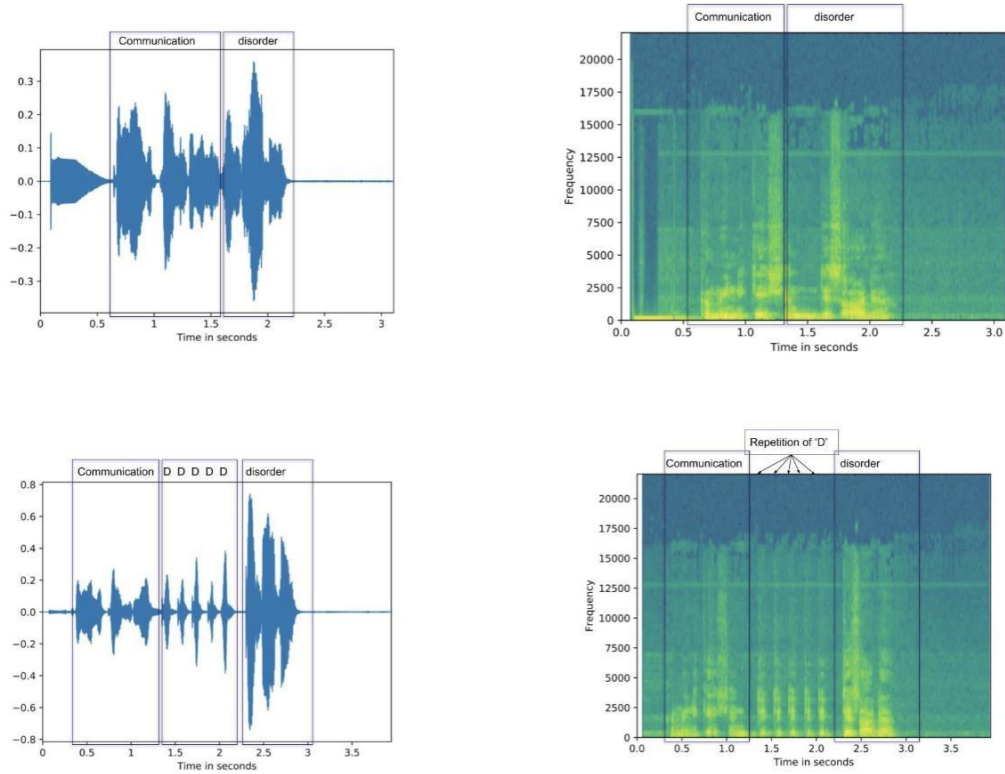
8

Figure 8: some of search paper example

2.6 Speech Stuttering Detection and Removal Using Deep Neural Networks

In this paper, state-of-the-art machine learning algorithms have been incorporated that have become more and more advanced over the past few years. In this study, the researchers used the UCLASS archives, which provide the time-aligned transcripts of stuttered speech in the. format. Their model has been optimized by hyper parameter tuning to maximize its accuracy by using different algorithms. Using random speech data with low to heavy stuttering from the same dataset, the algorithm is shown to significantly reduce the Word Error Rate (WER) for most of the test cases.[6]

2.7 Overview of Automatic Stuttering Recognition System

Stuttering is a disorder of speech communication. Many research has been done on stuttering recognition. There are 3 major classifiers used to classify types of dysfluencies between stutterers and nonstutterers. Those classifiers are Artificial Neural Networks (ANNs), Hidden Markov Model (HMMs), and Support Vector Machine (SVM). Each classifier provides different accuracies but HMMs is proven to give highest accuracy 96%, then SVM 94.35%, and ANNs 94.9%[7]

2.6 Comparative Study Between the Related Works

	LikeSo	VirtualSpeech	Orai	Speeko	PST
Platform	iOS	iOS, Android, VR headset	iOS, Android	iOS	Web
Fees/Full access	Not free to install	Needs subscriptions	Needs subscriptions	Needs subscriptions	Free
Network connection for installation and execution	-	Yes	Won't work without subscription, once the connection lost it logged out	Yes	Yes
User walkthrough	-	Yes	No	No	Yes
Personal plan customization	-	Yes	Yes	Yes	No
Disability consideration	Not considered	Not considered	Not considered	Not considered	For people with myopia/hyperopia in seeing, and blind people
Browsing files	No		No	No	Yes
User interaction	-	Realistic audience	-	-	Gif audience
Extra tools to run up.	-	Yes	No	No	The application will only need any device and microphone if needed.
Notifications	App	App	App	App	App+email
Performance report	Provide a report on the talking time, the pace, and filler words.	Provide a report on the eye contact, filler word, and pace.	Provides a report on audios, and overall performance.	Provide set of achievements.	Provides a report on uploaded audio, recorded audio ,and overall performance.

Table 1: related work vs PST 10

3 Proposed Project

3.1 Methodology

The PST is a web application working on windows operating system, written using python and the data is processed and analysed using MySQL database. The reasons for selection are written with details on section 5 (alternative solutions). The main techniques of PST are:

- Signal processing which is a special part of speech processing. It will be used to process voice audios and extract the important information like stuttering or silence moments. Then, provide a report to the user with those moments for self-improvement.

- Deep learning (DL) is a part of Machine learning (ML) which is also a part of Artificial intelligence (AI). ML provides the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it to learn for themselves [8]. Whereas DL has a multi-layered design known as Neural Networks. This structure is similar to human brain and tries to derive inferences like human would do. This is done by trying to find patterns of new information and classifying them to the existing ones[9]. DL will be used to train a voice disturbance detection model. After choosing the suitable dataset, the data will be divided into two parts for training and testing with the hope of making a model providing satisfying results.

3.2 Project requirements

3.2.1 System features

This section lists system features, stimulus/response sequences, and the functional requirements associated with it.

3.2.2 User sign in / sign up

3.2.2.1 Description of feature

This feature is going to be the first thing when the application opens. The user can sign in with email and password, or directly select continue with Google or Facebook. Also, if the user forgets their password, then they can choose forgot password feature. If the user doesn't have an account yet, they can sign up.

3.2.2.2 Stimulus/Response Sequences

Stimulus: The user enters email and password, then click sign in.

Response: The system will validate the information by comparing the email and password data from the server, and then load home page.

3.2.2.3 Functional Requirements

- The system shall allow the user to sign up.
- The system shall validate the email and password.
- The system shall send email if the user forgets password.
- The system shall enter home page after successfully sign in.

3.2.3 Session type

3.2.3.1 Description of Feature

In this feature, the user should choose between two sessions, either before presentation, or in the presentation day.

3.2.3.2 Stimulus/Response Sequences

Stimulus: The user chooses which session he/she wants.
Response: The system will retrieve the correct page from the server.

3.2.3.3 Functional Requirements

- The system shall load the correct page for the chosen session.

3.2.4 Record speech

3.2.4.1 Description of feature

In this feature the user can record their voice, then the system should store the voice and analyse it.

3.2.4.2 Stimulus/Response Sequences

Stimulus: The user will open the mic of the device and start record their voice.

Response: The system will store the voice in the database and analyse it.

3.2.4.3 Functional Requirements

• The application shall record the voice.

3.2.5 User input

3.2.5.1 Description of feature

In this feature, the user can input pdf, and mp3 files. Then the system should analyse the voice to show the awkward silence moments or stuttering.

3.2.5.2 Stimulus/Response Sequences

Stimulus: The user clicks on upload file to upload the pdf file, and clicks on upload voice to upload mp3 sound.

Response: The system will analyse the voice.

3.2.5.3 Functional Requirements

- The application shall take the documents (pdf) as an input.

- The application shall read the voice as input (.mp3).

- The application shall know the awkward silence moments and stuttering.

3.2.6 Viewing user performance

3.2.6.1 Description of feature

In this feature, the user can see his/her performance by displaying a report of their performance at any time and day.

3.2.6.2 Stimulus/Response Sequences

Stimulus: The user will open the report tab to see the report history.

Response: The system will retrieve the stored data and display them as chart, and table of time of the awkward silence moments and stuttering.

3.2.6.3 Functional Requirements

- The application shall display the report to inform user performance.

- The application shall list to the user the times when silence or stuttering happened from the record.
- The user could track his improvement and training history.

3.2.7 Settings

3.2.7.1 Description of feature

In this feature, the user is able to update his/her information, turn on/off notifications, log out, contact the developers, enable microphone access, view a walkthrough video about the application, enable auto reader, and change font size.

3.2.7.2 Stimulus/Response Sequences

Stimulus: The user will click on the setting icon and choose one of the actions to be done.

Response: After checking the actions selected by the user, the system will decide what actions to do based on what was selected.

3.2.7.3 Functional Requirements

- The system shall allow user to edit his/her profile from the application.

- The system shall turn notifications on/off according to user action.

- The system shall retrieve the walkthrough video from the database.

- The system shall be flexible on changing the font size from the application.

4.2 Non-Functional requirements:

- The application should be easy to understand and use.

- The application should be available for many users at any time.

- The application should not fail if anything wrong happens.

- The application should be as accurate as possible to provide clear statistics.

3.3 Alternative solutions/approaches

In this section, the report demonstrates the alternative development solutions/approaches.

3.3.1 Application Type

This section discusses several applications types alternatives, along with their advantages and disadvantages, as given in *Table 2*, and then explains why an application type is selected.

Alternative Solution	Advantages	Disadvantages
Desktop application	1. The notifications will be limited so that the presenter can concentrate on their presentation. 2. Offers offline capabilities 3. More secure 4. Fast processing	1. Hard to carry it, if the presenting venue don't have its own device.
Mobile application	1. Easy to carry anywhere.	1. The presenter may be distracted during presenting due to mobile notifications 2. Not all places allow the presenters to present using mobile phones.
Web application	1. Always up-to-date 2. Runs easy 3. A Better User Experience 4. Flexible Access	1. Internet Reliance 2. Website Dependency
Web application is selected as an application type.		

Table 2: application type

From the table of the comparison above, web application will be more convenient for the presenters to use, because they would have some difficulties installing and repairing. Also, you can access the web application at any time whenever you need to work if there was internet.

3.3.2 Operating System Type

This section discusses different operating systems for web devices, along with their advantages and disadvantages, as given in *Table 3*, and then explains why the chosen operating system is selected.

Alternative Solution	Advantages	Disadvantages
Microsoft Windows	1. More popular for installing apps 2. Installation is more straightforward 3. Easy to use	1. High computer resources, if you are installing windows OS then your computer should have high ram capacity
MacOS	1. Mac users get good response from its support team as compared to other operating systems. 2. Similar GUI for all the products	1. Not fixable for installing apps due to security reasons 2. High price
Linux	1. Open source 2. No anti-virus software needed 3. Powerful command prompt	1. Learning Linux is difficult 2. Few software and games applications
All OS types are selected as a Web operating system		

Table 3: operating System Type

From the table of the comparison above, all OS types are chosen since web application is independent, and it is better for people to use too.

3.3.3 Programming Language for Building the Speech Recognition Model

This section discusses multiple programming languages for building the speech recognition model, along with its advantages and disadvantages, as given in *Table 4*, and then explains why the chosen programming language is selected.

Alternative Solution	Advantages	Disadvantages
Python	1. Most popular programming language used in Machine Learning area 2. Extensive collection of libraries that support Machine learning and deep learning areas, in addition to libraries for signal processing. 3. Strong community support 4. Handles large training datasets 5. Easy to learn 6. Used for a variety of purposes 7. Open source	1. Speed limitations since it is interpreted and not compiled 2. Memory consuming
Java	1. Strong community support 2. Easy to learn 3. Provides many libraries for Android applications	1. Heavy code weight
JavaScript (JS)	1. Client-side JavaScript is very fast 2. Simplicity 3. Popularity	1. Client-Side Security 2. JavaScript is sometimes interpreted differently by different browsers
Cascading Style Sheet (CSS)	1. It helps to make consistent and spontaneous changes. 2. It improves the loading speed of the page. 3. It has better device compatibility.	1. There could be cross-browser issues while using CSS.
Hyper Text Markup Language (HTML)	1. Is widely used 2. Every browser supports HTML Language 3. Easy to learn and use 4. HTML is light weighted and fast to load	1. It cannot produce dynamic output alone, since it's a static language. 2. Making the structure of HTML documents

		becomes tough to understand.
		3. Errors can be costly.
Python, HTML, CSS, and JS are selected for building the speech recognition model		

Table 4: programming language for building the speech recognition model

From the data provided by the table above, Python combined with HTML, CSS, and JS are used as programming language, python is the most popular programming language used in Deep Learning areas, it has an extensive collection of libraries for Machine Learning and Deep Learning, it has a strong community, it can handle large training datasets, is easy to learn, and it is open source. Python requires less code and can compile even when there are bugs.

3.3.4 Database for user information and vocal commands

This section compares different databases, along with their advantages and disadvantages, as given in *Table 5,* and then explains why the chosen database is selected.

Alternative Solution	Advantages	Disadvantages
SQLite	1. Supports Python programming language 2. The default database for Android applications 3. Easy to learn and setup 4. Can be used with almost all programming languages without any compatibility issues	1. Not scalable
Oracle	1. Ported to all different platforms 2. It is good to be used as to take a proper backup of your entire oracle online backup as well as recovery too. 3. It easily manages the multiple databases within the same transaction	1. Complex 2. Difficult to Manage 3. Cost of Oracle Database is too high
MySQL	1. It is cross platform database server. 2. Open-source database 3. It is globally recognizing the most secure and reliable database	1. Doesn't support SQL check constraints. 2. Doesn't have a good developing and debugging tool compared to other databases.
MySQL is selected as a Database.		

Table 5: database for user information and vocal commands

MySQL is described as a cross platform database server, open-source database, and it is globally recognizing the most secure and reliable database.

3.3.6 Deep Learning vs. Machine Learning

This section compares the difference between deep learning and machine learning, along with their advantages and disadvantages, as given in *Table 6*.

Alternative Solution	Advantages	Disadvantages
Machine Learning	1. Improve accuracy and efficiency to make better decisions. 2. it can handle multi-dimensional and multi-variety data.	1. ML algorithms need enough time to learn and develop to fulfil their purpose with a fair amount of accuracy and relevance. 2. High error-susceptibility.
Deep Learning	1. Features are automatically determined and optimized for desired outcomes, they do not need to be extracted in advance. 2. Robustness to natural variations in the data is automatically learned. 3. It is flexible to be adapted to new problems in the future. 4. It requires a very large amount of data in order to perform better than other techniques.	1. It is extremely expensive to train due to complex data models. 2. It is not easy to comprehend output based on mere learning and requires classifiers to do so.
Deep Learning is selected		

Table 6: deep learning VS ML

4 Preliminary Expected Outcomes

This section explains the expected outcomes in PST before completing the implementation. The application will support the following functionalities:

- Save the users information.

- Represent the document file for the presenter.

- Take an existing voice file as an input.

- Allow the users to record a voice directly.

- Process the recorded voice by using signal processing technique.

- Send a notification if the users be in silence moments or stuttering.

- Allow the users to track their performance improvement.

- Present a report to the users after each voice file processing.

- Allow the users to see their report performance as a table of time or voice.
- Allow the users to change the font size.

5 Appropriateness of analysis

In this section, we analyse PST from different points of view, including its interaction with the external environment, activities sequence, relationships between components, and the expected outcomes. The following diagrams illustrate the proposed system design: Context diagram, Use Case diagram, Activity diagram, and Entity Relationship diagram.

5.1 Context diagram

This section shows the context diagram of PST. It shows the entities in PST is interacting with.
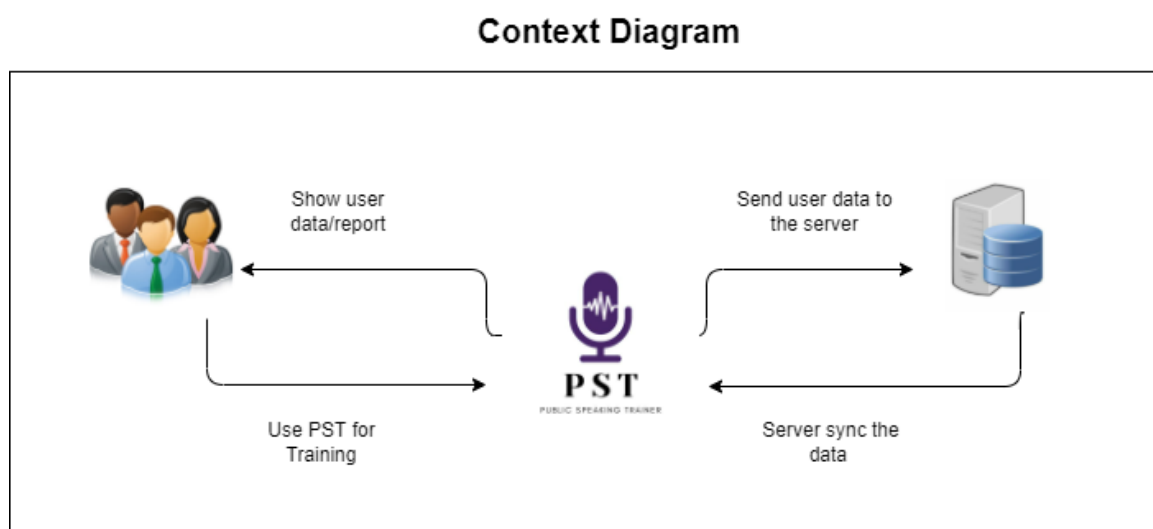
Figure 9: Context Diagram

5.2 Use case diagram

This section shows the Use Case of PST web application. The Use Case shows the interactions between the actors and the system.
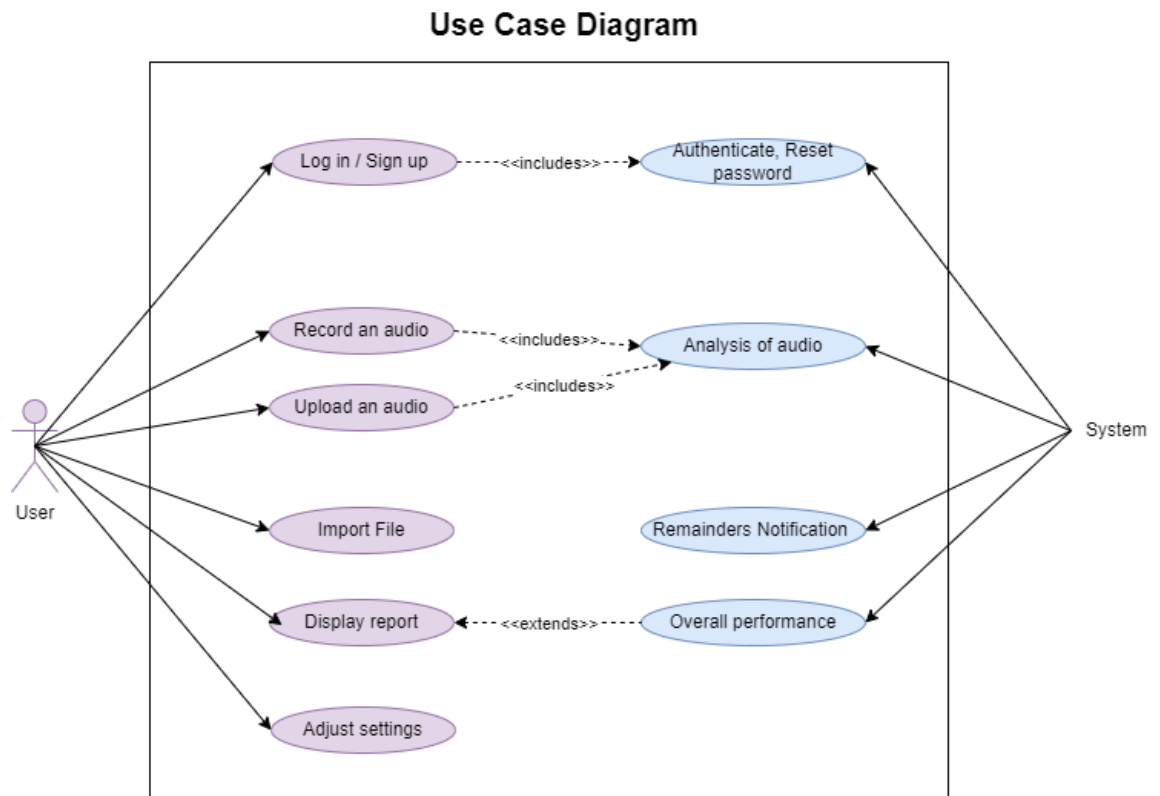
Figure 10: Use case Diagram

5.3 Activity Diagram

This section illustrates the behavior of the system as a series of actions or activities which is another form of a Flowchart diagram. In the activity diagram shown in *Figure 9*.

Figure 11: Activity Diagram

5.4 Entity Relationship Diagram (ERD)

This section demonstrates the Entity Relationship Diagram (ERD) for PST. It gives a complete overview of the database entities and their attributes, including their primary keys and foreign keys. As shown in *Figure 4*, it also shows the relationships among the entities and the types of cardinalities associated with each relationship.

Figure 12: ERD

6 Tools and Techniques Used During Project Proposal
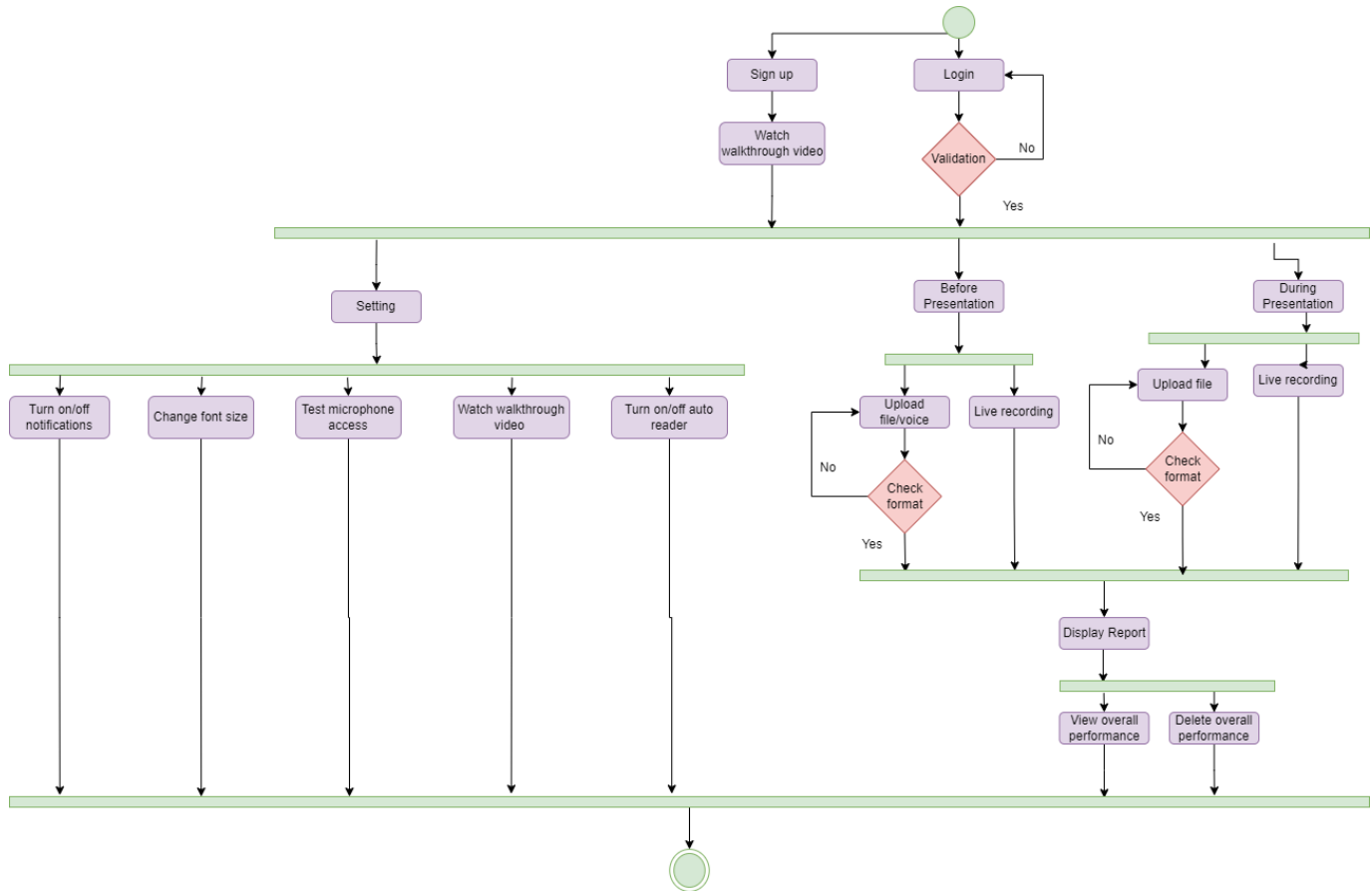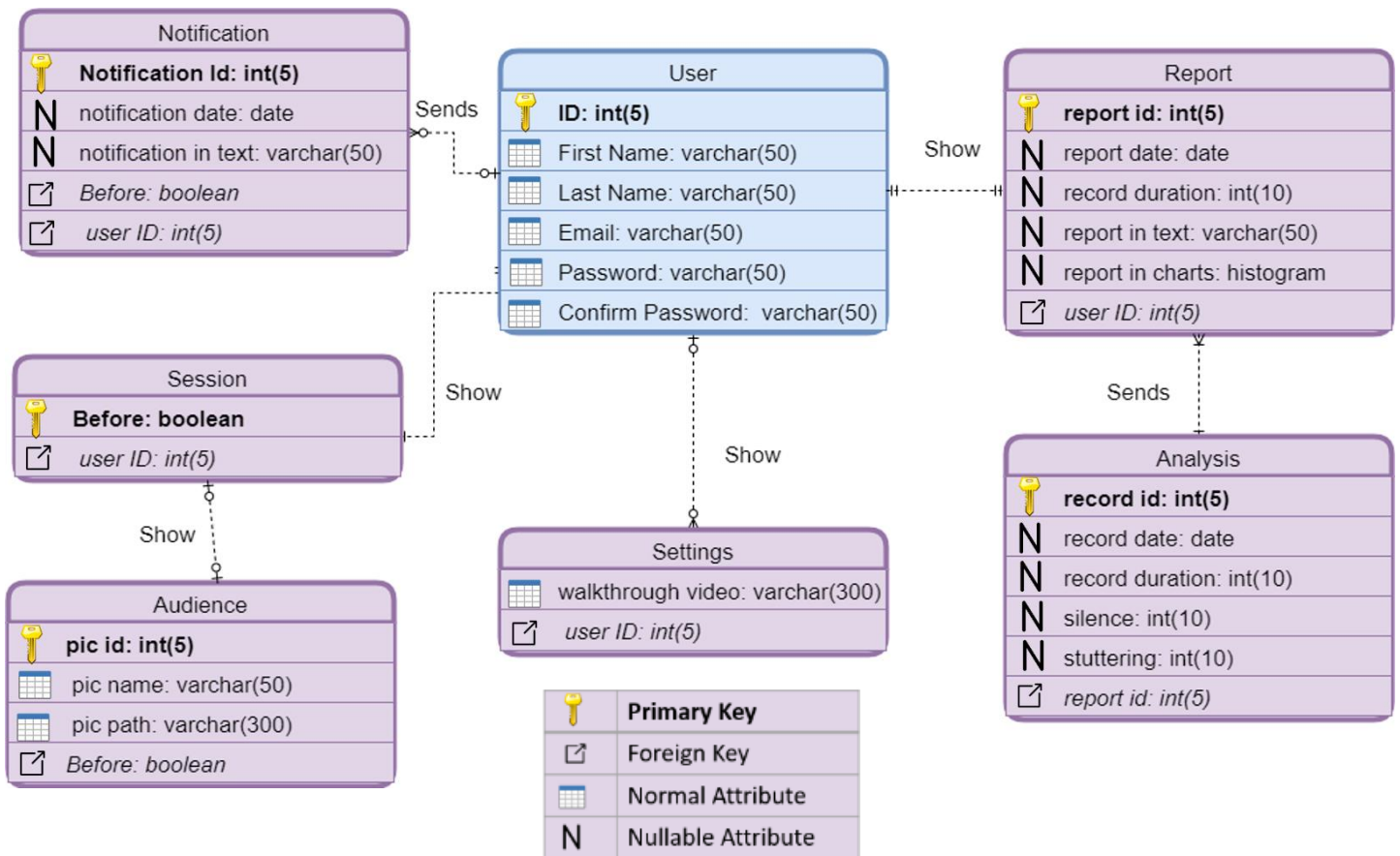
In order to carry out the proposal phase of PST, many tools and techniques have been used which have helped us to perform multiple functionalities. Such functionalities include; writing the reports, preparing the presentations, designing the proposed system, drawing the diagram and many other useful functionalities. *Table 7* lists and describes each tool that was used during the project proposal phase.

Tool/Technique	Description of Tool/Technique
Microsoft Office Word	Microsoft Office software that enables word processing, opening, saving, editing documents and many important features that help produce well-formatted documents. In our project, this software is used to prepare and finalize the milestone's reports neatly.
Microsoft Office PowerPoint	Microsoft Office software that enables the creation of slideshow presentations. We have used this software to prepare the presentations for each milestone to help interpret our work and ideas in a clear and attractive way.
Google Drive	Google Drive is an online file hosting service that offers cloud storage, synchronization between files. It has been used in our project to effectively share the work progress among the project team members and the supervisor.
Google Scholar	A web search engine for scholarly literature, conference papers, journal article and more. that is freely accessible. It is used in our project to investigate related work to our proposed project.
Cite This For Me	Cite This For Me has been used to cite references from multiple information and data resources in APA format to be referenced in the project reports.
Grammarly	Grammarly is an online grammar checking, spell checking, and plagiarism detection platform. It is used in our project to help checking milestone's reports.

Draw.io	Draw.io is an online drawing tool a free online diagramming tool that supports multiple diagrams and data viewing tools creation like; class diagram, flowcharts, and many more UML diagrams. This tool has been used in our project to help design the Use Case Diagram of related work.
creately	Creately is an online intuitive tool to create and share various types of technical diagrams. This tool was used in our project to create the Use Case and Activity Diagram.
Mockitt App	Mockitt App is a software tool that is mainly aimed for wireframing, prototyping and describing interactions of interfaces for various types of applications. This software was used in our project to prototype the interfaces in the system including the web app interfaces and the mobile app interfaces.

Table 7: tools and techniques used during project proposal

7 Tools and techniques required for implementation

At this stage of implementing PST, many tools and techniques have been used and utilized. It includes hardware, programming languages, frameworks and platforms …, etc. In the given *Table 8-9* lists and describes the tools and techniques used at this implementation phase along with the experience and knowledge gained.

7.1 Tool/Technique Primary

Tool/Technique Primary	Description of Tool/Technique
Personal Computer	The use of a personal computer is a requirement as all the work is done through it, in addition to building the PST web application and testing it.
Plug in Microphone	Plug in Microphone needs Plug-In Power to help two devices connected together. In our case we need a USB wire that establish the connection between the Microphone and the PC.
MySQL	MySQL is a relational database management system based on SQL - Structured Query Language. The application is used for a wide range of purposes, including data warehousing and logging applications. The most common use for mySQL however, is for the purpose of a web database. This will be used with BST web application to store the data of the users who will be using the system.

PyCharm	PyCharm compiler can help the programmer to know more about your code. Rely on it for intelligent code completion, on-the-fly error checking and quick-fixes, easy project navigation and more.
Dataset	This dataset (SEP-28k) contains over 28k clips labelled with five types of events, including blocks, prolongations, repeating sounds, repeating words, and interjections. The audio is primarily from public podcasts in which people who stutter interview other stutterers.[10]
PyQt5	Is a Python binding of the cross-platform GUI toolkit Qt, implemented as a Python plug-in. Which includes a substantial set of GUI widgets as a feature of drag and drop.

Table 8: primary tools and techniques used for project implementation

7.2 Tool/Technique Detailed

Tool/Technique Detailed	Description of Tool/Technique
tkinter	Tkinter Frame widget organizes a group of widgets using Python. It acts as a container by which other widgets can be placed. Widgets are organized into rectangular areas of the screen for the Python application.
Reset password libraries	The "Forgot your password" feature should allow users to recover their accounts by email. We will create a function that retrieves the user's email address and prompts them to input new passwords from an email link sent.
Upload file + Upload voice	Flask library is used to upload files from the user. ffmpeg-python library lets you upload or save audio in different file formats (more on this later), slice audio, calculate the length of audio files, fade in or out, and apply cross-fades. Librosa for audio analysis, it is used mainly when dealing with audio data, for Automatic Speech Recognition.
Notifications	Plyer is a Python library for accessing features of hardware/platforms. It will be used to send notifications from PST to the computer.[11]

Chart builder	matplotlib.pyplot is a library that contains collection of functions that can help to make some change in a figure. For example, creates a figure, creates a plotting area in a figure, plots some lines in a plotting area, decorates the plot with labels and more.
Time Library + Date Library	DateTime is a Robot Framework standard library that supports creating and converting date and time values.as well as doing simple calculations with them.It supports dates and times in various formats, and can also be used by other libraries programmatically.Python's standard DateTim objects can be used both in input and output. In input they are recognized automatically, and in output it is possible to get them by giving DateTim value to result_format argument.
Narratore	Our program will need the access to window narrator, so it can read every single word in the application and provide the support for people with special needs.
Microphone tester	This function aims to confirm that the connected microphone is working just fine. And Pyaudio library might be useful for this.[12]

Table 9: detailed tools and techniques used for project implementation

8 Details of Proposed Design Conforming to the Problem Statement

This section presents the proposed design for the interface of PST web application. The PST web application is designed based on the consistency and flexibility standards for the users. This prototype combining the interfaces of our application in order to enhance the learnability, consistency, flexibility, recognition rather than recall and the visibility of the notifications.

Figure 13: welcome page

Figure 14: main page

- All users who are entering the application will start with the welcome page.
- Each user wants to access our application must have an account.
- In the main page the application contains drop list to let the users choose the suitable font size for to help myopia or hyperopia people as shown in *figure 14*.
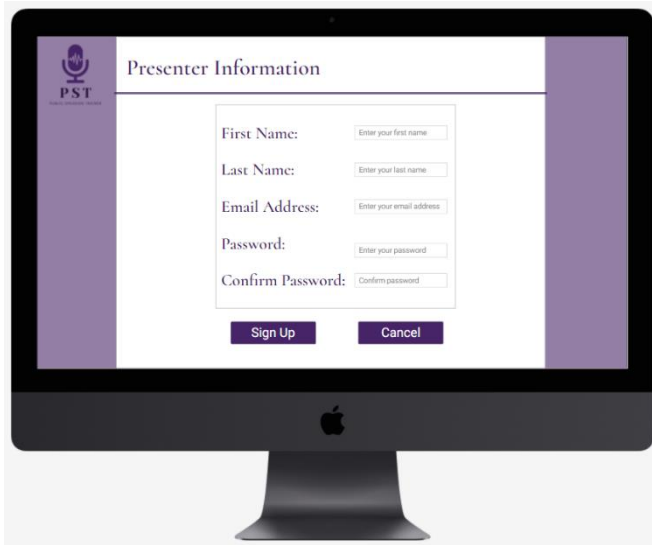- Moreover, in the main page the application will contain two options sign up or sign in as shown in *figure 14.*
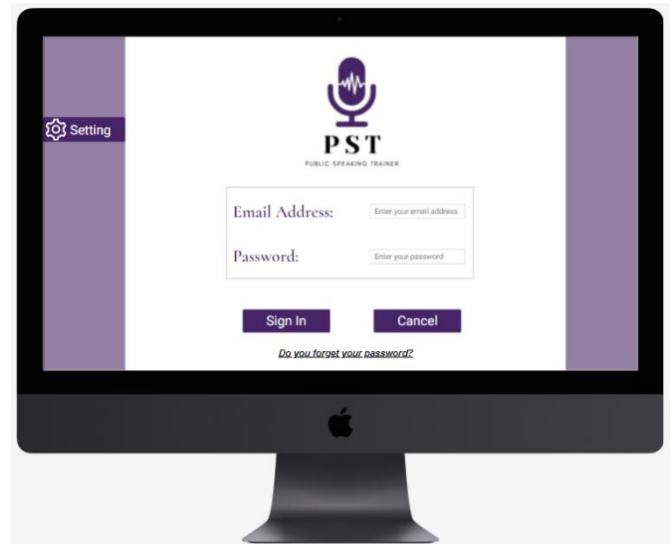
Figure 15: sign up page

Figure 16: sign in page

- If the users click on sign up, they must identify themselves and full their information as shown in *figure 15*.
- After the users clicking on sign in button from the main page, the users should enter their email and password as shown in *figure 16*.
- Moreover, in both pages there is a "Cancel" button to help the user go back to the previous main page easily.
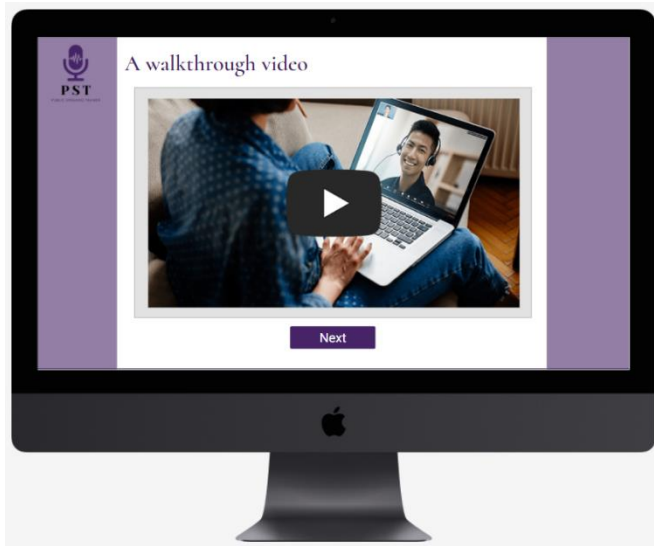
Figure 17: walkthrough page

Figure 18: main functionalities page
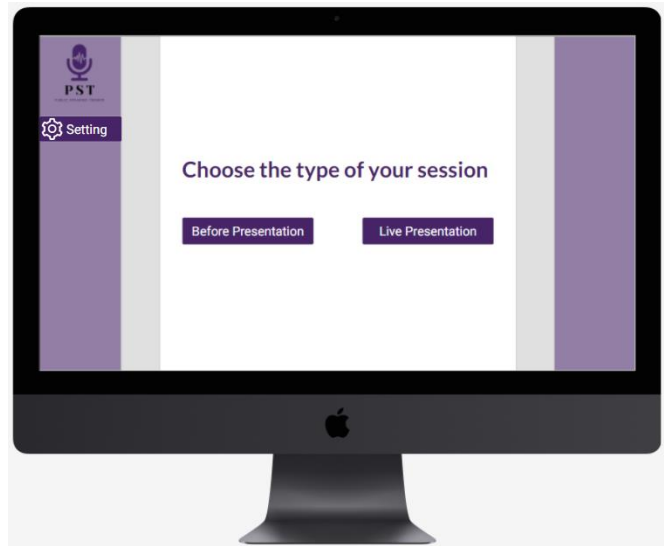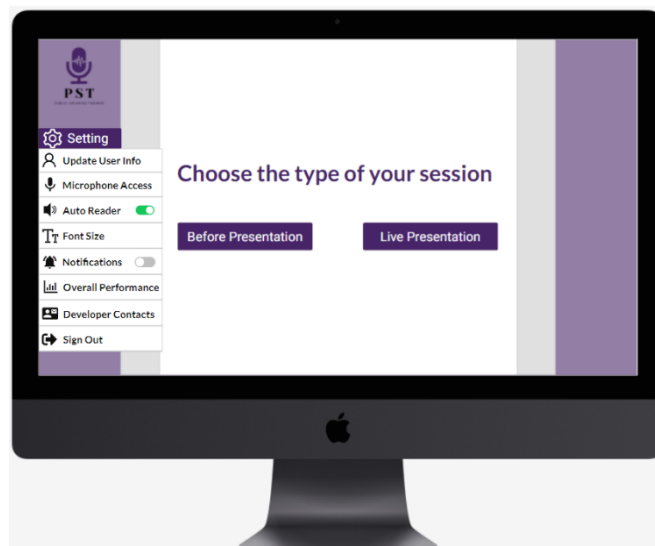

Figure 19: settings page

- After the users done with sign up, the users will see the walkthrough video to help them know how to deal with PST web application easily as shown in *figure 17*. The walkthrough video is only for new users who are singing up to PST web application for the first time.

- The "Next" button helps the users to skip the walkthrough video to go to sign in page.

- After the users done with sign in process, the users will see the main functionalities page as shown in *figure 18*.

- In *figure 18*, the users will have two types of sessions either before or live presentation session. Each one of the sessions will support different functionalities.

- Moreover, if the users click on "Settings", the user will get dropdown list. In the dropdown list the user will find sup functionalities as shown in *figure 19*.

- The users can test their microphone by clicking on "Microphone Access" which will enable the system to access the microphone and check if it is working or not.

- The users can turn on or off the "Auto Reader" functionality which will help blind people to hear the text on PST web application and by default it will be on.

- The users can turn on or off the "Notifications" functionality which will enable the users to receive some emails from PST application and by default it will be off.

- If the users want to sign out from the application, they can directly click on "Sign Out".
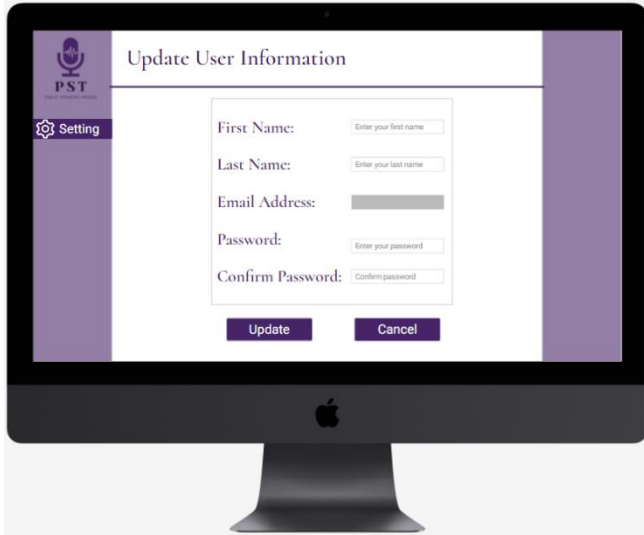
Figure 20: update user information page
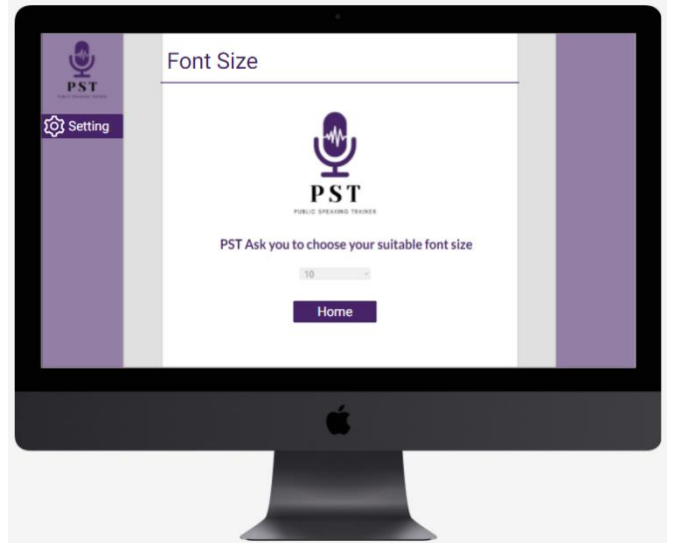
Figure 21: font size page
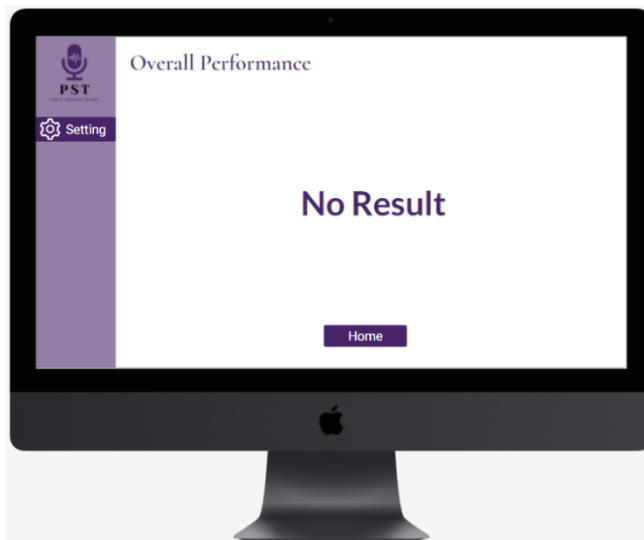
Figure 22: empty overall performance page

Figure 23: result of overall performance

- If the users choose "Update User Info" option, the users can update their information without updating the email as shown in *figure20*. Moreover, if the users click on "Update" button the update will be confirmed.

- If the users choose "Font Size" option, the users can change the whole font size of PST web application as shown in *figure 21*.

- If the new users click on "Overall Performance" button, the users will get no result since the users do not record or upload any voice till now as shown in *figure 22*.

- If the users choose "Overall Performance", the users will be able to see the overall performance as a chart only as shown in *figure 23*. This chart will represent the average of deducting per days for all recorded voices.

- Moreover, if the users click in "Home" button, they will go back to the main page and if the users click on "Clear History" the performance follow-up result will be removed.
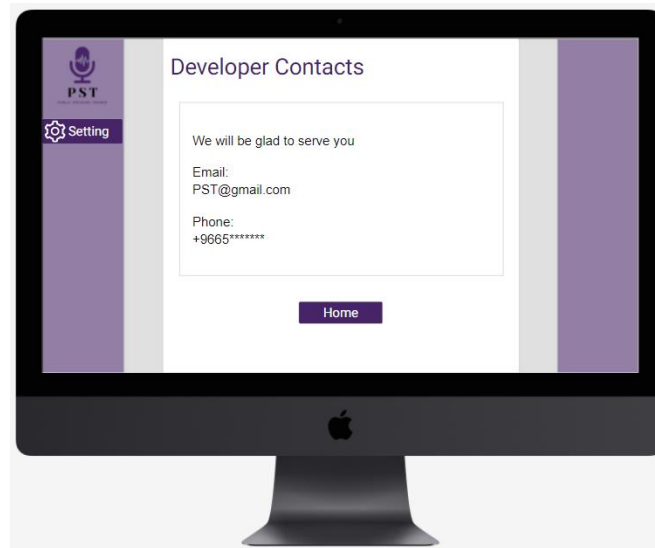

Figure 24: developer contacts' page

- If the users choose "Developer Contacts" the users will be able to contact the developers through their email or mobile number.

- Moreover, if the users click in "Home" button, they will go back to the main page.
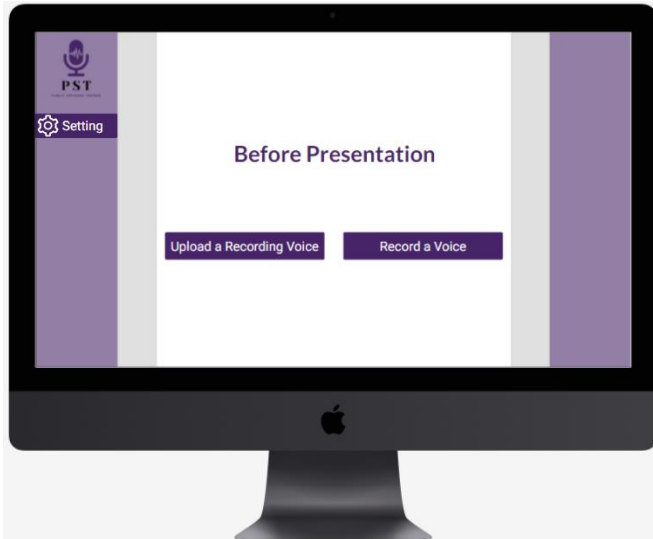
Figure 27: before presentation
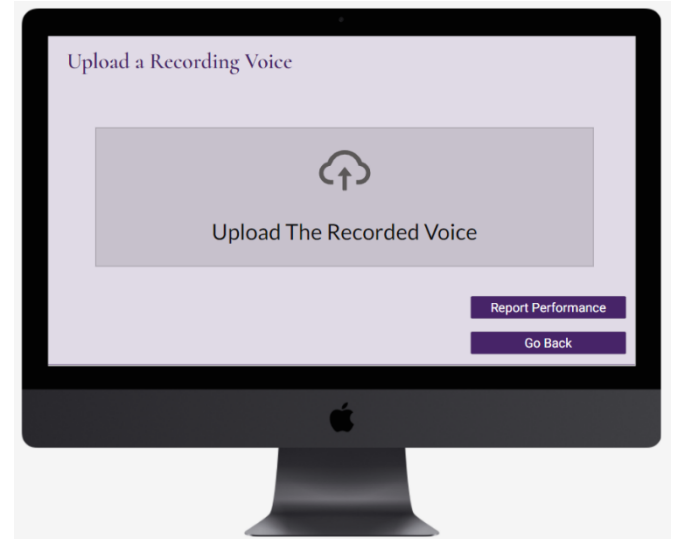
Figure 28: upload a recording voice page
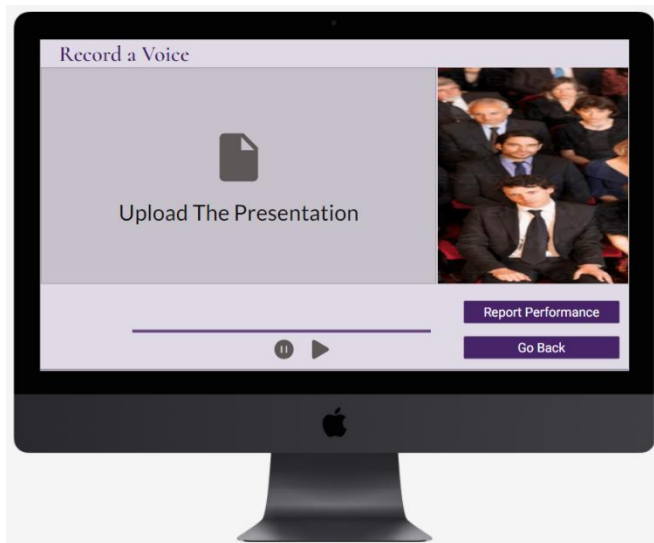
Figure 25: record a voice page before any interaction
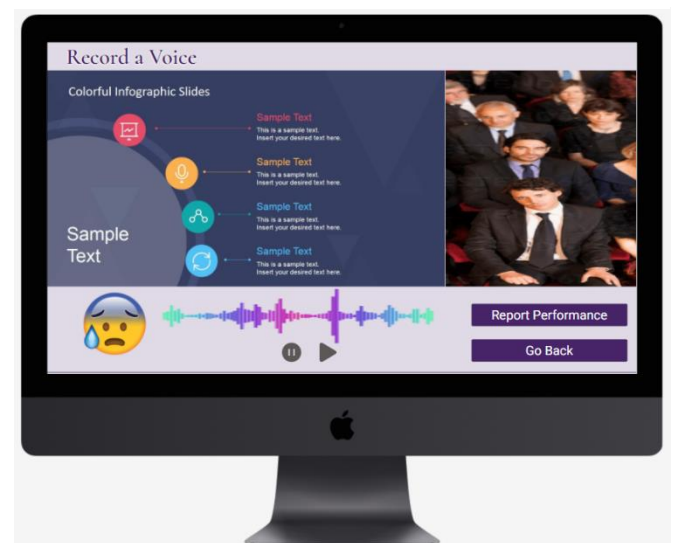
Figure 26: record a voice page after the interaction

- If the users choose before presentation type, the users will have two options. The first option is "upload a recording voice". The second option is "record a voice" as shown in *figure 25*.

- If the users choose "upload a recording voice", the users will be able to upload an existing recorded voice file as mp3 as shown in *figure 26*.

- If the users choose "record a voice", the users will be able to upload their presentation file as PDF file and start new recording voice. While the users are

42

recording, they will see virtual audience to increase the feeling of reality as shown in *figure 27*.

- While the users are recording and they have stuttering more than one time or they in a long silent moment, the PST web application will send to them a notification as an emoji to help hem recontrol themselves.

- Moreover, in both pages there is a "Go Back" button to help the user go back to the previous page easily.
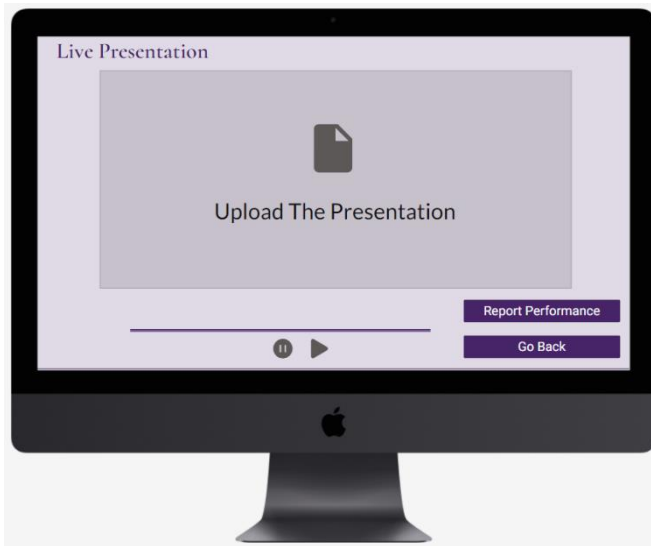
Figure 29: live presentation page before any interaction
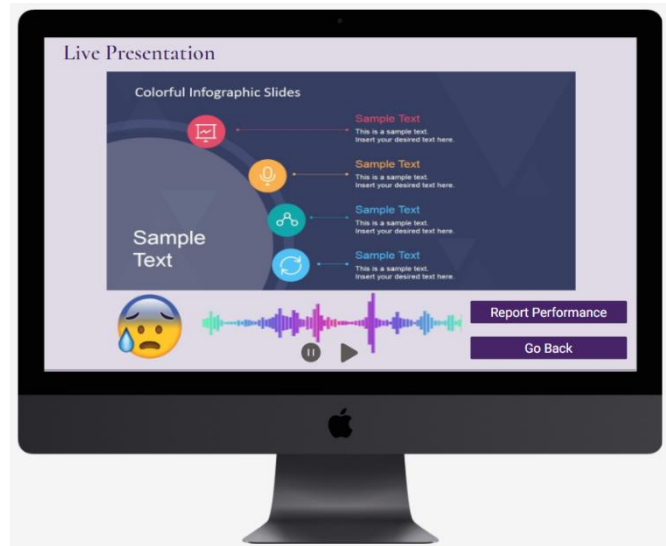
Figure 30: live presentation page after the interaction

- If the users choose "live presentation", the users will be able to upload their presentation file as PDF file and start live recording voice as shown in *figure 29*.

- While the users are recording and they have stuttering more than one time or they in a long silent moment, the PST web application will send to them a notification as an emoji to help hem recontrol themselves as shown in *figure 30*.

- Moreover, if the users click in "Go Back" button, they can go back to the previous page.
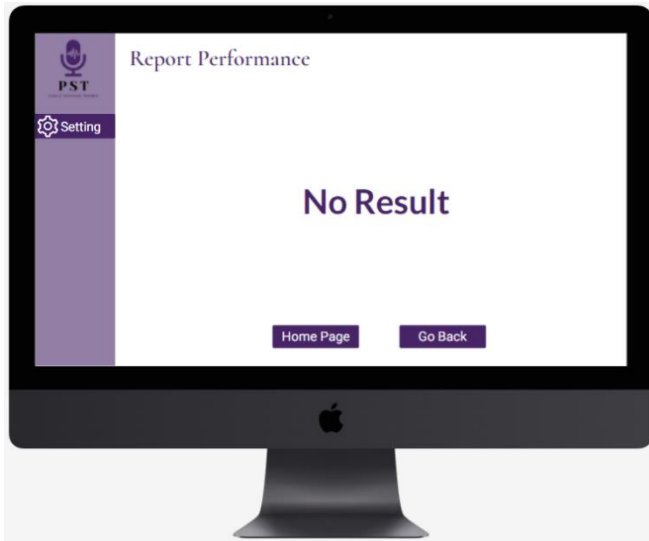
Figure 31: empty report performance page
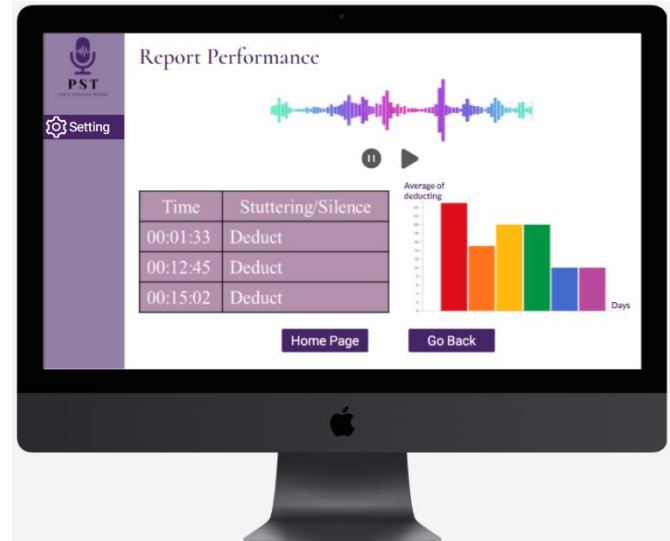
Figure 32: result of report performance page

- As shown in *figure 26, figure 27 and figure 29*, if the users click on "report performance" button without record or upload any voice file, the user will get no result as shown in *figure 31*.

- However, if the users click on "report performance" button after recording or uploading the voice, the user will see the report performance result as shown in *figure 32*.

- In order to support blind people, the PST web application will help them hear their report performance by using the mouse over in the column time. PST web application will link the blind people to the recorded voice file to enable them to hear the moments that the application deducted the stuttering and long silent moments.

- As shown in *figure 32*, the chart will show the average of deducting per days for all recorded voices.

- Moreover, if the users click in "Go Back" button, they can go back to the previous page and if the users click in "Home Page" button, they will go back to the main page.

9 Conclusion

Many presenters in the world face some difficulties while they are presenting. These difficulties appear because of the fear, anxiety and lack of self-confidence. these difficulties may lead the presenters to some problems. One of these problems is lack of audience understanding. Moreover, our application will support disabled people. In response to these problems, PST will be used to provide an effective way for presenting that help both able and disabled people. PST is expected to contribute to achieving various outcomes. Firstly, the application will enable the users to practice before and during the presentation. This will help the users to end up with a good performance result. Secondly, the application will process the recorded voice by using signal processing technique in order to find silence and stuttering moments. The result of that the users will be received a report in text, voice and chart form. Thirdly, the application will send a notification to the users during presentation if the users be in silence moments or stuttering to help the users to readjust themselves. Fourthly, the application will allow the users to change the font size. To help the people who have myopia/hyperopia problems. We hope by working on this project that we will be able to create a practical system from all perspectives and we wish this project will help us master new skills and techniques and explore various levels of knowledge.

Appendix A: Identified Task and a Tentative Work Plan for Project Proposal

The table below presents the project plan that is followed during the project proposal. It shows the different tasks that are performed weekly in order to accomplish PST proposal.

First Semester: Proposal Phase

Activity / Week		W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	W13	W14	W15
Explain the project idea	Determine the project motivation and background	░	░													
	Explore the problem statement and the scope of the project		░	░	░											
	Detect the project baseline requirements				░											
Analysis	Review and analysis related work		░	░												
	Discuss the detailed Project Requirements				░	░										
	Identify the alternative solutions of the problems					░	░									
	Submitting first milestone							▓								
	Presenting the milestone								▓							
Designing a computer-based system	Conduct an appropriate analysis to the system design conforming to the problem statement									░	░					
	Describing the tools and techniques to be used during project implementation										░	░	░	░		
	Identifying a realistic work plan for project implementation												░	░		
	Submitting second milestone														▓	
	Presenting the milestone															▓

Table 10: work plan for project proposal

Appendix B: Identified Tasks and a Realistic Work Plan for Project Implementation

The table below shows the different tasks that have to be performed weekly in order to accomplish the implementation of PST.

Second Semester: Implementation Phase

Task/Week	W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	W13	W14	W15
Software tools that are needed for implementing the project	■	■													
Hosting/subscribing to a server and creating the system database.		■	■												
Creating the application interface + Integration of libraries and tolls that will be used.			■	■											
Designing and developing the application (implementing most of the functionalities)				■	■	■									
Connection of the database with PST web application.					■	■									
Applying signal processing technique with the application.						■	■								
Analysing and concluding the results.							■								
Project Milestone 3 presentation and evaluation								■							
Continuing the implementation of the pending/incomplete functionalities								■	■	■	■				
Comparing, validating, verifying the overall result and outcome										■	■	■			
Final system integration, documentation and testing. + Applying deep learning technique to test PST model.													■	■	
Project Milestone 4 presentation and evaluation															■

Table 11: work plan for project implementation

References

[1] Say It Media. (2019, January 25). *Home*. Say It Like So. https://sayitlikeso.com/

[2] VirtualSpeech. (n.d.). *VirtualSpeech - Soft Skills Training with VR and Simulations*. Retrieved October 2, 2021, from https://virtualspeech.com/

[3] Ai-powered app for practicing your presentations. Orai. (n.d.). Retrieved September 29, 2021, from https://www.orai.com/.

[4] The #1 public speaking app. Speeko. (n.d.). Retrieved September 29, 2021, from https://www.speeko.co/.

[5] A. SHEIKH, S. H. A. K. E. E. L. (2021, July 12). *Machine Learning for Stuttering Identification: Review, Challenges & Future Directions*. Machine Learning for Stuttering Identification.
https://arxiv.org/pdf/2107.04057.pdf

[6] Rajput, S. (1970, January 1). *Speech Stuttering Detection and Removal Using Deep Neural Networks*. SpringerLink. https://link.springer.com/chapter/10.1007/978-981-16-6554-7_50?error=cookies_not_supported&code=4dba1f6c-d04b-459f-a5d3-56216ffa20e5

[7] Chee, L. S., Ai, O. C., & Yaacob, S. (2009, October). Overview of automatic stuttering recognition system. In Proc. International Conference on Man-Machine Systems, no. October, Batu Ferringhi, Penang Malaysia (pp. 1-6).

[8] *What is machine learning? A definition.* Expert.ai. (2021, May 26).

Retrieved   October   13,   2021,   from
https://www.expert.ai/blog/machine-
learningdefinition/#:~:text=Machine%20learning%20is%20an%20appli
cation,it%20to%20learn%20for%20themselves.

[9]Oppermann, A. (2020, August 11). *What is deep learning and how does it work?* Medium. Retrieved December 8, 2021, from https://towardsdatascience.com/what-is-deep-learning-and-how-does-it-work-2ce44bb692ac.

[10] Et Al, L. (n.d.). *Papers with Code - SEP-28k Dataset*. SEP-28k.

Retrieved   October   9,   2021,   from
https://paperswithcode.com/dataset/sep-28k

[11] O'Hanlon, C. (2021, March 3). *Send a desktop notification using Python*. Conor J O'Hanlon. Retrieved December 4, 2021, from https://www.conorjohanlon.com/send-a-desktop-notification-using-python/#:~:text=To%20send%20a%20notification%20to%20your%20desktop%20using,file%20and%20start%20by%20importing%20notification%20from%20plyer%3A.

[12] Real Python. (2021, January 16). *Playing and recording sound in Python*. Real Python. Retrieved December 4, 2021, from https://realpython.com/playing-and-recording-sound-python/#recording-audio.