Voice Authentication System

SAAD ZAHID QURESHI

BILAL JABBAR

Department of Computer Science
COMSATS University Islamabad,
Attock Campus-Pakistan

SESSION 2018-2022

VOICE AUTHENTICATION SYSTEM

Undertaken By:

SAAD ZAHID QURESHI
CIIT/FA18-BSE-086/ATK
BILAL JABBAR
CIIT/FA18-BSE-076/ATK

Supervised By:

Dr. SHAHZAD FAISAL

Co-Supervised By:

Mr. NAJAM DAR

A DISSERTATION SUBMITTED AS A PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF BACHELOR OF SCIENCE IN SOFTWARE ENGINEERING

Department of Computer Science
COMSATS University Islamabad,
Attock Campus-Pakistan

SESSION 2018-2022

UNDERTAKEN

We certify that this is my/our own work. The work has not, in whole or in part, been presented elsewhere for assessment. Where material has been used from other sources it has been properly acknowledged. If this statement is untrue, we acknowledge that we will have committed an assessment offence and shell be liable to punishable action under the plagiarism rules of HEC.

___________________	___________________
Saad Zahid Qureshi	Bilal Jabbar
FA18-BSE-086	FA18-BSE-076
Dated: ______________	Dated:______________

FINAL APPROVAL

Certified that we have read this project report submitted by Mr. (Saad Zahid Qureshi and Bilal Jabbar) and it is, in our judgment, of sufficient standard to warrant its acceptance by Department of Computer Science, COMSATS University Islamabad, Attock Campus, for the BSSE degree.

Committee:

1. External Examiner

____

Dr. Muhammad Ishtiaq
Assistant Professor
Department of Software Engineering,
Foundation University Islamabad

2. Supervisor

____

Dr. Shehzad Faisal
Assistant Professor
Comsats University Islamabad
Attock Campus

3. Co-Supervisor

____

Mr. Najam Dar
Lecturer
Comsats University Islamabad
Attock Campus

4. Head of the Department

____

Dr. Khalid M Awan
HoD CS Department
Comsats University Islamabad
Attock Campus

DEDICATION

This all work which is done by us is dedicated to our parents and Respected Teachers, without their prayers and effort we could not able to do anything.

ACKNOWLEDGEMENT

In the name of ALLAH, the most merciful and beneficent. I'm grateful to the ALLAH Almighty who provides all the resources of every kind to me, so that I make their proper use for the benefit of mankind. May He keep providing me with all the resources, and the guidance to keep helping the humanity.

I would like to thank our parents who kept backing us up in all the times, both financially and morally. I'm thankful to them for their prayers.

I would also like to thank my supervisor Dr. Shehzad Faisal for his guidance and encouraging me to work hard and smart and for his moral support to accomplish our project. His critical comments on my work made me think of new ideas and techniques in the field of optimization and software development. I have found him very helpful while discussing the optimization issues.

I'm very thankful to COMSATS University Islamabad, Attock Campus for giving me enough knowledge and skills that made me to innovatively.

ABSTRACT

The work leading to this thesis has been focused on establishing a text-independent speaker recognition system. The method of identifying the speaker based on the voice characteristics of a given speech is known as speaker recognition. SR is based on finding and extracting distinct aspects of a speaker's speech. Voice biometrics refers to the characteristics of a person's voice. Since the technology is evolving and there are developing different ways for security breaching. Our aim is to provide the security to the systems and to make the security more authenticate by using voice biometric rather than text-based password. This system is built using GMM algorithm. GMMs are often used in biometric systems, most notably in speaker recognition systems, due to their capability of representing a large class of sample distributions. For this, we are going to use MFCC features to decrease misclassification. The MFCC computation is a simulation of the human hearing system that aims to artificially recreate the ear's working principle, assuming that the human ear is a trustworthy speaker recognizer. that’s how user will be differentiated based on the features that extracted from the voice.

TABLE OF CONTENTS

List of Figures

List of Tables

Chapter 1

Introduction

Introduction

In the recent few years, it has become increasingly more popular to use voice recognition/speaker recognition for applications. Applications like Google's Alexa Now and Siri can transcribe audio and understand what a user is saying. For example, Shazam and projects like MusicBrainz have been able to identify music using only 8-10 seconds of recorded music. There are two approaches for voice recognition one is text dependent recognition and other is text independent voice recognition. Voice authentication system is a text independent speaker recognition which help to identify the speaker and authenticate. Our system that provides a secure and reliable for the banking system to authenticate their users by their voices by using this system with their existing systems. The approach is to have an individual speak different sentence multiple times and the system extract the features, make voice print and the next step train the model. The generated voice print after user speaks, authenticates the user by matching with old voice print. There are many good features in this system. The enrollment of new user voice is not much time taking. The user speaks casually for a certain time. Then a specific model will create and train on those voice sample(s). The digital representation of those samples is called voice print. At the time of logging next time into the system. The user needs to give a voice for authentication. The authentication will be accomplished by matching and comparing the current voice print with the available voice print saved at enrollment time. The authentication process is fast enough that it will take only a few seconds.

1.1 Problem statement

The fingerprint biometric is currently used as a security feature. But if a person (user) is disabled and couldn't use fingers. He couldn't use those systems. As well as we know the current situation all over the world. Different deadly viruses can spread through physical interaction and multiple persons use same device (fingerprint sensor). By the touching of that virus can be spread. But for our system user doesn't need to interact barely with devices like fingerprint sensor. Main purpose is to make a secure and the reliable system that allow the consumer to enter into the system through their voice. The system tells in no time that given voice is belong to the authentication user or not on the basis of individual information included in voice. This is more easy and reliable way than traditional way of authentication using textual demographics. The existing system is not as secure as our proposed system because the chance to unauthorized access is higher as compared to our proposed system.

1.2 What is Speaker recognition?

The method of identifying the speaker based on the voice characteristics of a given speech is known as speaker recognition. SR is based on finding and extracting distinct aspects of a speaker's speech. Voice biometrics refers to the characteristics of a person's voice.

A SR system is used to identify and differentiate speakers, as well as extract unique features that used to train model. The method of identifying the speaker from a particular utterance by comparing voice biometrics of a given sample of the speaker is known as Speaker Identification (SI). Speaker Verification refers to the use of speech for authorization. Security, call tracking, and forensic sciences are three main application areas for SR. Speaker Verification and Recognition systems are also used to replace passwords and other forms of user authentication (voiced password). Speaker Identification is used in forensic science to compare the voice samples of the person alleged to be with other evidence such as phone conversations or other recorded evidence. Speaker detection is another name for this approach. The most significant part of employing Speaker Identification systems is for automating operations such as forwarding customers' mails to the correct mailbox, detecting talkers in discussions, alerting discourse acknowledgement frameworks of speaker changes, and so on.

1.2.1 Actual system

Voice authentication system is a text independent speaker recognition which help to identify the speaker and authenticate. Our system that provides a secure and reliable for the banking system to authenticate their users on their voices by using this system with their existing voice-based models. The approach is to have an individual speaks different sentence multiple times and the system extract the features, make voice print and train the model and the next step train the model. When the user speaks again in the future, the new voice print created will be compared to their previous voice prints, allowing the person to be authenticated.

1.3 Project scope

The project scope outlines what the system will and will not accomplish.

- The system may accept previously recorded voice or runtime as input. Other sources, such as music or extensive talks, will be rejected by the system.
- All new users must register at least once.
- The system may do a basic microphone test before each speech recording to reduce recording mistake.
- To train the network, each user needs enroll in only a few times.
- To confirm, the system will perform speech recognition by comparing the user's live template to their individual voice templates previously saved in the network.
- The system will simply validate a user's identity, not identify them.

1.4 Project Objectives

The goals sought to attain by implementing this system are outlined in the project objectives mentioned below.

- To design and implement a more inventive and secure login mechanism for accessing a personal or networked computer system.
- Create a system that can validate a user's voice regardless of gender or age to authenticate them.
- To investigate and use artificial intelligence (AI) approaches to speacker recognition.
- Extend the capabilities of a computer system with minimum keyboard input.

1.5 Application

This system can be used in many fields like:

1. In the banking sector to protect your bank account from unauthorized access, where highly sophisticated data resided.
2. For security, to solve crimes with voice recognition. Like when it comes to a crime scene there are some forensic situations, in which only audio evidence is available to investigators, and that’s where voice biometrics can be deployed to great effect.
3. Buying products and services with the sound of your voice. This is one of the most popular and mainstream applications of biometrics in general is mobile

payments, and voice recognition has also made its way into this highly competitive area.

4. For call tracing, where suspect's location can be traced through their call voice.

1.6 limitations of project

There are certain limits to the speech identification method as well, because we know that the human voice is not invariant over time, therefore a biometric template must be used as time passes. The human voice changes throughout time as a result of colds, hoarsens, stress, and other emotional states. Voice identification, on the other hand, has a larger risk of mistake than fingerprint recognition since the human voice is not as distinctive as a fingerprint. The technology requires more processing power than fingerprint matching to compute rapid furrier transformations. As a result, in its current condition, speaker verification is not suited for mobile applications or battery-powered devices.

1.7 Feasibility report

During the analysis of the system, we get to know that there are four types of feasibility studies carried out to develop quality software. Four types of feasibility studies are given below:

- Economic feasibility
- Technical feasibility
- Legal feasibility

1.7.1 Economic feasibility

This software's needed are common and easily available. This work has been carried out on our personal computers. The logic, design and troubleshooting were analyzed manually. The project general software so there are no economic problem of implementation of hardware and software's needed to be installed in any organization

Hardware resource: Python free available.

Software resource: Microphone was required.

1.7.2 Technical feasibility

For the fulfillment of our project, we have used deep learning algorithms, which are commonly used for speaker recognition projects. There were some problems to understand these very complicated algorithms, but we find a lot of help from internet.

all troubleshooting and error were removed with through knowledge, discussion and research, which result in a well-managed, equipped and proper working project.

1.7.3 Legal feasibility

Although our project was not the type of project that needed legal feasibility study concerning violation of rules and regulations, but to avoid it we also conducted it. Our project wouldn't violate any rules or regulation as for as state's is concerned, and is not violating any social, legal or logical tradition or any other code of ethics. Few points are given below:

- The project wouldn't violate any religious, social or tradition government's laws.
- It wouldn't be against any common believe or wouldn't violate any common sense of consideration. If any.

There wouldn't be anything against tradition or code of ethics.

Chapter 2

Literature Review

2 Literature Review

2.1 Introduction to biometrics

Biometrics is an automated way of identifying a person based on physical or behavioral features. Biometrics include fingerprint recognition, fingerprint geometry, face recognition, iris recognition, and voice recognition, among other things. Biometrics is becoming the foundation of a wide range of highly secure identification and personal verification solutions. The necessity for highly secure identification and personal recognition technology becomes obvious as the number of security breaches and transaction frauds rises. Biometric-based solutions can ensure the secrecy of financial transactions and personal information. Biometrics are required by the federal, state, and municipal governments, as well as commercial uses.

Comparing a registered or enrolled biometric sample to freshly collected biometric samples is required for biometric authentication (for example voice signal or fingerprint captured during Login). A model captures a sample biometric feature during registration or enrolment and stores it for further comparison. Biometrics are also utilized in identification mode, when a biometric system searches a database to identify a person from the total enrolled population. For example, to ensure that a person has not claimed for benefits under two distinct identities, an entire database can be examined. This is referred to as "one-to-many" matching. The biometric system may be utilized in verification mode, where it authenticates users based on previously enrolled patterns. It's known as "one-to-one" matching.

2.2 Speaker recognition

The speaker detection method relied on acoustic aspects of speech that were discovered to vary between people. Acoustic patterns reflect both anatomy (throat and mouth size and shape) and acquired behavioral tendencies (like pitch and speaking style). Text dependent, text independent, and text prompted are the three types of spoken input used by speaker recognition systems. The majority of speaker verification systems rely on text-based input, which entails choosing and enrolling one or more voice passwords. When there is a risk of impostor, text prompted input is employed. HMM (Hidden Markov Model), pattern matching algorithms, Neural Networks, KNN, and decision trees are among the technologies used to process and store voiceprints. The technology can detect human speech patterns. As a result, speech recognition falls within the heading of behavioral biometrics.

A speech is a complicated function of the speaker and his surroundings that a conventional microphone can readily record. Unlike other biometrics technologies such as fingerprints, speech recognition systems do not have any fixed, static, or physical properties. There is only information based on an act in speaker recognition. The most up-to-date method for automatic speaker verification is to create a stochastic model of the speaker based on features collected from the available training audio. We can distinguish between high-level and low-level information in speech recognition.

2.2.1 Speaker verification vs speaker identification

Speaker verification is the use of machine to verify a person's claimed identity from his voice. And in speaker identification, there is no prior identity claim and the system decide who the person is, what group the person is a member of or that the person is unknown. Speaker related differences are the result of combination of anatomical differences inherent in the vocal tract (in throat) and the learned habits of different individuals. In speaker recognition all the differences can be used to discriminate between speakers.

2.2.2 TD-SV vs TI-SV

Speaker verification is the process of confirming a speaker's stated identification based on the speaker's speech signal (voiceprint). Text-Independent Speaker Verification (TI-SV) and Text-Dependent Speaker Verification (TD-SV) are the two types of speaker verification systems (TD-SV). The speaker must recite the registered or supplied password exactly in TD-SV. The method of authenticating the identification of a speaker without regard to the speech content is known as text independent speaker verification. It is more convenient than TD-SV since the user may communicate freely to the system.

There are two primary reasons why a speech recognition system should provide the customer with a new password phrase for each new test occurrence.

1. The user is not required to remember a set password.
2. Replaying recordings of the user's voice is not an easy way to circumvent the system.

2.2.3 Generic system flow

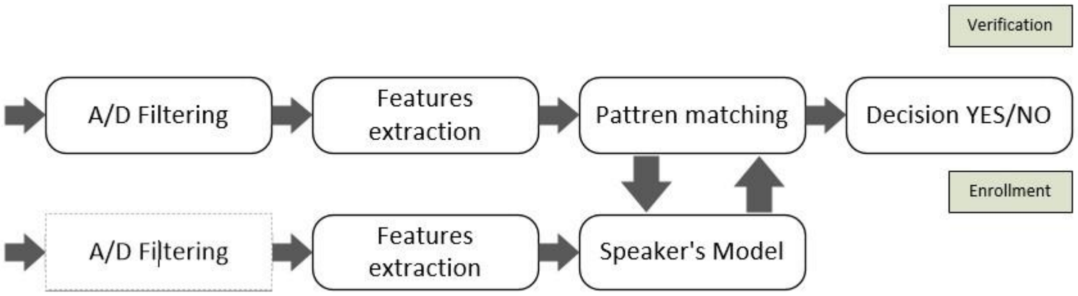


Figure 2.1 Generic system flow

The above figure describes the generic flow of the system, how the system work. For user model training the first step is to remove the noise and distortion from the voice. Second step is to extract features from that voice. Third is to train model on that voice data. Fourth is to make decision either user is enrolled or not.

Now for user testing (login), first step is to remove the noise and distortion from the voice. Second step is to extract features from that voice. Third is to train model on that voice data. Fourth is pattern matching and last is decision making either logged in or not.

2.3 Comparison table

Table 2.1 Comparison table

Functionality	Voice authentication system (Proposed)	LastPass	Jumio Identity verification
Voice Biometrics	✓	✗	✗
Normal Login	✓	✓	✓
Facial Biometrics	✗	✗	✓
Portability	✓	✓	✓
Google Assistant	✗	✗	✓
Customer's Satisfaction	✓	✓	✓

In the above given table there are few app given that were used for secure authentication into your account.

LastPass is a secure password manager that keeps track of all of your usernames and passwords in a vault. Your vault serves as the gateway for all of your data. LastPass remembers your passwords when you save them to your vault. LastPass inputs your login and password for you when you need to log in to a website!

Jumio Identity Verification automates the verification process and helps firms boost conversion rates, comply with AML and KYC laws, and better detect fraud – all while providing a decisive yes/no judgement in seconds. Jumio makes life simpler by incorporating Jumio into the customer verification process, reducing the impact of lengthy verification processes and eliminating costly man hours spent manually receiving and reviewing papers.

2.4 Speaker Modeling

Stochastic and template models are the 2 types of models. The speech production process is treated as a parameter in the stochastic model, which presupposes that the parameters of the underlying stochastic process can be calculated in an exact and well-defined way. The Hidden Markov Model (HMM) is a well-known stochastic model. By preserving a number of sequences of characteristics vectors generated from the same individual uttering the same phrases, the template model attempts to represent the speech production process in a non-parametric manner. The normalized cepstral coefficient is used in the project to employ the template model.

2.4.1 Gaussian Mixture Model (GMM)

Gaussian Mixture Model (GMM) is a frequently used approach in modelling feature in voice recognition task because gaussian models are smooth functions that perform well in modelling real signals. GMM is a combination of gaussian distributions with varying weights, means, and covariance.

The most effective probability function for text-independent speaker recognition, when there is no prior information of what the speaker will say, has been GMMs. Additional temporal information might be useful in text-dependent applications if there is a strong prior understanding of the spoken text. hidden Markov models can be included (HMMs) for the probability functions However, until date, the utilization of more sophisticated likelihood functions, such as those based on probability distributions For text-independent speaker recognition tasks, HMMs have demonstrated no advantage over GMMs.

2.4.2 MFCC features

Mel frequency cepstral coefficients (MFCC) were first proposed for detecting monosyllabic words in continuously uttered phrases. The MFCC computation is a simulation of the human hearing system that aims to artificially recreate the ear's working principle, assuming that the human ear is a trustworthy speaker recognizer. Frequency filters spaced linearly at low frequencies and logarithmically at high frequencies have been employed to keep the phonetically crucial aspects of the speech signal, with MFCC features founded in the observed difference of the human ear's critical bandwidths. Tones of various frequencies are frequent in speech communications; each tone has an actual frequency, f (Hz), and the subjective pitch is calculated using the Mel scale. The frequency spacing on the mel-frequency scale is linear below 1000 Hz and logarithmic above 1000 Hz. 1000 mels is the pitch of a 1 kHz tone at 40 dB over the perceptual hearing threshold, which is used as a reference point.

Chapter 3

System Requirement Specification

3 Requirements

3.1 Functional requirements

The functional requirement are the functionalities of the system and its components. In other words, functional requirements are the services that the software must provide. A functional is just inputs, behavior and output of the system.

3.1.1 Voice input

The very important functional requirement of this system is voice input without voice this system is useless. Because all the other functionalities are dependent on this requirement. That’s why it is more critical.

3.1.2 Feature extraction

The functional requirement of our system is to provide voice that is used to extract feature. And the GMM model will use these features to train model and create .gmm file. Feature is the basic requirement for our system.

3.1.3 Speaker modeling

After feature extraction, the system use the these features to create model on that features system will train the model that will be used for verification of the particular user.

3.1.4 Features matching

After the model training the, when he user test their voice then his/her voice features will be compared with the already trained model and will make a decision.

3.2 Non-functional requirements

It defines the non-functional attributes of system such as performance, security, scalability, maintainability and usability.

3.2.1 Uniqueness

Speaker should train model at least once.

3.2.2 Security

No unauthorized access to the system should be possible. Only the authorized users will be able to login.

Passwords stored in the database should be encrypted form.

3.2.3 User-friendly

The proposed system user interface should be simple, easy to understand and easy to use by any user.

3.2.4 Response Time

The system will provide a reasonable short time response. For example, the speaker should be able to get a response to his login/logout request in 2-3 seconds.

3.2.5 Portability

This app should be executed/run on every web browser with an internet connection.

3.2.6 Scalability

In the case of a lots of user using the web app, the app still performs very well. They handle many users at a time and query speed for data is same in case many users queries on the task at a time.

3.2.7 Robustness

In case user give wrong input or do any unexpected action, the application should be able to recover itself as soon as possible.

3.2.8 Validation

The name, email and password entered by the user in registration, update or login should be validated first by using a front-end programming language before we store them in a Table.

- The user should provide his unique username.
- The user should provide his unique email.
- The user should enter his password.
- The user should enter the password again to confirm that password.

3.2.9 Performance

User perform any task, then they required response should be very fast, so by keeping in the view this demand of user we cope this issue and our query time is less than 1 second.

3.2.10 Efficiency

This application uses minimum resources of your device. The application size is under 50 Mb as the storage used for the storing user data and minimum space is required to install.

3.2.11 Maintainability

Any sort of change in functionalities, design in Application for admin is easy because latest code refactoring and other maintainable technique we use like every module have its own package so easily change any sort of error, improvements, etc.

3.3 System requirements

- Python 3.8
- RAM 8GB
- Processor core i5 at least
- OS windows 10

Chapter 4

Project Design

4 Project Design

4.1 System Architecture

Speaker recognition system is mainly consisting of two major modules speaker enrollment and speaker verification. Both modules have some common process and some their own specific process.

4.1.1 Speaker Enrollment Module

This module extracts specific features from speaker voice and make model for that speaker, save these features into database which are further used for verification module. This module also called training module. MFCC are used for feature extractions.

4.1.2 Speaker verification Module

This module is used for verification of enrolled user. To verify a user based on his voice require that user must be able to must be enrolled with the system. When a user want authentication against his claimed identity his module extracts feature from his voice and store in a vector. These features are matched with already stored features of that speaker and on the basis of that match accept or reject decision is made.

4.2 Agile Development Model

This module is an incremental model as well. The product that will be created will be incremental and in quick cycles. As a consequence, each release is tiny and gradual, building on prior capabilities. Each release is rigorously tested to guarantee consistent product quality. One of the most well-known agile development life cycle methods is extreme programming.

4.2.1 Standards of agile development

- Standard of the agile method
- Client satisfaction and ongoing software development.
- For the client's competitive edge, changing requirements are welcomed.
- Focus on releasing functioning software as often as feasible; delivery priority will be given to the lowest possible time frame.
- Throughout the project, the developer and business personnel must collaborate.
- The greatest approach to share knowledge and establishes a team is through face-to-face contact.

- Working software is the most important indicator of progress.

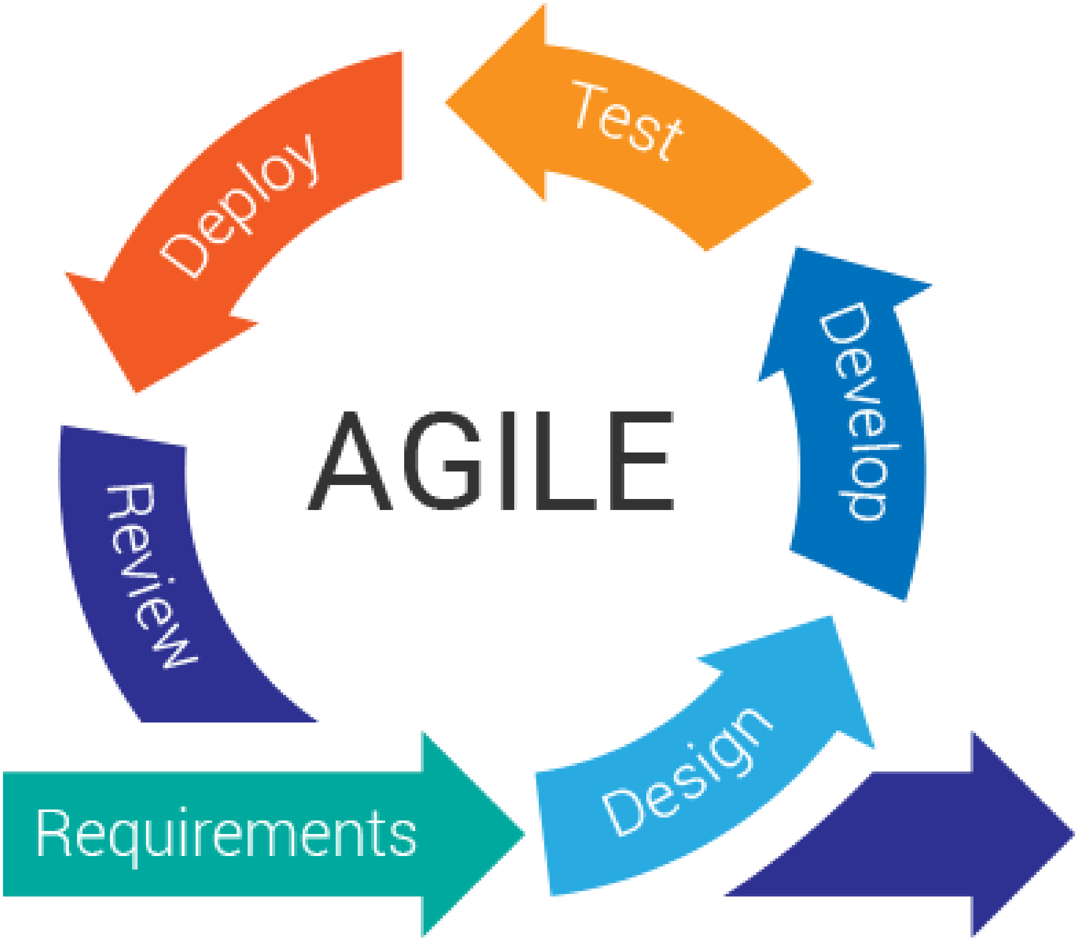


Figure 4.1 Agile Development

This system will be divided into different modules. There are two modules of the system. Enrollment part, and sign in part with voice. We will develop and test each part in different modules. So that’s why we use different sprint to complete this project.

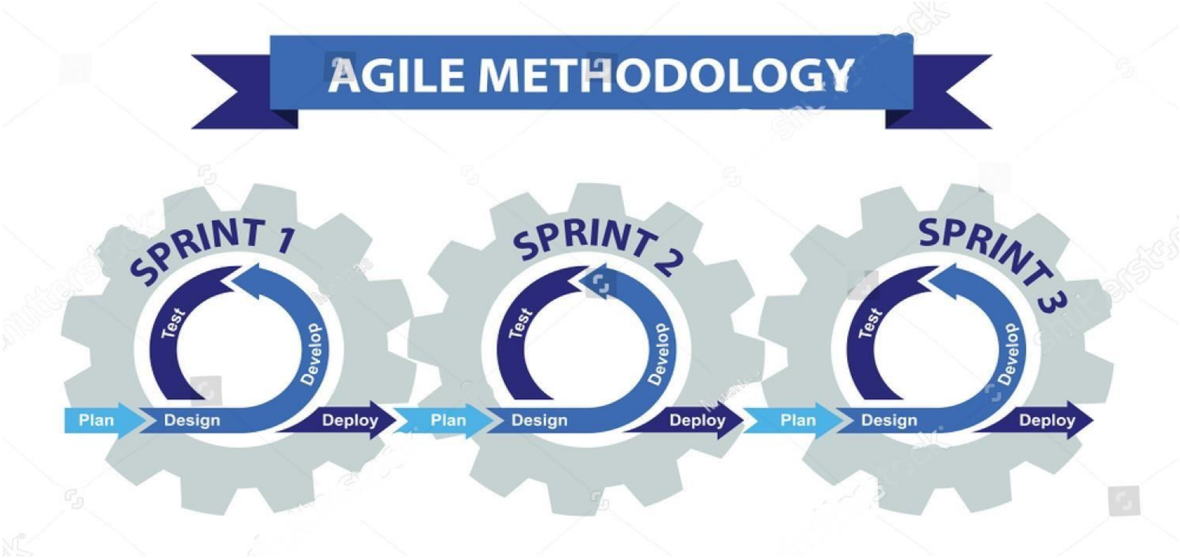


Figure 4.2 Agile Sprints

4.3 Systems control flow

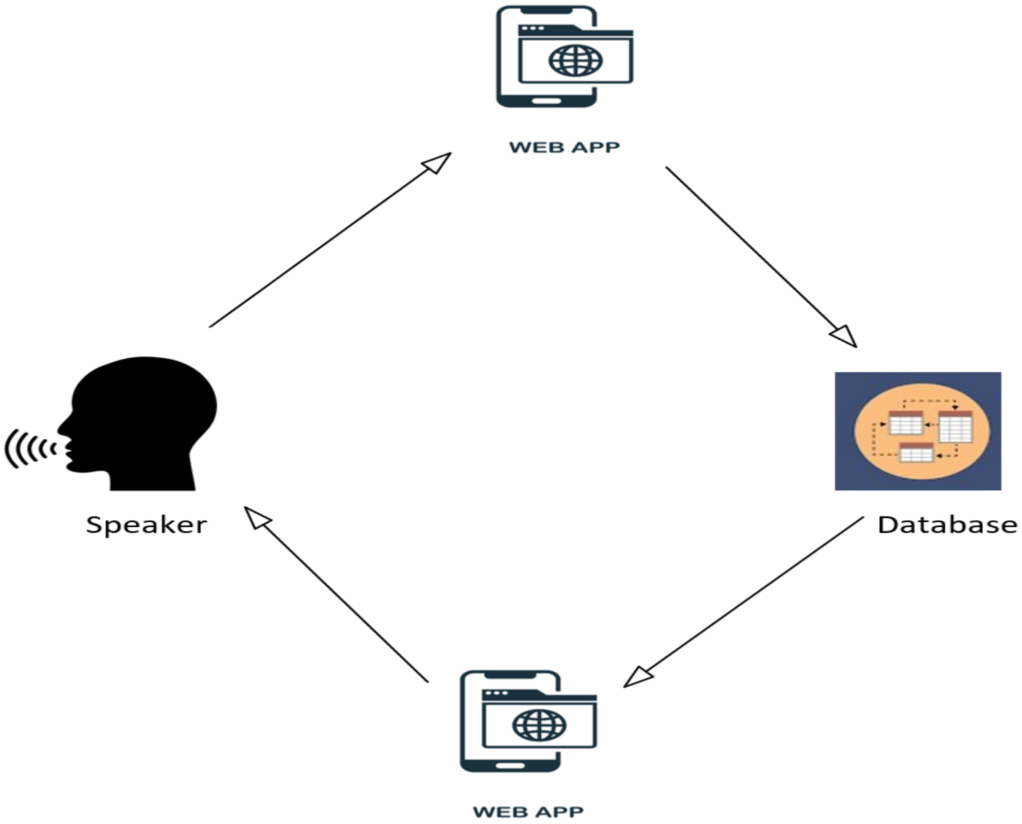


Figure 4.3 System architecture

4.4 Diagrams

4.4.1 Block diagram

The primary geometric shapes utilised in block diagrams were boxes and circles. Blocks linked by straight and segmented-line displaying connections reflect the main parts and functions.

The block diagram for our app for speaker detection is shown below.

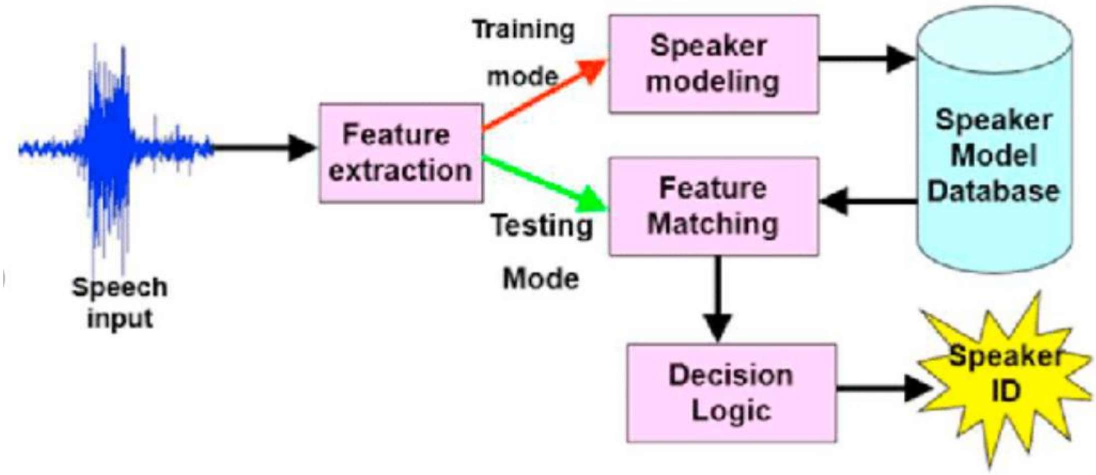


Figure 4.4 Block diagram

Description

1. **Speech input**

   Speaker's voice is very important for our proposed system when speaker came to register with the system by using voice.

2. **Features extraction**

   After recording the voice, the system will extract the features (MFCC features) from the user's voice for two purposes one for training and other for testing purpose. On the basis of these testing and training the use can get into the system.

3. **Speaker modelling**

On the basis of extracted features the system will be trained for that specific user.

4. **Feature matching**

   After training the model the user tests the his\her voice then the extracted features of voice will be compared with the already saved in the database. And then make the decision either to provide access into the system or not.

5. **Decision Logic**

   After testing the model make decision either to provide system authorization or not.

6. **Speaker Model Database**

   In the speaker model database where all the model of the all the user resides after enrolment and will me compared with the new features at the time of login.

4.4.2 Use case diagram

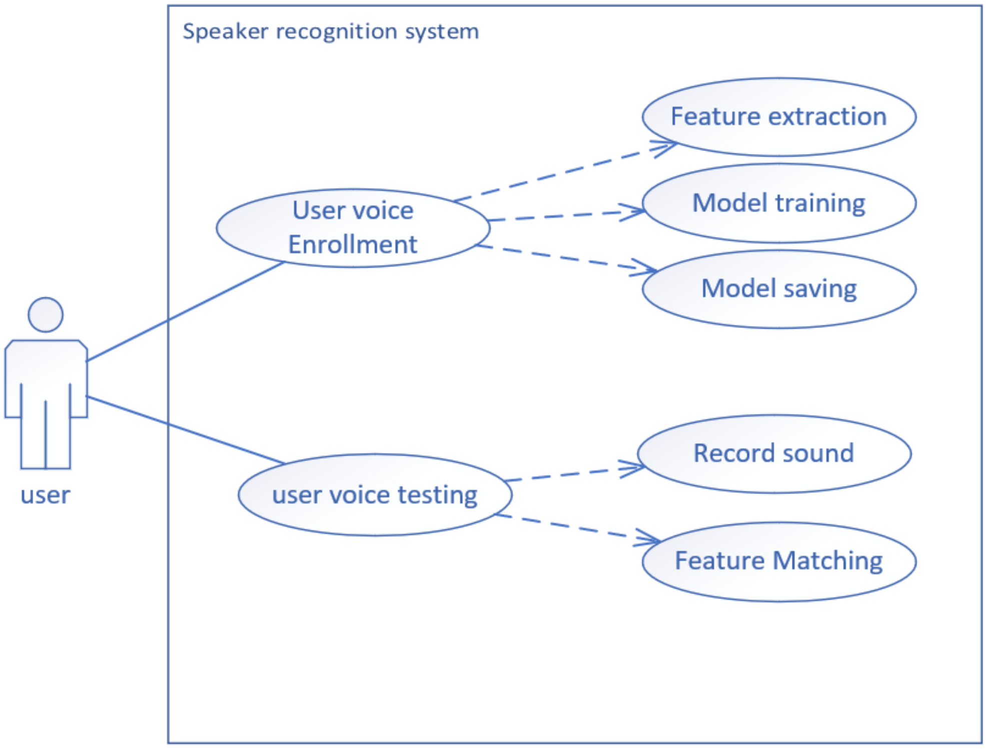


Figure 4.5 Use case diagram

Use case name: speaker Enrollment

Actor: speaker

Pre-condition:

Post-condition: speaker will be enrolled

Details: this use case describe how speaker enroll himself from the system.

Success scenario:

1. The user start enrollment
2. Enter personal detail and voice signal
3. the system extract features
4. store features in database

alternate scenario:

1. username not entered
2. voice not recorded
3. user not exists

Use case name: speaker verification

Actor: speaker

Pre-condition: speaker must be enrolled

Post-condition:

Details: this use case describe how speaker verify himself from the system.

Success scenario:

1. The user start verification
2. the system extract features
3. match these features from stored featured
4. display accept or reject

alternate scenario:

1. username not entered

2. voice not recorded
3. user not exist

4.4.3 Activity diagram

Activity diagram shows the activity that system or user should perform while interacting with the system.

4.4.3.1 Activity diagram for Model Training

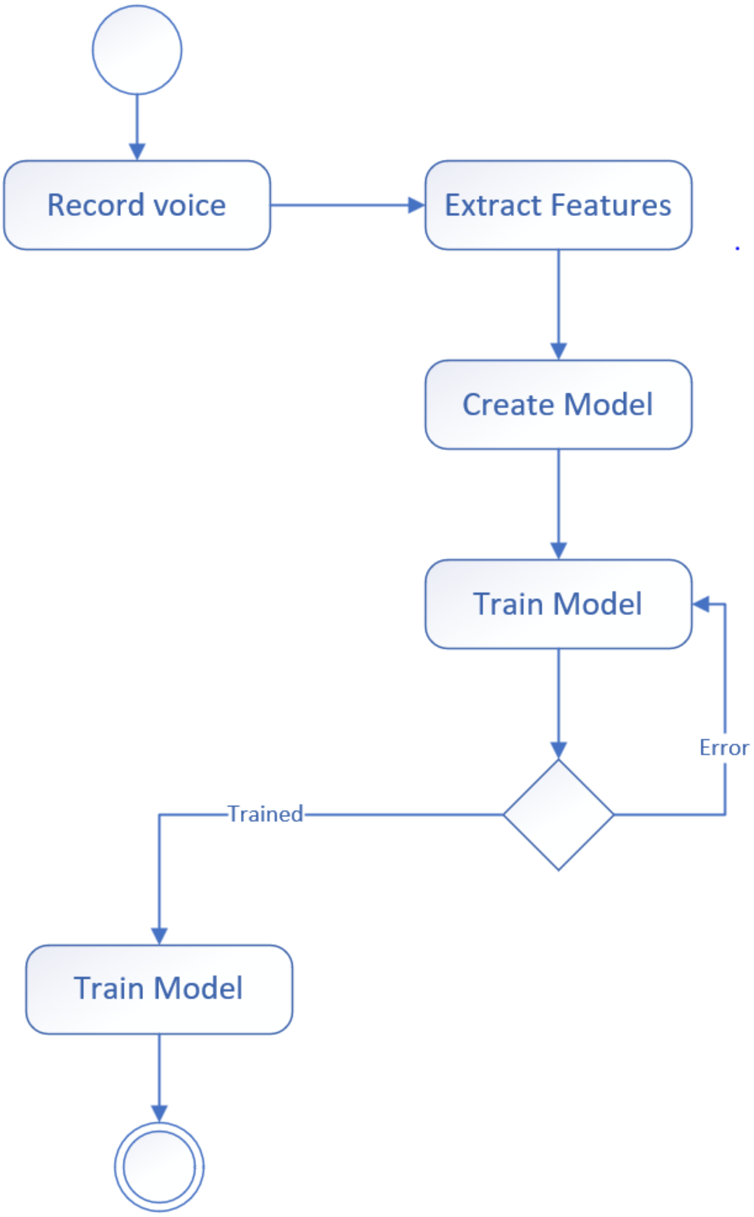


Figure 4.6 Activity diagram for model training

This is the activity diagram for our system while user enrols into the system at very first time. All the above given activities will be performed at these stages to train the model over the new user's voice.

4.4.3.2 Activity diagram for Model Testing

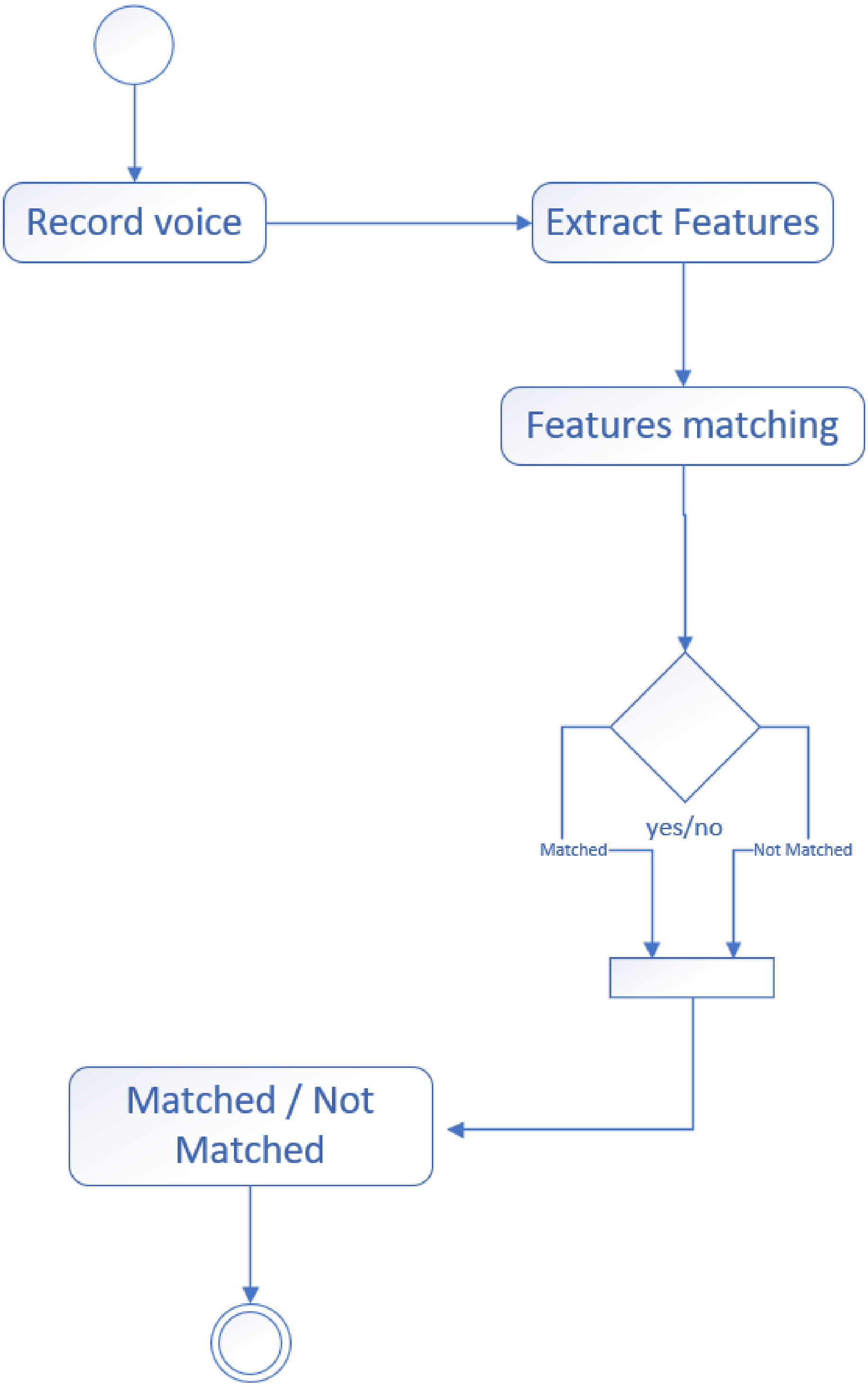


Figure 4.7 Activity diagram for model testing

This is the activity diagram for user testing or when user login into the system. All the above activities should be performed to make the decision either the voice of the user is matched or not.

4.4.4 Sequence diagram

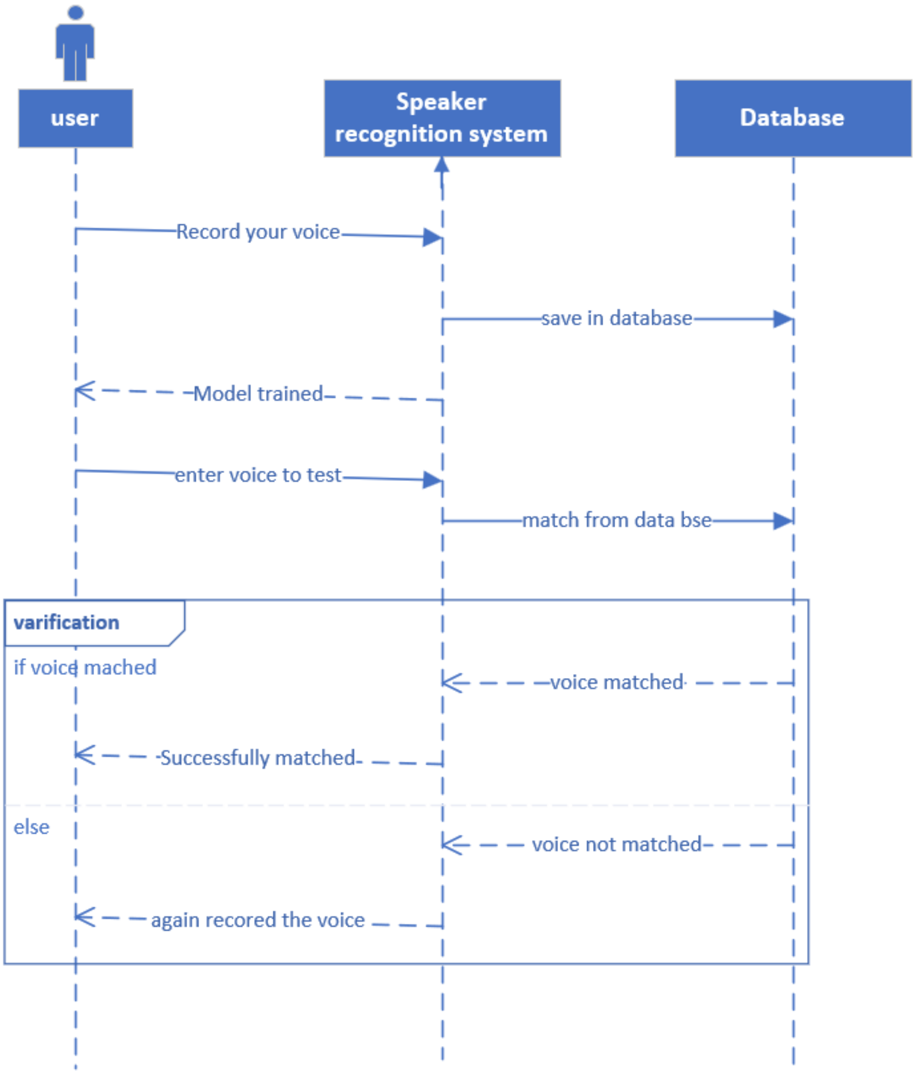


Figure 4.8 Sequence diagram

This is sequence diagram of our system which show the control flow of the system and response to the user, when user interact with the system.

Chapter 5

Implementation

5 Implementation

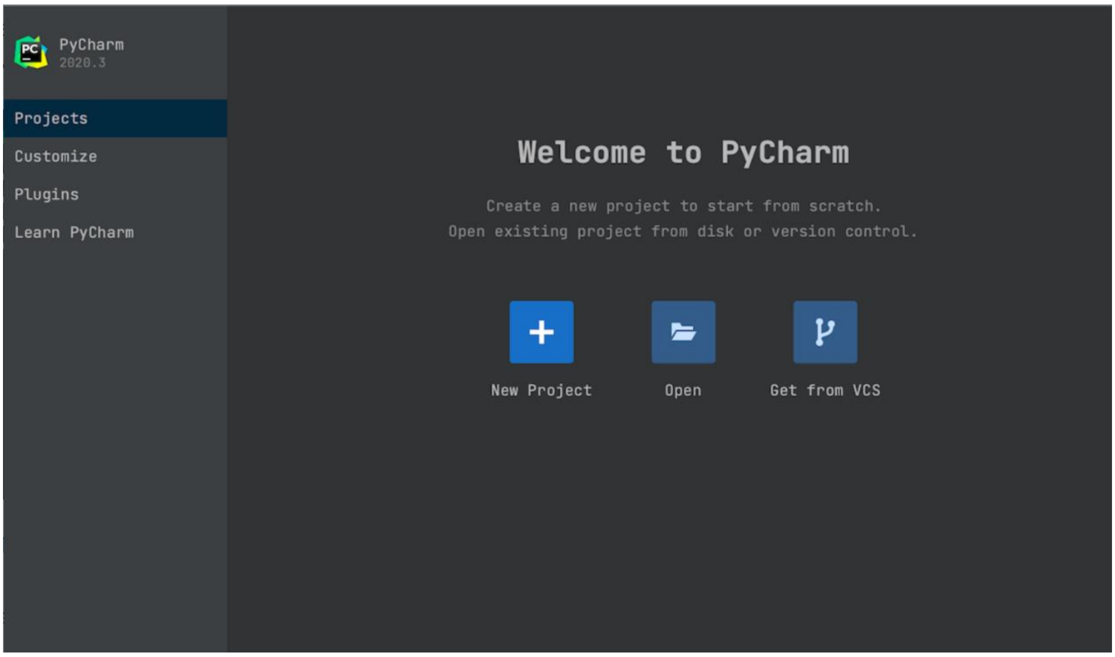
This chapter describe the implementation of the whole application. The implementation is the most important phase of software development life cycle in which we implement the whole idea and requirements. If we are able to implement the functionalities successfully so we will definitely be able to satisfy the users. Each and every functionality as well as implementation described in this chapter.

5.1 Tools & Technologies

Tool and Technologies used in our system are given below:

5.1.1 PyCharm

We are using PyCharm Community Edition. It is cross-platform IDE (integrated development environment) for python. It is free of cost and open source. It is purely for Python. It can be use on Windows, MacOS and Linux.



Welcome to PyCharm

Create a new project to start from scratch.
Open existing project from disk or version control.

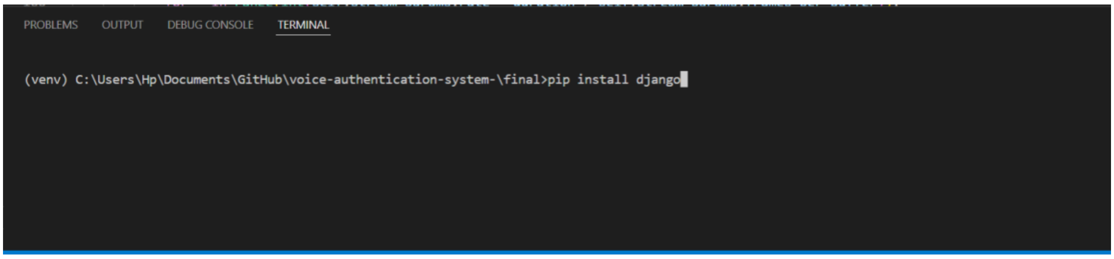
Figure 5.1 PyCharm

Features

1. Intelligent Code Editor
2. Availability of Integration Tools
3. Integrated Debugging and Testing
4. Project and Code Navigation

Django

Django is most famous and richly featured web framework, which is actually based on Python**.** which enable us to source development and keep stable the website. Django is secure scalable and versatile web framework. Django framework installing using command.



```
(venv) C:\Users\Hp\Documents\GitHub\voice-authentication-system-\final>pip install django
```

Figure 5.2 Django installing

5.1.2 Language

- HTML/CSS
- Python

5.2 Development Stages

5.2.1 Pre-Planning

First of all, we must decide ahead of time what application we will create, and answers to questions such as "What is the purpose of this app?" must be provided. Who do you want to reach out to? Which platform should we begin with? Is the software free or does it cost money?

Once we've answered all of these issues, creating a solid design for the voice authentication system application will be a piece of cake. Try to identify features that are missing from competitors' apps so that we may put them in ours to make it unique. Once we have all of the facts, the next step is to figure out how much it will cost, as well as how long it will take.

5.2.2 Assessment of Technical feasibility

To determine technological feasibility, we must examine if the backend technology, such as Firebase, can handle the app's functionality. To determine whether our App's concept is technically feasible, we must first gain access to public data using public

APIs. We must also decide which platform we will develop our app on initially. Depending on the platform, different requirements will apply to developing an app.

5.2.3 Designing and development of application

Before moving to coding part, we are going to designing of user interface it creates interaction with the user and it makes the good look and feel of our e-voting application and it also informs us that how to make a user-friendly application and easy to use the functionalities in our application.

5.2.4 Testing and deployment of application

For user interface testing, we need to target some voter or user to test our application. For this purpose, we give user experience to put our application in the hands of user and after that we do Beta testing. The feedback from those beta users will help to determine that our app functions work well in the real-world Environment. After successful testing the of application, our app is ready to deploy/launch in Online market like Google Play Store.

5.3 User Interface

In our App interface we have two modules which are:

- Enrolment Module
- Login Module

5.3.1 Sign in Page

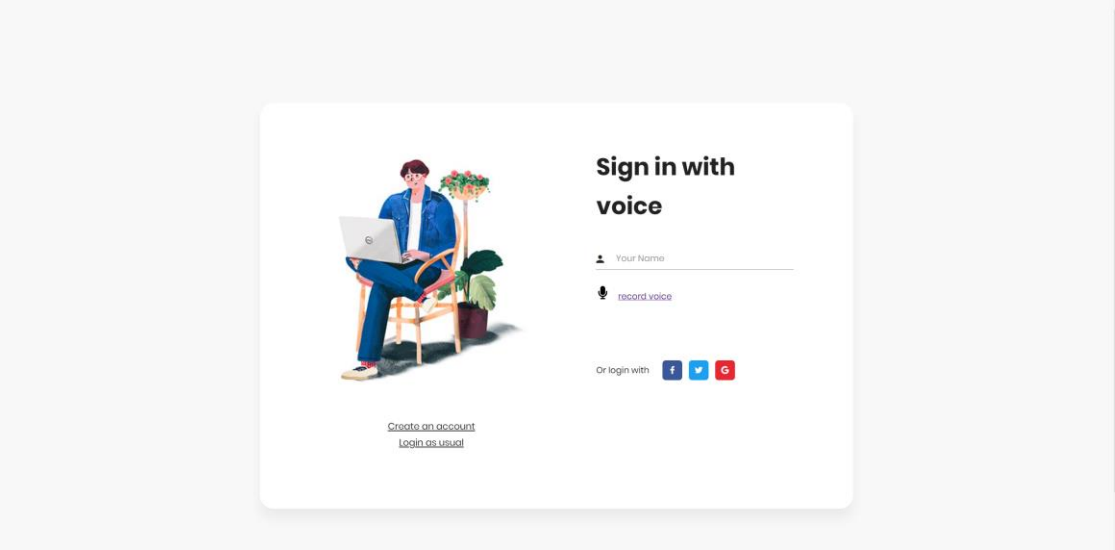


Figure 5.3Sign in page

Figure 5.3 Show sign in screen where user can log in into the system by their name and voice and then then system will verify that voice sample against that name.

5.3.2 Signup Page

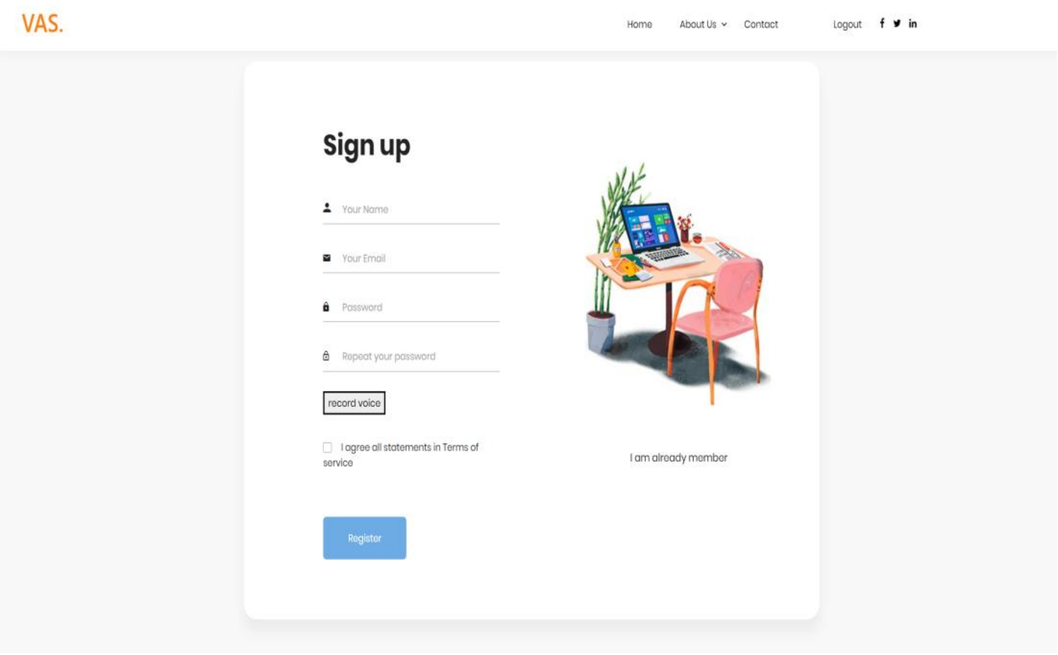
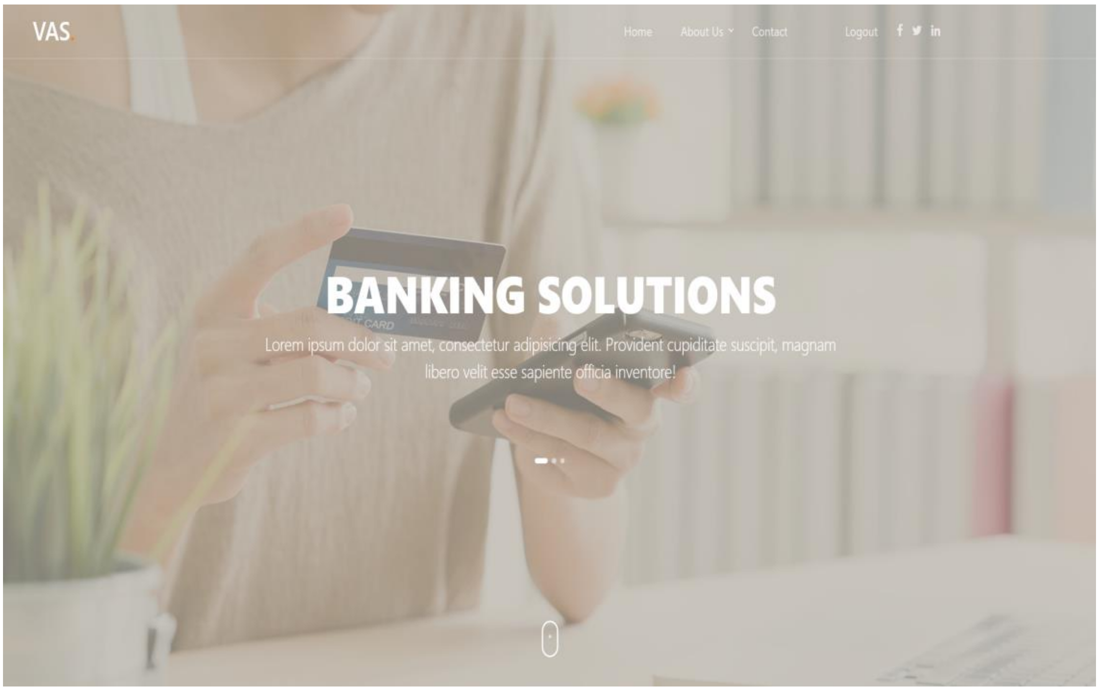


Figure 5.4 Signup Screen

This page will get the user data along with user voice record for the purpose of model training.

5.3.3 Home page



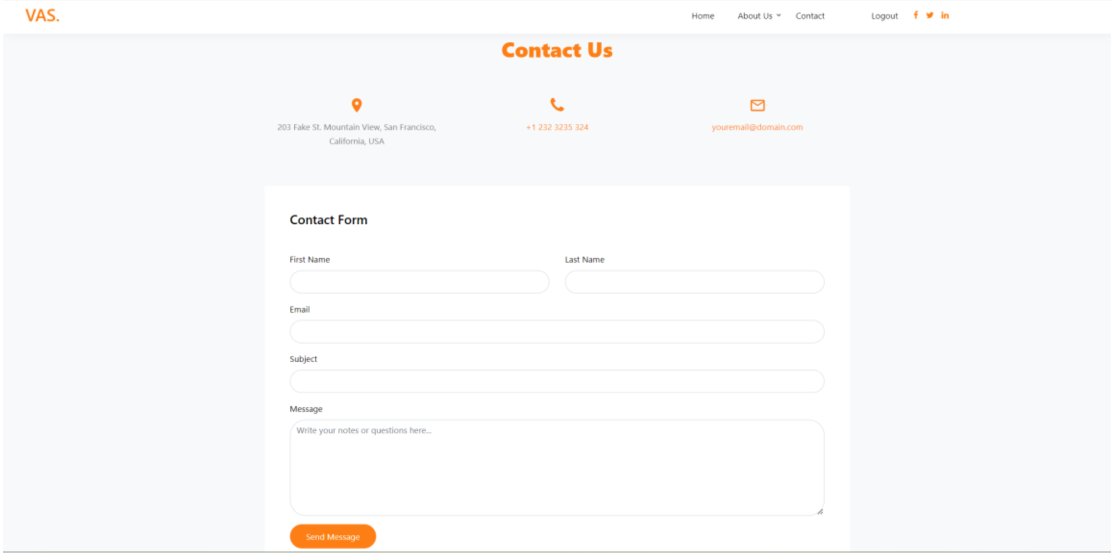
BANKING SOLUTIONS

Lorem ipsum dolor sit amet, consectetur adipisicing elit. Provident cupiditate suscipit, magnam libero velit esse sapiente officia inventore!

Figure 5.5 Home Screen

Figure 5.3 Shows the home page of our website VAS (Voice Authentication System) which will be displayed after successful login.

5.3.4 Contact us Page



Contact Us

203 Fake St. Mountain View, San Francisco, California, USA

+1 232 3235 324

youremail@domain.com

Contact Form

First Name

Last Name

Email

Subject

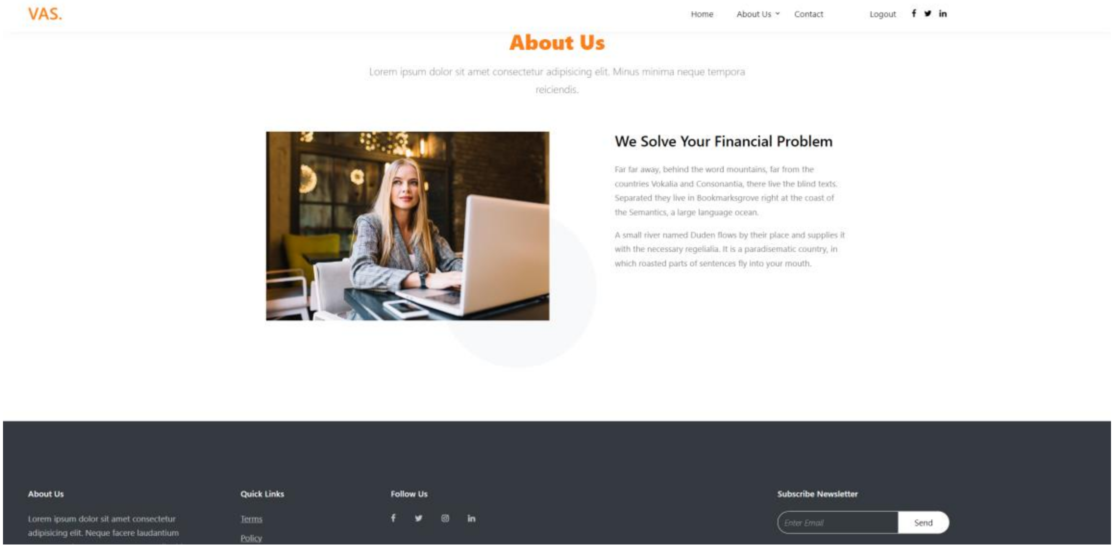
Message

Write your notes or questions here...

Figure 5.6 Contact us page

Figure 5.4 shows the contact us of the system that will used to connect with admin.

5.3.5 About us Page



About Us

Lorem ipsum dolor sit amet consectetur adipisicing elit. Minus minima neque tempora reiciendis.

We Solve Your Financial Problem

Far far away, behind the word mountains, far from the countries Vokalia and Consonantia, there live the blind texts. Separated they live in Bookmarksgrove right at the coast of the Semantics, a large language ocean.

A small river named Duden flows by their place and supplies it with the necessary regelialia. It is a paradisematic country, in which roasted parts of sentences fly into your mouth.

About Us

Lorem ipsum dolor sit amet consectetur adipisicing elit. Neque facere laudantium

Quick Links

Terms

Policy

Follow Us

Subscribe Newsletter

Figure 5.7 About us page

Chapter 6

Evaluations

6 Evaluations

6.1 Test

Testing is the process of measuring the stability of our system stability by focusing on bugs, errors and their occurrence. Visual studio provides multiple ways to handle the exceptions. After the completion of project, it will be provided for testing purposes.

6.1.1 Unit Testing

The “Speaker recognition system” has two modules. And testing is applied on each individual module at the time of development because both modules are connected with each other. Testing applied on each module to check its functionalities or not.

6.1.2 Function Testing

The complete App is passed through a process to check its working or not. We not only work on functionalities of the application but also work on user friendly and attractiveness of interface. After integrating the system, testing was done on the functions which the application will perform.

6.2 Objectives

The main objectives of passing the App from long process of testing is to fix the bug and improve the performance of app. “Speaker recognition system mainly used to handle the and improve security”. To improve the quality of system we apply many testing methods and many tests.

6.3 Test cases

Test case contain clearly defined test step for testing a feature an application.

Test case focus on “What to test and How to test”.

6.3.1 Enrolment Test case

Table 6.1 Test Case for voice input -A

Test Case ID: voice input -A	**Test Designed by:** Saad
Test Priority (Low/Medium/High): High	**Test Designed date:** 20-05-2022
Module Name: training	**Test Executed by:** Bilal
Test Title: to enroll into system	**Test Execution date:** 20-05-2022
Description: Test the successful voice training process	

Pre-conditions: User must have a mic and device to record voice.

Step	Test Steps	Test Data	Expected Result	Actual Result	Status (P/F)
1	Click on record voice button		System show timer to record voice	System displayed a timer to record voice	Pass
2	Provide clear user voice	Clear voice	User Successfully entered the voice	The record the voice successfully.	Pass
3	Click on start training Button	Recorded voice	The system will start to train model on that voice.	The display menu to test voice.	Pass

Post-conditions:

The model successfully trained and features store in database.

Table 6.2 Test Case for voice input -B

Test Case ID: voice input -B	**Test Designed by:** Saad
Test Priority (Low/Medium/High): High	**Test Designed date:** 20-05-2022
Module Name: training	**Test Executed by:** Bilal
Test Title: to enroll into system	**Test Execution date:** 20-05-2022
Description: Test the successful voice training process	

Pre-conditions: model must be trained on the that voice.

Step	**Test Steps**	**Test Data**	**Expected Result**	**Actual Result**	**Status (P/F)**
1	Click on record voice button		System show timer to record voice	System displayed a timer to record voice	Pass
2	Provide clear user voice	Distorted voice	Error message shown because voice is so much noisy and distorted	The user entered clear voice.	Fail
3	Click on record voice button	Noisy voice	The system displays an error massage "**enter clear voice**".	The system displays the page to record again	Fail

Post-conditions:

The model is not trained.

6.3.2 Login Test case

Table 6.3 Login voice input -A

Test Case ID: Login voice input -A	**Test Designed by:** Saad
Test Priority (Low/Medium/High): High	**Test Designed date:** 20-05-2022
Module Name: test	**Test Executed by:** Bilal
Test Title: to login into system	**Test Execution date:** 20-05-2022
Description: Test the successful voice testing process	

Pre-conditions: User must have already registered.

Step	Test Steps	Test Data	Expected Result	Actual Result	Status (P/F)
1	Click on record voice button		System show timer to record voice	System displayed a timer to record voice	Pass
2	Provide clear user voice	Clear voice	User Successfully entered the voice	The record the voice successfully.	Pass
3	Click on start training Button	Recorded voice	The system will start to train model on that voice.	The display menu to test voice.	Pass

Post-conditions:

The user successfully logged into the system.

Table 6.4 Login voice input -B

Test Case ID: Login voice input -B	**Test Designed by:** Saad
Test Priority (Low/Medium/High): High	**Test Designed date:** 20-05-2022
Module Name: test	**Test Executed by:** Bilal
Test Title: Login into system	**Test Execution date:** 20-05-2022
Description: Test the successful voice testing process	

Pre-conditions: model must be trained on the that voice.

Step	Test Steps	Test Data	Expected Result	Actual Result	Status (P/F)
1	Click on record voice button		System show timer to record voice	System displayed a timer to record voice	Pass
2	Provide clear user voice	Distorted voice	Error message shown because voice is so much noisy and distorted	The user entered clear voice.	Fail
3	Click on record voice button	Noisy voice	The system displays an error massage “**enter clear voice**”.	The system displays the page to record again	Fail

Post-conditions:

The user voice mis matched.

Chapter 7

Conclusion and Future Work

7 Conclusion and future work

7.1 Conclusion

The voice-based security system can play vital role in the security of personal or organizational systems. The given system is developed to provide the authentication login system to the employees of the organization (i.e., Bank). To provide the touchless experience to the employees which can be very useful and hygienic to due to current diseases (i.e., Corona) and in normal situations this can be use as security layer with traditional login procedures.

7.2 Future work

For the future work, this system could be developed for different environment. Caller voice recognition for the organizations suspect detection for crime control purpose etc. And currently it works well in controlled environment, Later, it can be enhanced using large number of dataset and samples.

- Further you can make it text dependent voice recognition
- And also, we can make an android app.

Appendix A

References

[1] n. kodaagoda, “voislock-Human voice authentication using Hidden Markove Model,” dec, 2008.

[2] s. b. w. a. frederick t.sheldon, “voice biometriuc identity authentication for Iot devices,” 2020.

[3] R. Tanwar, “an approach to ensure secuirty using voice authentication system,” *2019.*

[4] P. Dhakal, “A Near Realtime Automatic Speaker Recognition Architecture for voice based user interface,” 2019.

[5] b. p. a. k. Atul anand, “Speaker Identification over Call Record,” 2018.

[6] V. Bhapkar, “Speaker Identification using Machine Learning,” medium, [Online]. Available: https://medium.com/analytics-vidhya/speaker-identification-using-machine-learning-3080ee202920. [Accessed 04 April 2022].

[7] A. Subramanian, “Building a Speaker Recognition Model,” 17 July 2021. [Online]. Available: https://blog.dominodatalab.com/building-a-speaker-recognition-model.

[8] M. Fabien, “Basics of Speaker Verification,” 02 april 2020. [Online]. Available: https://maelfabien.github.io/machinelearning/basics_speech/#overview-of-speaker-verification-pipeline. [Accessed 23 MAY 2022].