**COMSATS University Islamabad (CUI)**



**Department of Computer Science**

**Assignment-03**

CLO-3

**Software Design Description**

**(SDS DOCUMENT)**

**for**

**Speech2face**

Version 1.0

**Submitted By:**

**Shahzaneer Ahmed SP21-BCS-087**

**Shayan Zameer SP21-BCS-088**

**BCS-4-AB**

**Supervisor**

**Mr. Tehseen Riaz Abbasi**

Submission Date: (2nd December - 2022)

Original Version 1.0

Bachelor of Science in Computer Science (2020-2025)

***Software Design Description for <Speech2face> Page ii***

Contents

[1. Introduction 3](#_Toc37345)

[1.1 Scope 3](#_Toc37346)

[1.2 Modules 3](#_Toc37347)

[1.2.1 Module 1: Profile Management 3](#_Toc37348)

[1.2.2 Module 2: Place Voice Record 4](#_Toc37349)

1.2.3 Module 3: Sound to Vector Model .......................................................................................... 1 1.2.4 Module 4: Vector to Image Model .......................................................................................... 1

1.2.5 Module 5: Image View Customization ...................................................................................... 1

1.2.6 Module 6: Features Enhancer .................................................................................................. 1

1.2.7 Module 7: Insight Panel .......................................................................................................... 2

1.2.8 Module 8: Feedback Panel ...................................................................................................... 2 1.2.9 Module 9: Help and Support ................................................................................................... 2

1. **Design methodology and Software Process Model ...............................................................3**
   1. Design Methodology ................................................................................................................... 3 2.2 Software Process Model .............................................................................................................. 3
2. **System overview .......................................................................................................................4**  ............................................... 4
   1. System Architecture .................................................................................................................... 4 3.2 System Architecture .................................................................................................................... 6

#### 3.3 Process flow/Representation........................................................................................................ 7 4. Design Models.........................................................................................................................15

4.1 Class Diagram ........................................................................................................................... 15 4.2 Sequence Diagram ..................................................................................................................... 16

1. **Data design .............................................................................................................................24**
   1. Relational Database Model ........................................................................................................ 24
   2. Data dictionary .......................................................................................................................... 24
2. **Algorithm & Implementation ...............................................................................................25**
   1. Sign up and Login ..................................................................................................................... 25
   2. Adding Voice Record ................................................................................................................ 25
   3. Voice to Vector Modeling ......................................................................................................... 25 6.4 Vector to Image Modeling ......................................................................................................... 26
   4. Image View Customization ....................................................................................................... 26
   5. Features Enhancer ..................................................................................................................... 26 6.7 Insight Panel .............................................................................................................................. 26
3. **Human Interface Design........................................................................................................28** 7.1. Screen images ............................................................................................................................ 28

7.2. Screen objects and actions ......................................................................................................... 37

1. **Conclusion ..............................................................................................................................37**
2. **References ...............................................................................................................................37** 9.1 Related System: ......................................................................................................................... 37
   * + https://www.github.com/topics/sound-classification. ........................................................... 37
     + https://speech2face.github.io/ ................................................................................................ 37 • Google Facenet ...................................................................................................................... 37

9.2 YouTube Resources:.................................................................................................................. 37 https://www.youtube.com/watch?v=aKYlSIs3UDY&t=334s ........................................................... 37

1. **Plagiarism Report ..................................................................................................................38** **11.** **SDS Work Division ................................................................................................................38**

**ii**

***Software Design Description for <Speech2face> Page 3***

# 1. Introduction

Speech2face is a web and mobile application-based software that is mainly built to recognize the face general structure, ethnicity and gender with the audio waves. Currently there is

no full fledge software in the world that helps in this regard although there are deep learning libraries on which much work is done in the past. It was nearly impossible to predict the appearance of some person with his/her voice. But Artificial Intelligence has made it.

It will automate the attendance system and reinvigorate the authentication system. Moreover, it will prove itself as an asset to the security and intelligence agencies by recognizing criminals faces with the voice notes and eventually help them resolving complex cases.

Speech2face can also be used as a general-purpose software for recognizing the individuals in old audios and images where their appearance is not clear.

## 1.1 Scope

Speech2Face will be a Research and development-based Product with the main functionality to convert the voice into vector form and the vector form to image form and thus assisting in providing insights about the details of the person whose voice is under observation. It will be developed using different technologies and Deep Learning and machine learning techniques will be used. The Image generated can be modified in accordance with the users’ perspective.

***General user***: They can retrieve their images from their voice notes.

***Security Agencies***: It Can help them identify the criminals and speed up the process of resolving complex cases.

## 1.2 Modules

### 1.2.1 Module 1: Profile Management

This module focuses on account creation and management of Users and Security Person Accounts. After signing into their accounts, an individual user can toggle portals. Each portal will manage its respective domain.

1. **Sign Up**

The users will have to sign up for a new account to use the application.

1. **Log In**

The users will log in to their accounts when they want to use it.

1. **Sign in Via Phone**

The Users will log in using their phone number.

1. **Sign in via Guest**

The users will login as a Guest.

1. **Sign in Via Voice**

The User will login by using their voice.

1. **Update Profile**

The users will be able to view and edit their personal information saved in the application.

1. **Logout**

The users can logout from their accounts in the application.

1. **Delete Profile**

The users will be able to delete profile that is saved in application.

### 1.2.2 Module 2: Place Voice Record

This module focuses on Adding voice to the system and later the voice will be be converted to image.

1. **Record Voice**

The users can record voice by using mic in system.

1. **Upload Existing Voice**

The User can upload previous voice that is saved in system.

1. **Upload Existing Video**

The users can Upload Existing video that is saved in system. The voice can be fetch from video and late converted to image.

1. **Update Voice**

The users can Update voices that is saved in system.

1. **Delete Voice**

The users can delete saved voice in the system.

1. **Update Video**

**4**

The users can delete saved videos in the system.

#### 1.2.3 Module 3: Sound to Vector Model

This Module Focuses on how the image will be converted to Vector by using deep learning.

1. **Sound to Vector**

In this use case the sound will be converted to vector.

1. **Generation of Vector Model**

In this use case the vector model will be generated.

#### 1.2.4 Module 4: Vector to Image Model

This Module Focuses on how the Vector will be converted to Image by using deep learning.

**1.Vector to Image Model**

In this use case the Vector will be converted to image.

**2.Generation of Image**

In this use case the Image will be generated.

#### 1.2.5 Module 5: Image View Customization

This Module Focuses on how the image will be customize from user end.

**1.Brightness Control**

In this use case the brightness of image can be adjusted according to user’s need. **2.Saturation Management**

In this use case the Saturation of image can be adjusted according to user’s need.

**3.Skin Color Management**

In this use case the color of skin can be adjusted in image.

**4.Filters**

In this use case User can add filters to image.

#### 1.2.6 Module 6: Features Enhancer

This Module add some special access to Security person to enhance image after generation.

**1.Nose Enhancement**

In this use case the Nose of person can be adjusted.

**2.Eye Enhancement**

In this use case the eyes of person can be adjusted.

**3.Face Shape Enhancement**

In this use case the Shape of person can be Enhanced.

**4.Eyebrow Enhancement**

In this use case the eyebrows of person can be adjusted.

**5.Beard Maker**

In this use case beard can be add to image.

#### 1.2.7 Module 7: Insight Panel

This Module adds access to Security person to check the system.

**1.View Report**

In this use case the User can view report.

**2.Download Report**

In this use case the user can download report.

**3.Share via Socials**

In this use case the user cam share report on social media.

#### 1.2.8 Module 8: Feedback Panel

This Module Specifies some settings and configuration and also related about feedbacks.

**1.Sent Feedback**

In this use case the User can sent feedback to Management team and rate the system. **2.View Feedback history**

In this use case the user can view the history of their feedbacks.

**3.Dark Mode**

In this use case the user can apply dark mode to the system.

**4.Change language**

In this use case the user can change the language of System.

**5.Privacy Policy**

In this use case the user can Privacy Policy of system.

#### 1.2.9 Module 9: Help and Support

This module will cater the application users’ queries related the application itself. Using artificial intelligence and machine learning, a bot will handle the questions from the users. The users may want to engage with an actual person, so the bot can redirect the user to live chat as well. Contact information of support team will also be available on the portal.

**1.Chat with AI Bot**

The bot will handle queries of the users regarding the use of application.

**2.Contact Support Team**

The support team’s contact information will be available.

**3.Change Bot’s Language**

The Bot’s language can be changed by the users.

**4.View Bot’s Query history**

The query history can be vied by user.

1. **Design methodology and Software Process Model** 
   1. **Design Methodology**

The design methodology we will be using is **Object Oriented Approach** because it increases the reusability of the code, and it would be easier for the team members to work together without any confusion. Therefore, we are using Flutter for mobile and HTML, CSS and JavaScript for web-based application which follows object-oriented approach. Also, it is easier to describe the code using UML diagrams. So, OOP is the best approach which fits our Project.

* 1. **Software Process Model**

The software process methodology that we will use is **Incremental Process Model.** Our application has limited number of functionalities and most of the requirements are known since the beginning. There is a minor chance that the requirements would be changed in future. Therefore, the most suitable process model we could select is incremental process model.

1. **System overview**

The System consists of following 9 modules and it aims to convert the speech into image of the user. It allows the feedback and rating to improve its performance.

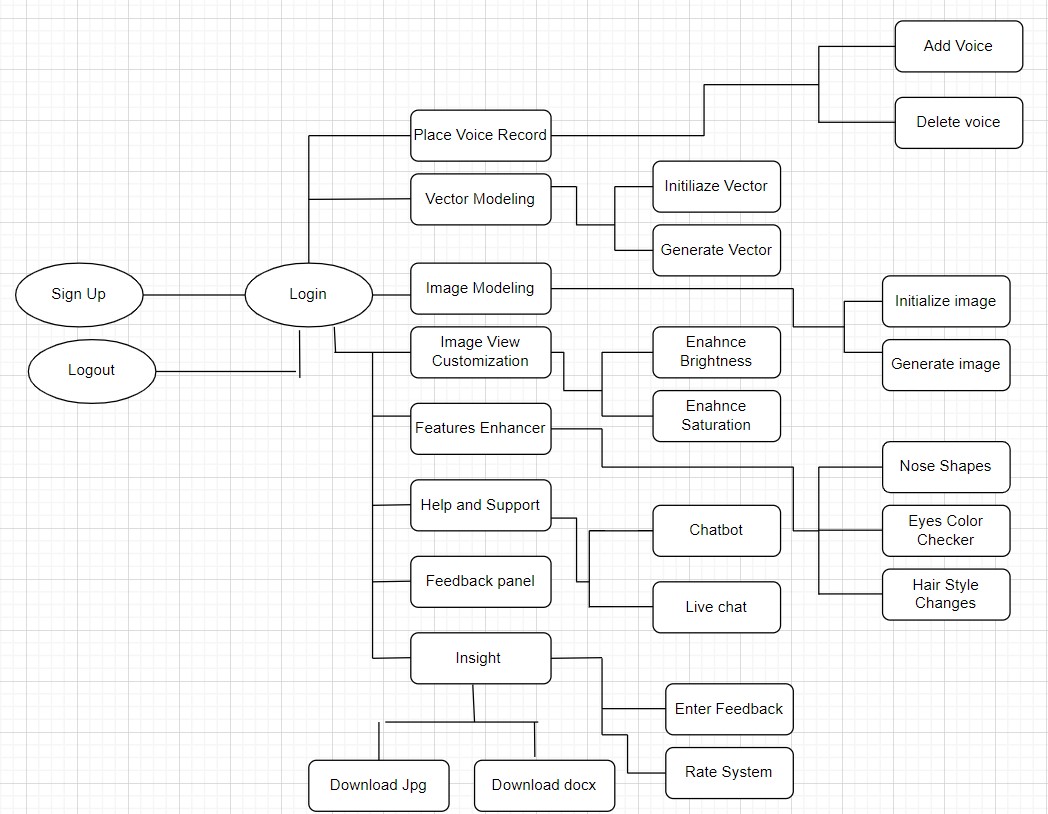


* 1. **System Architecture**

This system will consist of 9 modules further divided into some major components. These will be implemented in the form of classes and will be designed as a functional unit. User is the core components of the system.

All other components are dependent on them. The user and system are responsible for **Profile management**. The **Help and Support** will be dependent on the **FAQ** chatbot bot which will answer the queries of the users. Simple user chat-box and live chat options will also be possible and will depend on the user and the system.

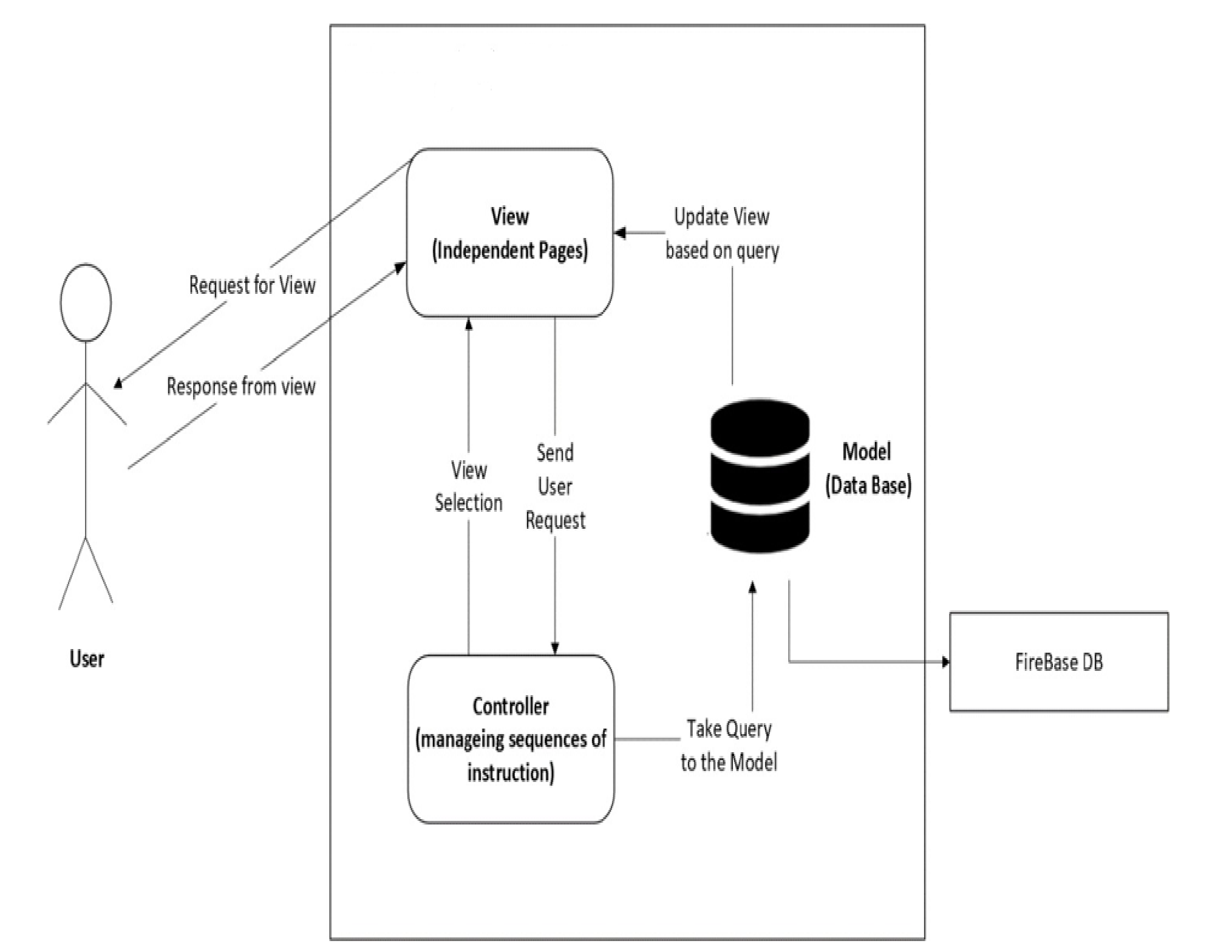
Components will be dependent on users as well like **Speech to face Vector Modelling** and **Vector to Image Modelling**. They will use machine learning and Deep Learning techniques to transform the speech into image. The Generated Image will be modified by **Features Enhancer** and **Image Customization**.



**Block Line Diagram: Speech2Face**

**3.2 System Architecture**

The System has a **Model View Controller (MVC) architecture.** The system will have static pages or menus which will be viewed by the user. All the independent pages of the View will be connected to the Controller of the system which will manage the sequences of instructions added by the user. The Model of the system is the database which will be used to store data. The model will update anything on the view based on the queries. The Controller will take the query to the model in order to fetch data.

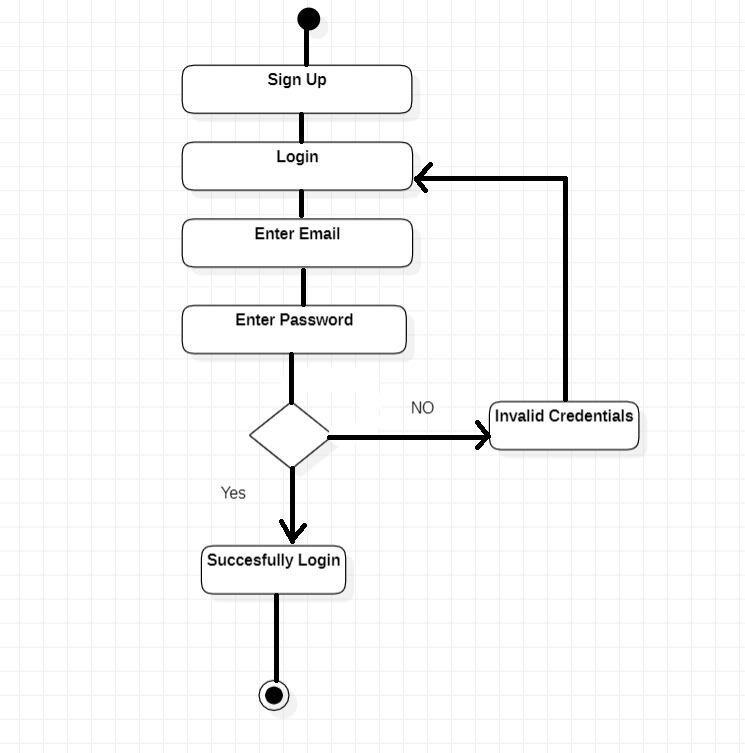


***Figure 2: System Architecture***

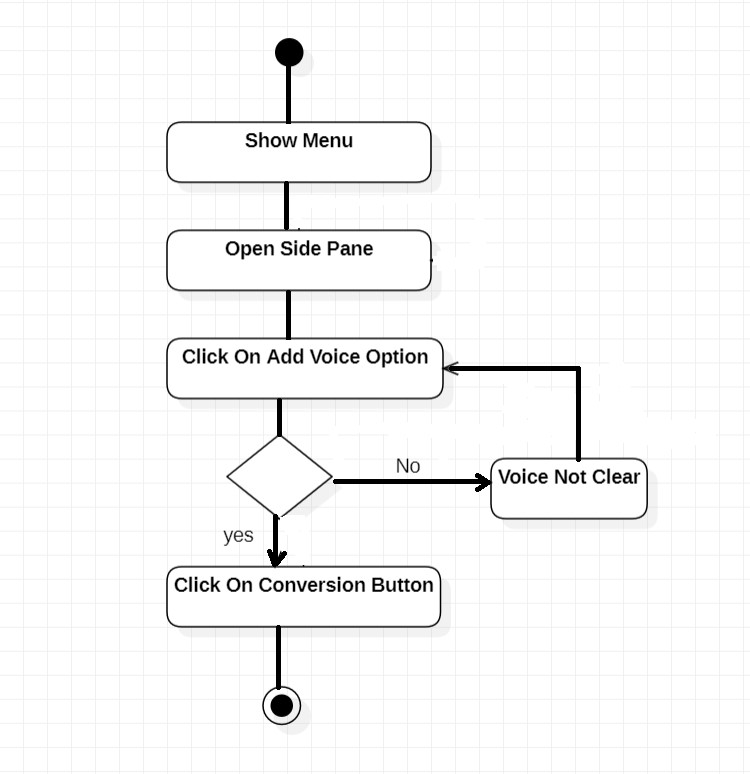
**3.3 Process flow/Representation**

Following are some of the activity Diagrams of the application “**Speech2Face**”

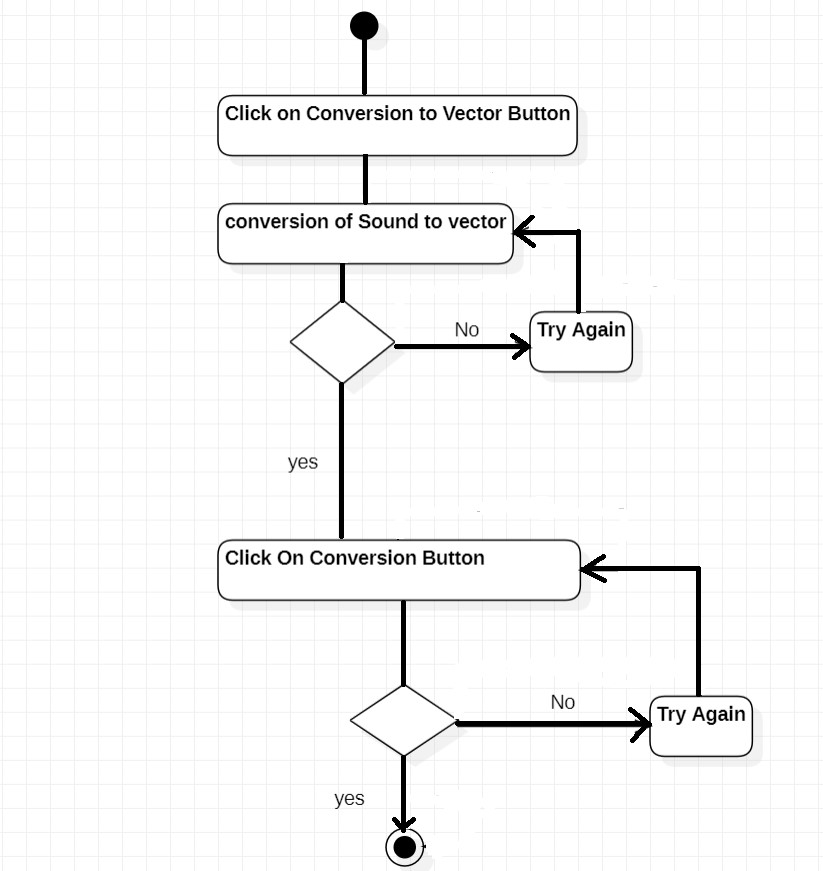
#### Module 1: Profile Management



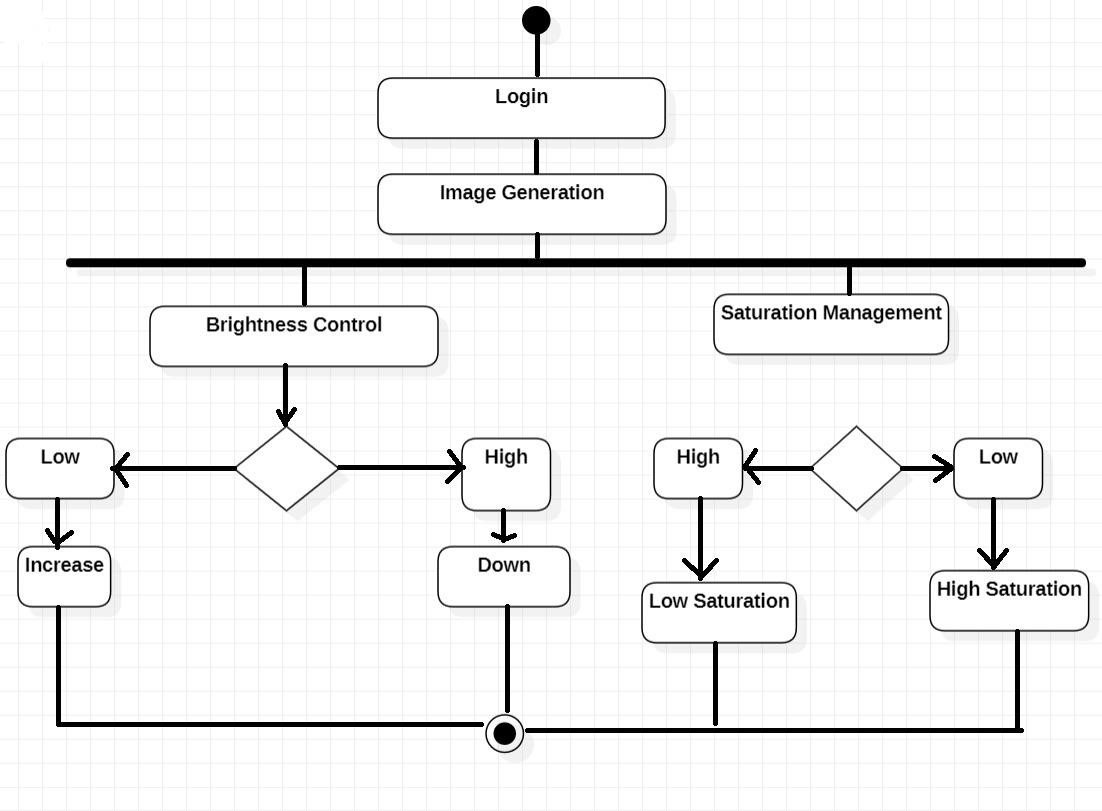
#### Module 2: Place Voice Record



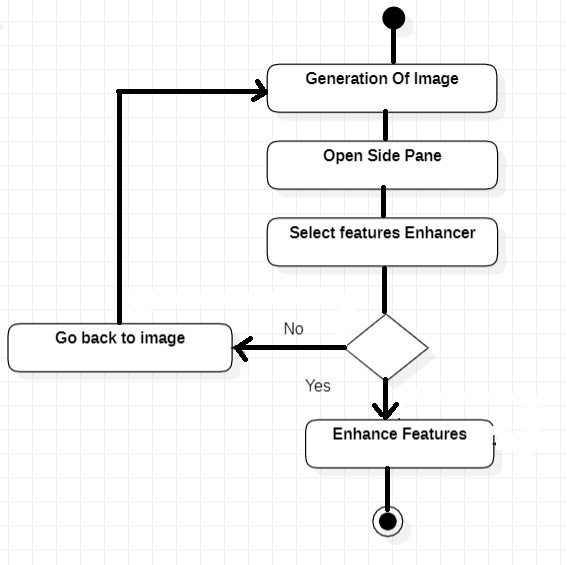
#### Module 3 and 4: Vector and Image Generation Via Voice



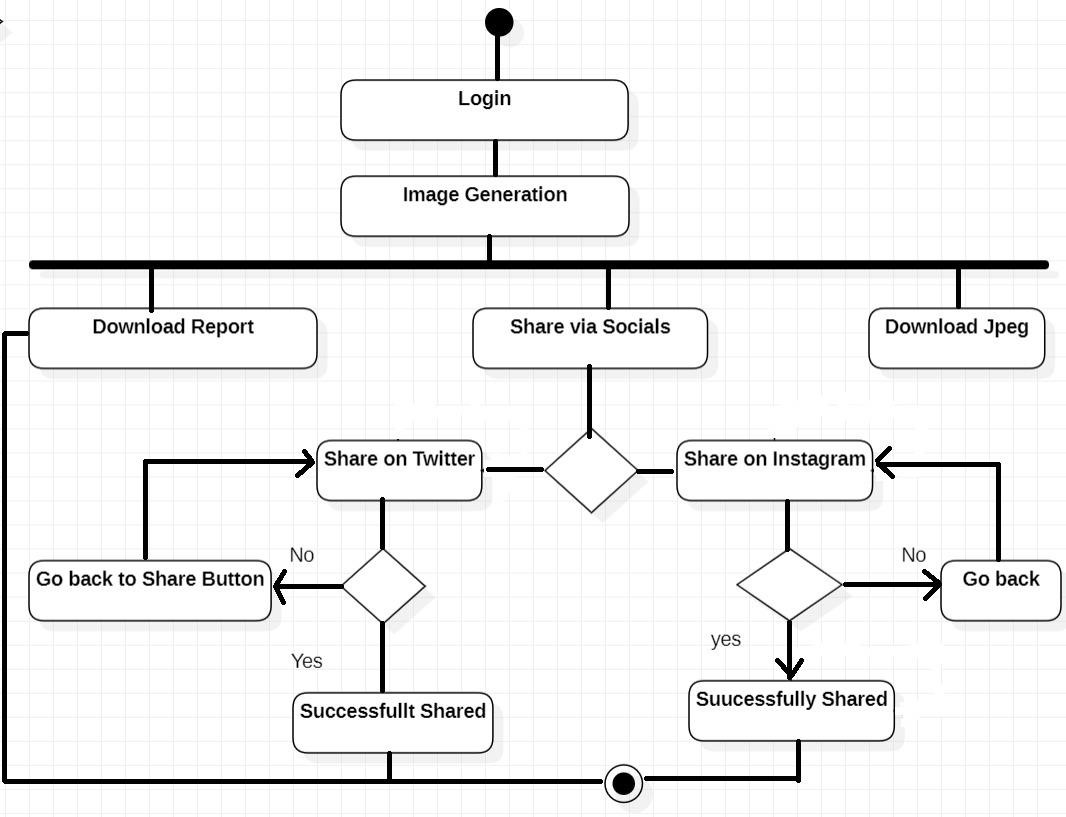
#### Module 5: Image View Customization



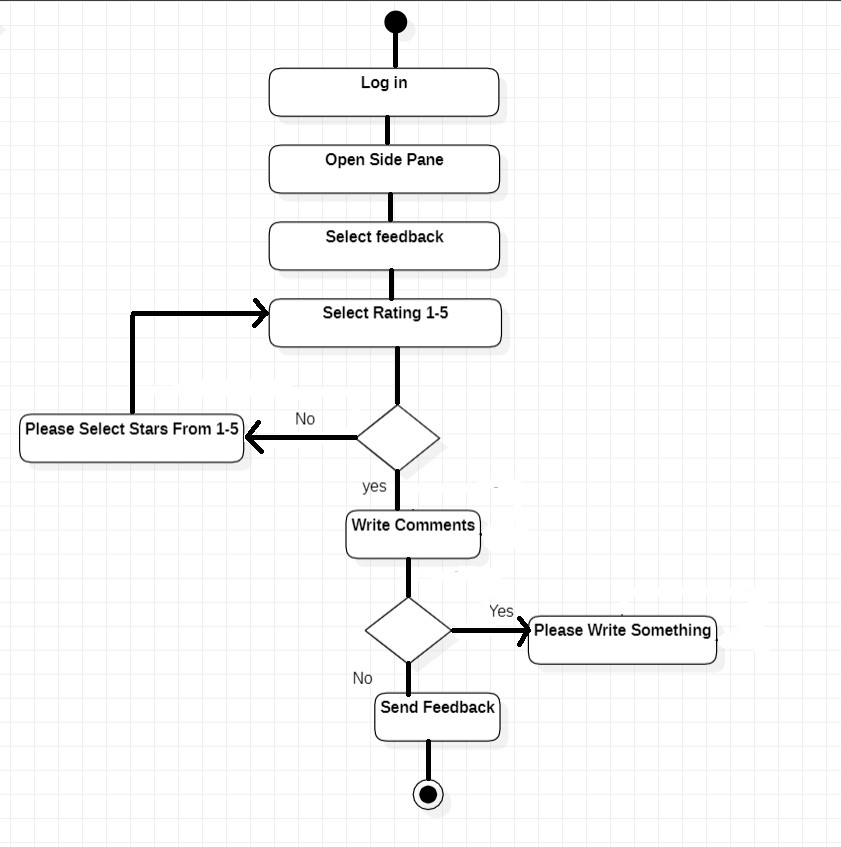
#### Module 6: Features Enhancer



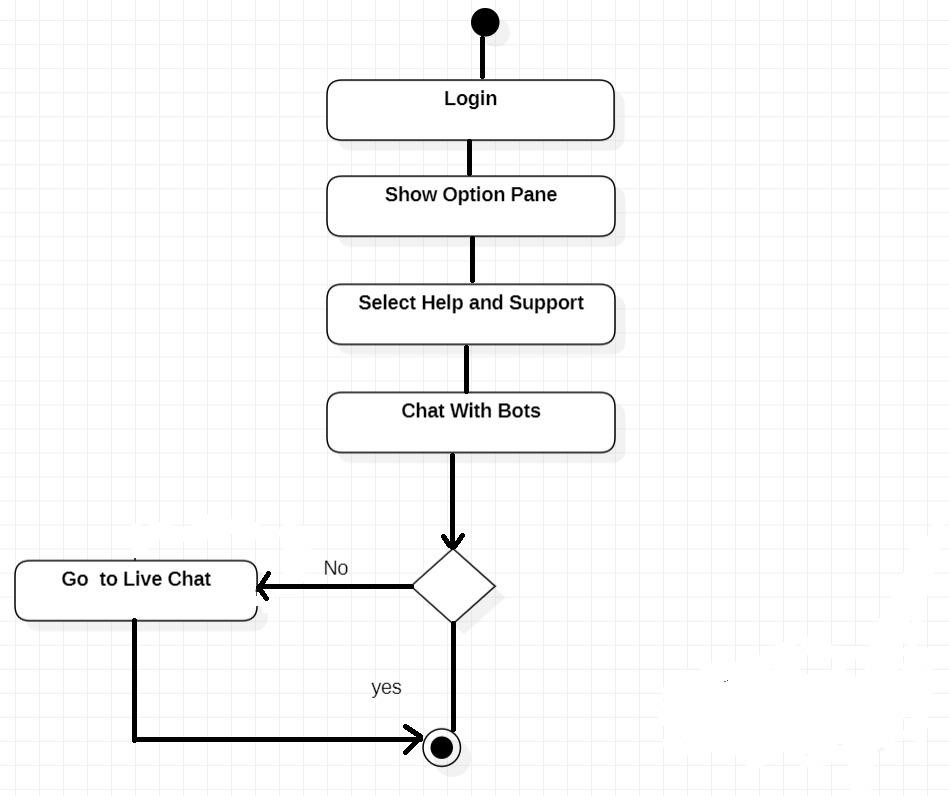
#### Module 7: Insight Module



#### Module 8: Feedback Panel



#### Module 9: Help and Support



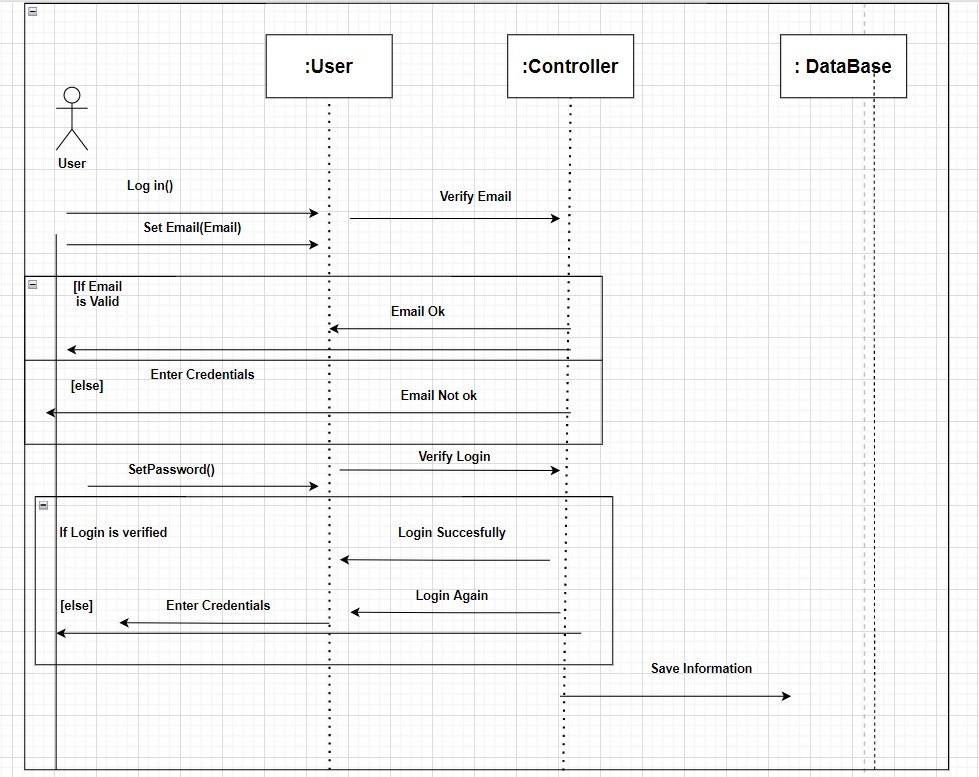
1. **Design Models** 
   1. **Class Diagram**

The class diagram consists of 21 classes which fetch or insert data from database. The proposed system has MVC architecture at application level so separate classes have been developed for model and controller. Following are the classes for Speech2face:

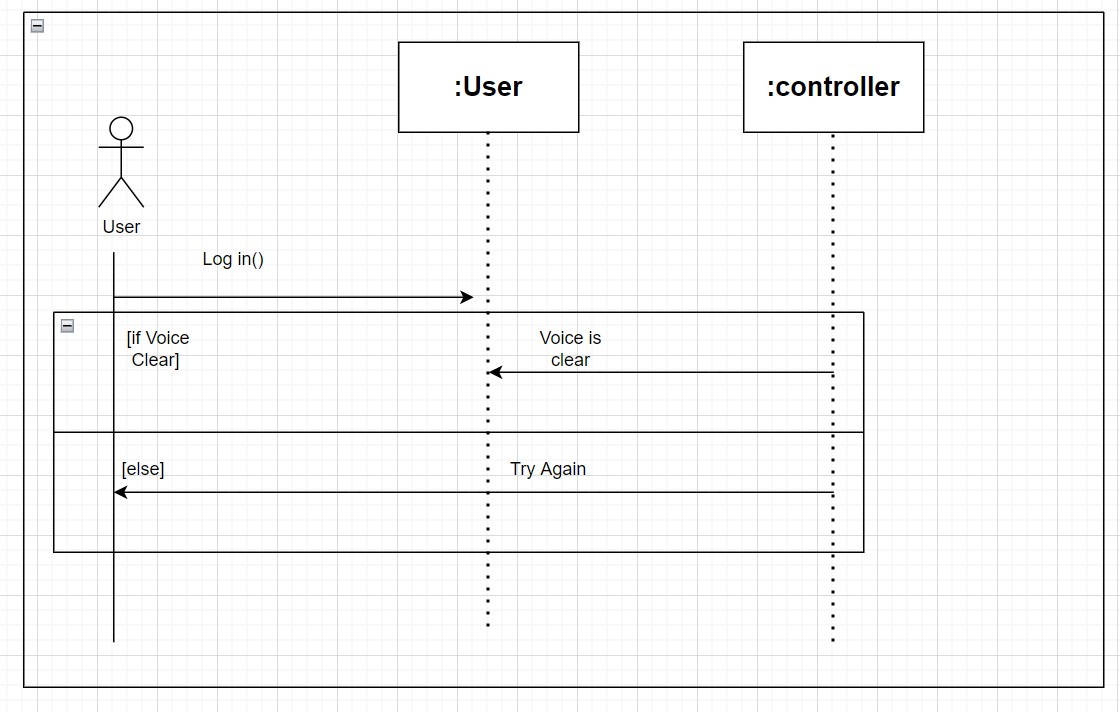
* + 1. User
    2. Security Agency User
    3. User Authentication
    4. Voice Integration Login Model
    5. Speech to Vector Modeler
    6. Vector to Image Modeler
    7. Image Filters
    8. Saturation Controller
    9. Brightness Controller
    10. Hue Controller
    11. Share Via Socials
    12. Nose enhancer
    13. Hair Modeling
    14. Eye Color Modeler
    15. Ads
    16. Chatbot
    17. Chatbot Connection
    18. Notification
    19. Notification Setting
    20. Feedback
    21. About Us

* 1. **Sequence Diagram**

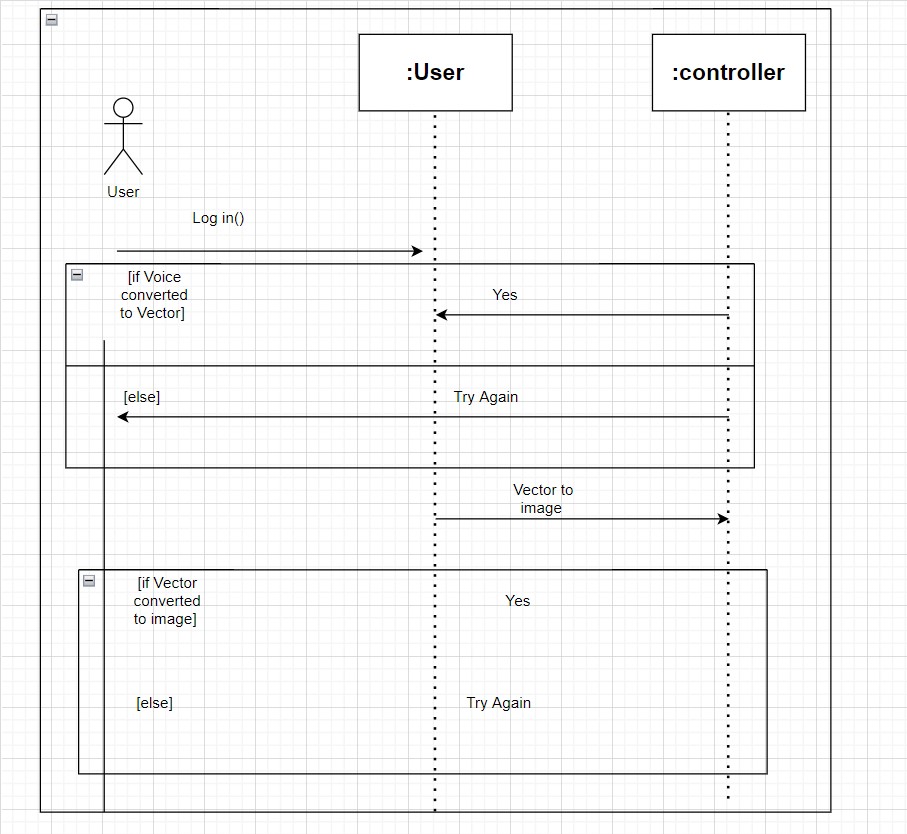
#### Module 1: Profile Management



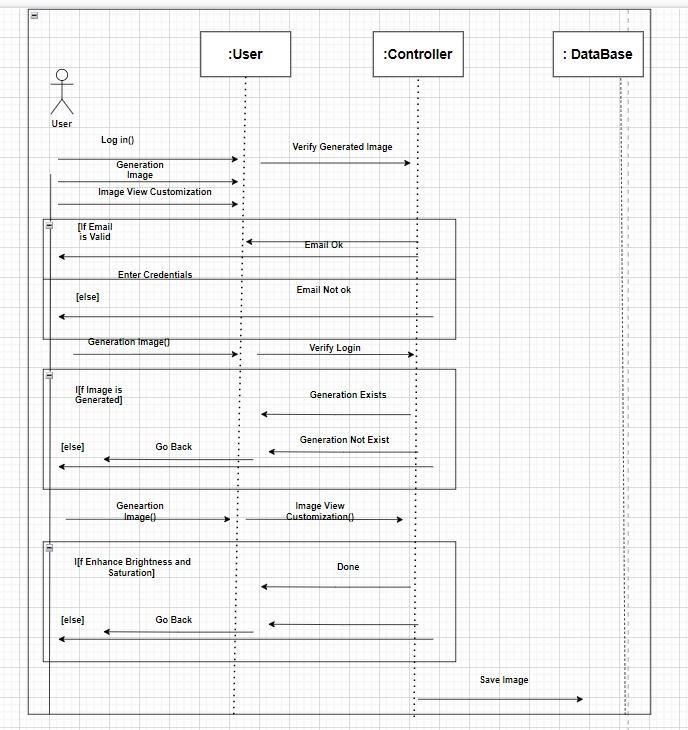
#### Module 2: Place Voice Record



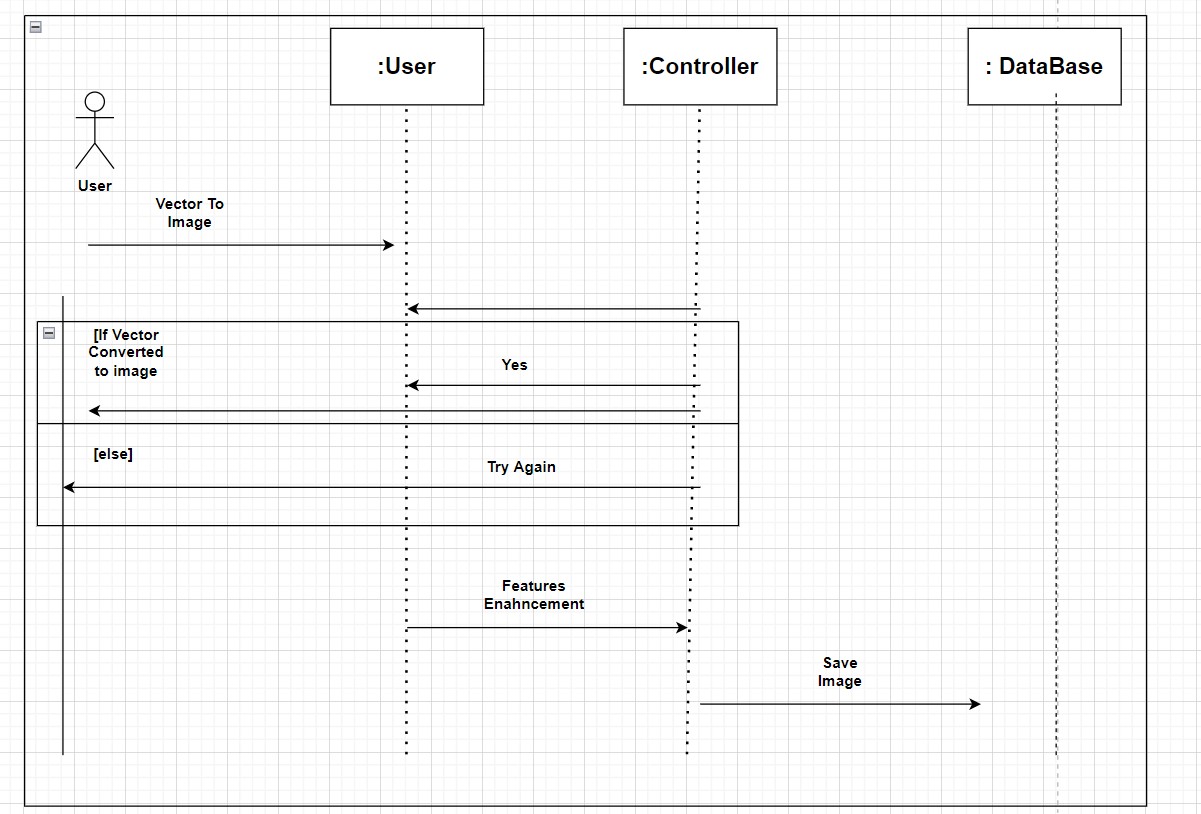
#### Module 3 and 4: Vector and Image Generation Via Voice



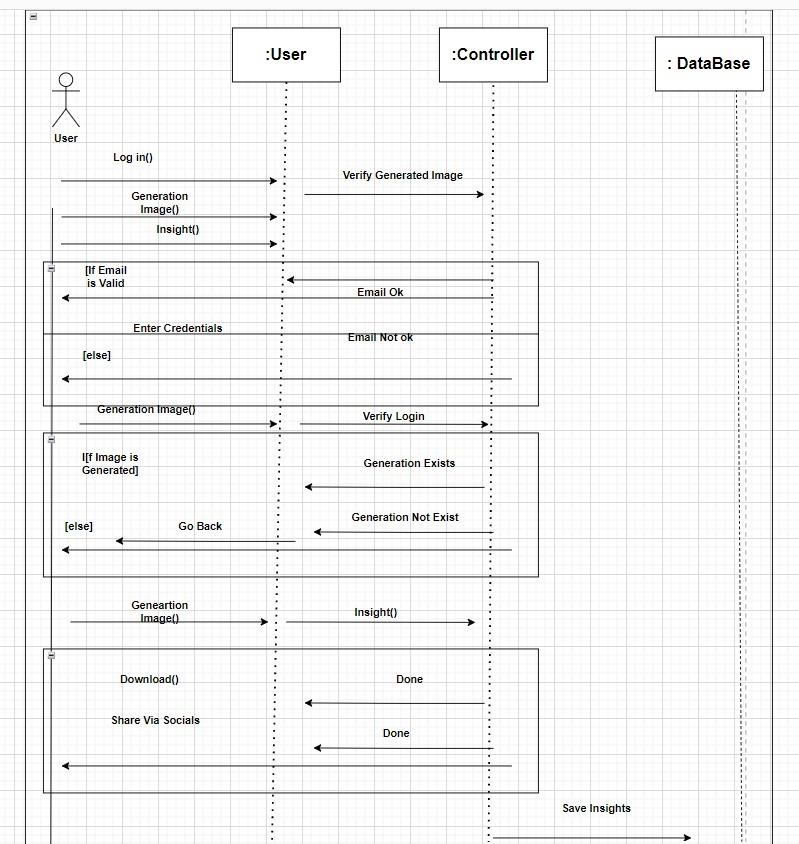
#### Module 5: Image View Customization



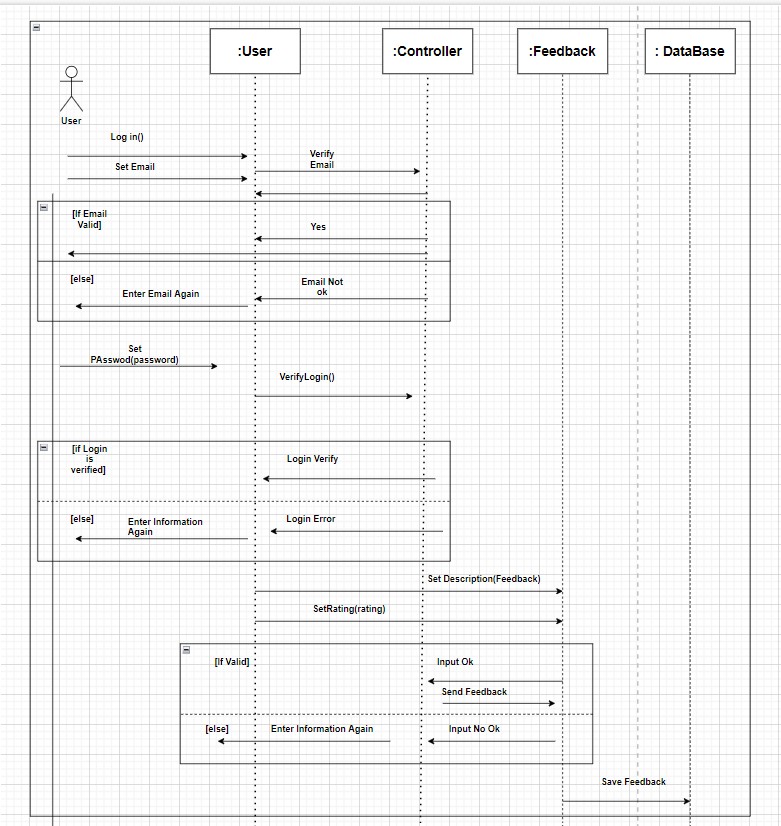
#### Module 6: Features Enhancer



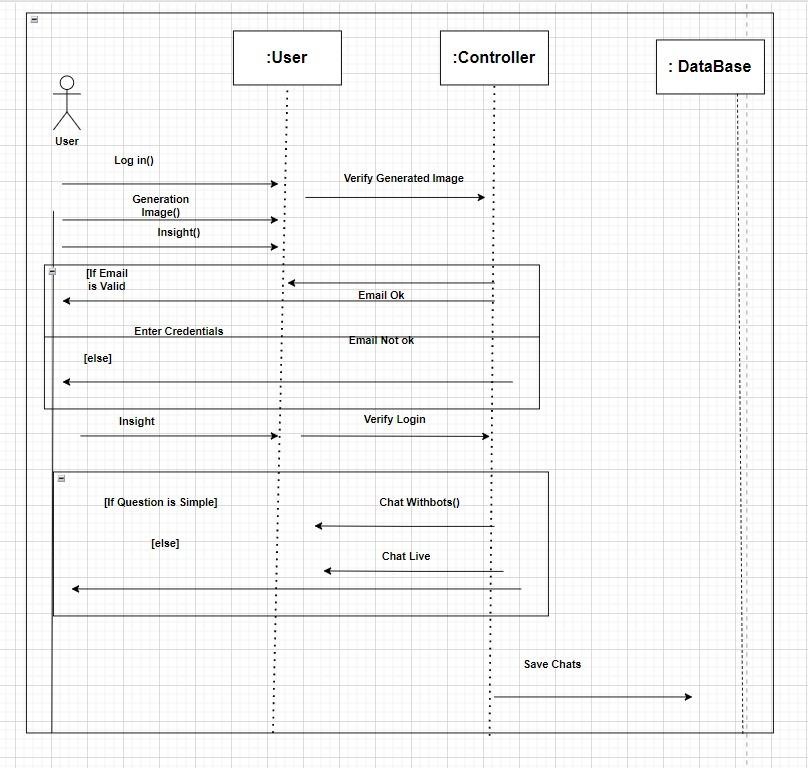
#### Module 7: Insight Module



#### Module 8: Feedback Panel



#### Module 9: Help and Support



1. **Data design** 
   1. **Relational Database Model**

User(Email,Username(PK),password, confirm password)

VoiceForLogin(Username(FK),VoiceID(PK), VoiceNote)

VoiceRecordAgainstAccounts(Username(FK), Voice ID(PK), ArrayOfVoices)

Insights(Username(FK),age, race,skin-color,PDF,jpg)

VectorModels((Username(FK),generatedVectorImageID)(PK), generatedVectorImage)

ImageModel(Username(FK), ImageModelID(PK),ImageModel)

ImageViewCustomization(Username(FK),ImageCustomID(PK),Brightness,Saturation,Hue)

FeaturesEnhancer(Username(FK),Image CustomID(PK),NoseValues,EyeColor,hairColor,skinColor)

* 1. **Data dictionary**

|  |  |  |  |
| --- | --- | --- | --- |
| **Collections** | **Properties** | **Data Type** | **Description** |
| Name | First name and last name | String | It will contain full Name of the user |
| Email | User Email | String | It will contain the email of the user in specific format. |
| Voice for login | Account login voice | Voice | It will contain the voice of the user for login. |
| Password | Password for login | String | It will contain the country of the user. |
| Voice Array for conversion | Place Holder for audio notes | Array of Voice | It will contain the array of voices a user want to test. |
| Vector | Vector against Voice | Vector (SVG) | It will contain the vector file generated from voice |
| Image | Image against the Vector | Image (jpeg) | It will contain the image generated from vector. |
| Age | Age predicted via Insights by ML | String | It will contain the age predicted by ML |
| Race | Race Predicted via Insights by ML | String | It will contain the race predicted by ML |
| Insights | Details by insight module | String | It will contain the whole description of the insights generated. |
| Feedback | Improving the System with Feedback | String | The Feedback Property helps us in improving the system |
| ID | User id | String | It will contain user ID |
| Filters | For Image Customization | Filter (double values) | It will contain the numeric values for image views |

1. **Algorithm & Implementation** 
   1. **Sign up and Login**

DO

INPUT email

IF(verifyEmail NOT EQUAL true)

DISPLAY “Invalid Email”

WHILE (verifyEmail EQUAL false)

INPUT password

IF(account NOT found)

DISPLAY “Invalid login credentials! Please try again!”

ELSE

DISPLAY “Successfully logged in!”

* 1. **Adding Voice Record**

IF(Login)

Do

INPUT VOICE RECORD

INPUT USER\_ID

IF(VOICE NOT DELETED)

DISPLAY “VOICE SAVED”

* 1. **Voice to Vector Modeling**

IF (Login && voiceSaved)

InitialzeVectorModeling

WHILE(!Modeling successful)

Initialize Again

ELSE

Generate Vector

DISPLAY ”Generate and Save Vector Image”

* 1. **Vector to Image Modeling**

IF (Login && VectorIsSaved)

InitialzeImageModeling

WHILE(!Modeling successful)

Initialize Again

ELSE

GENERATE Image

DISPLAY “Image Saved”

* 1. **Image View Customization**

IF (Login && ImageSaved)

ShowCutomizeButton

IF(CustomizedButtonClicked)

ShowCustomizeScreen

WHILE(Brightness || saturation || hue)

setController;

MODIFY Image

DISPLAY Image

* 1. **Features Enhancer**

IF(login && ImageGenerated)

ShowFeaturesEnhancer

SELECT NOSE , EYE, HAIR , SKIN COLOR

WHILE (NOSE || EYE || HAIR || SKIN is Modifying)

setController;

MODIFY Image

DISPLAY Image

* 1. **Insight Panel**

IF(login && ImageGenerated)

ShowInsightsPanel

SELECT (Download via JPG || DOCX || Share via socials) ExecuteSelectedCommand

**Following are the API used by Speech2face -System:**

|  |  |  |
| --- | --- | --- |
| **Name of API** | **Description** | **Purpose** |
| IBM Watson API | IBM Watson API is an AI powered chatbot for user support. It turns FAQ content into a helper bot. | Eliminates the need for a human to be available 24/7 to resolve redundant and common queries. |
| Facenet API by Google | Google’s facenet is a powerful dataset governed by Deep learning modules. | Google is working on speech2face functionality and it has maintained great data sets so we will use it to transform voice  into image |
| Google Ads API | The Google Ads API lets advertisers manage large Google Ads accounts and lets developers make applications  that use these ads | A way to provides user free services but the development team still gets revenue out of it. |

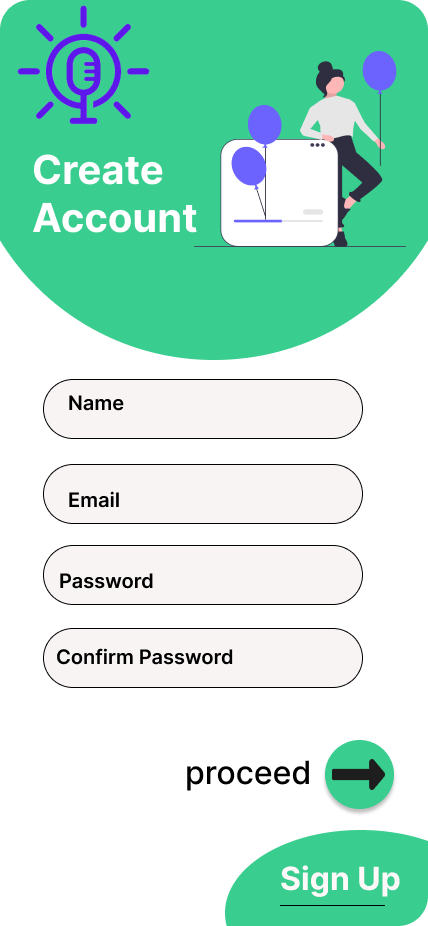
1. **Human Interface Design**

* 1. **Screen images**

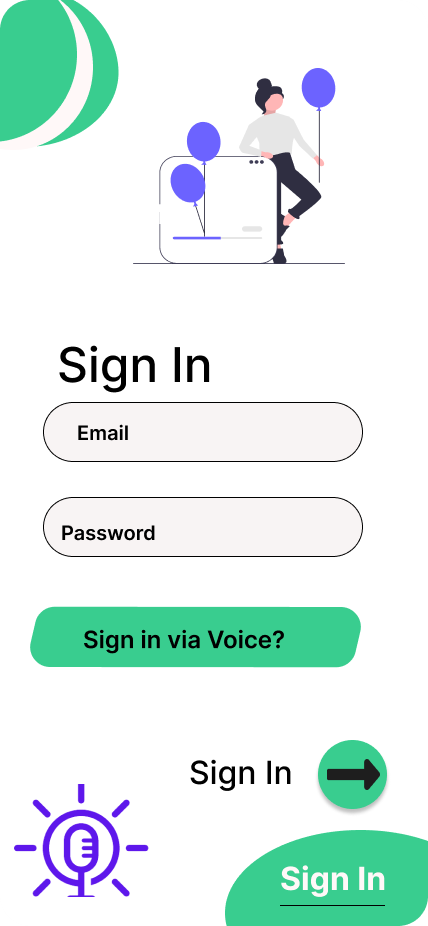
The system has a user-friendly interface. It is accessible and easy to understand. The design of the system is responsive so that it is compatible with most devices. The color schemes used in the design are also added while keeping in mind the integrity of the design so that the application looks attractive to the user. Every task can be performed by using minimum steps. All the fields that the user will provide shall be validated.

The screen images for our app are given below:

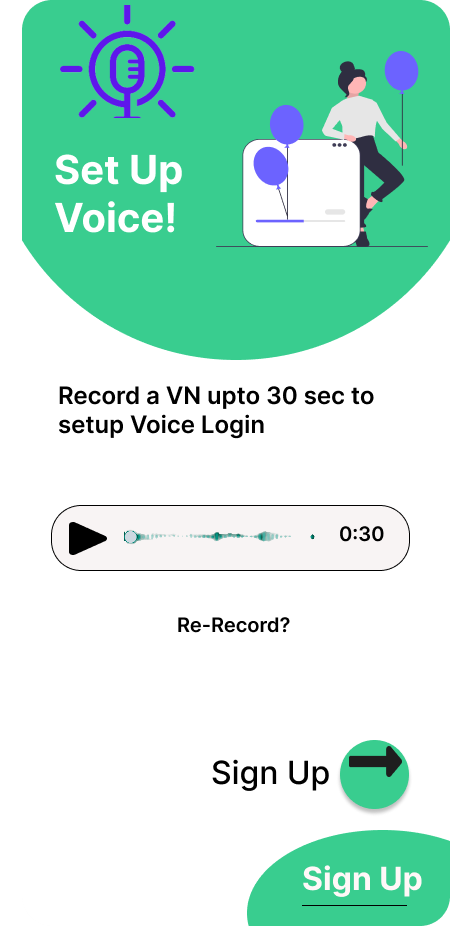
#### Splash Screen: Speech2Face



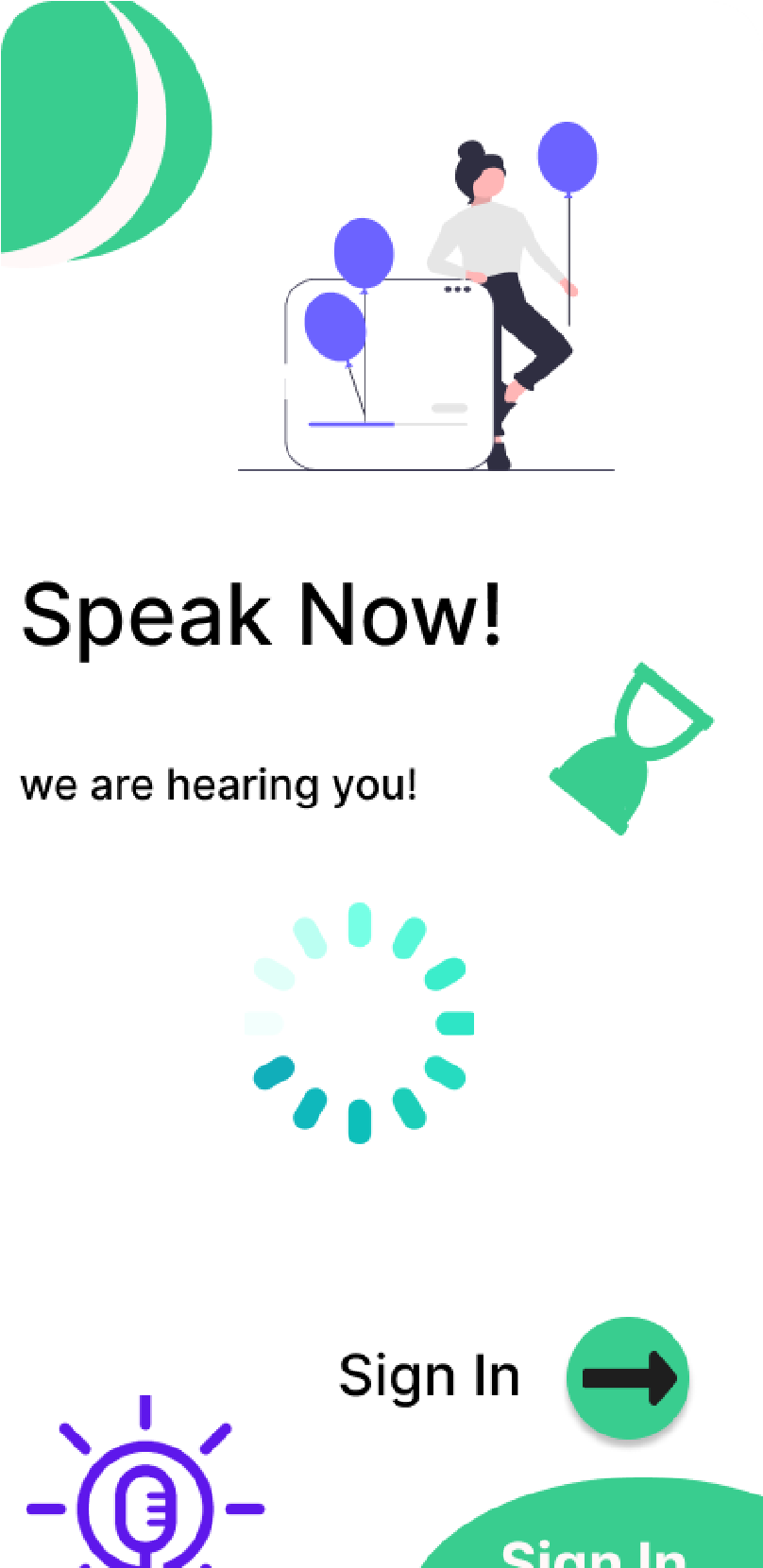
#### Sign Up Screen



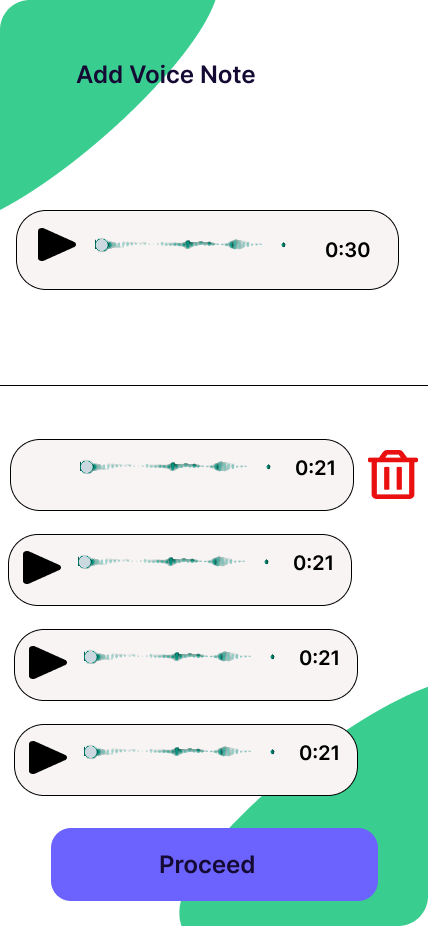
**Sign In Screen**



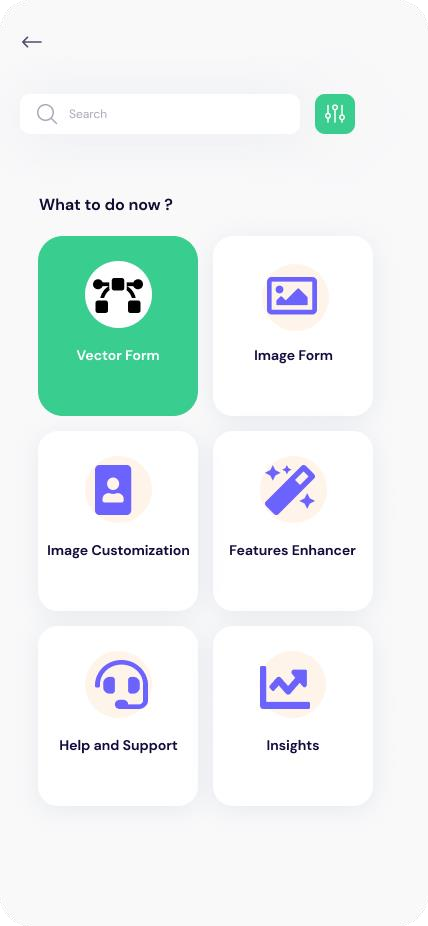
**Sign Up via Voice**



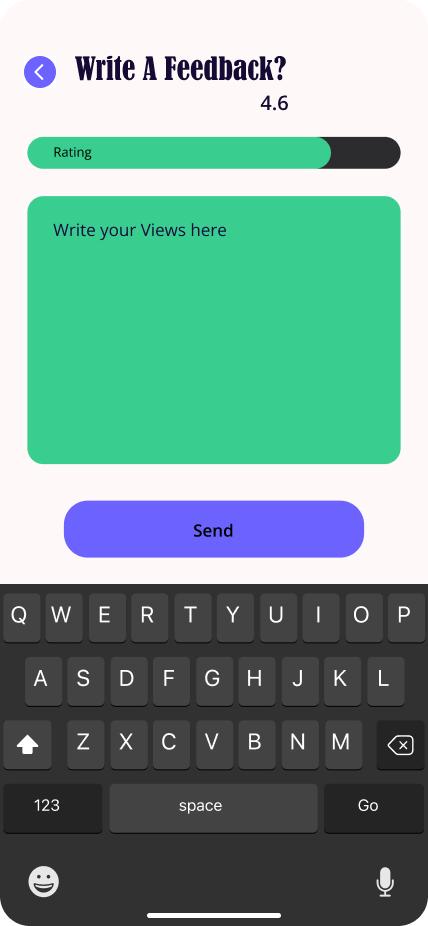
**Sign In Via Voice**



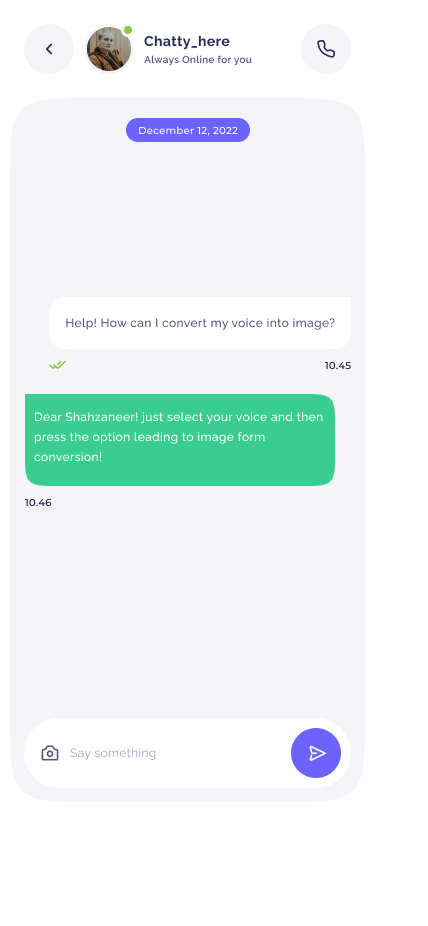
#### Place Voice Records



**Main Options Screen**



#### Feedback Screen



**Chatbot**

**7.2. Screen objects and actions**

Following are the Screen Objects that will have specific actions:

* **Login Button:** Login button is used to login the user to the application.
* **Side Option pane Button:** The User will be able to open the side pane to view the available options.
* **Vector Generation Button**: The user will be able to generate vector Image.
* **Image Generation Button**: The user will be able to generate Image from vector.
* **Image Customization:** The user will be able to customize the image.
* **Face Enhancer:** The user will be able to enhance the features of the image formed.
* **View Insights:** The user will be able to view the insights.
* **Add Voice:** The user will be able to add Voice records against his profile.
* **Settings:** The user will be able to send feedback.
* **FAQ:** The user will get their queries solved by opening the FAQ menu using FAQ button.
* **Send Feedback:** The user will be able to send feedback using feedback button.

1. **Conclusion**

The Proposed system can be developed by the developers by using the Software Design Specification Document as it contains all the necessary information that is needed for the coding and implementation phase is defined in detail in the document.

1. **References** 
   1. **Related System:** 
      * [**https://www.github.com/topics/sound-classification.**](https://www.github.com/topics/sound-classification)
      * [**https://speech2face.github.io/**](https://speech2face.github.io/)
      * [**Google Facenet**](https://www.geeksforgeeks.org/facenet-using-facial-recognition-system/%23:~:text=FaceNet%20is%20the%20name%20of,for%20Face%20Recognition%20and%20Clustering.)
   2. **YouTube Resources:**

**https://www.youtube.com/watch?v=aKYlSIs3UDY&t=334s**

1. **Plagiarism Report**

N/A

1. **SDS Work Division**

|  |  |
| --- | --- |
| **Shahzaneer Ahmed**    **(SP21-BCS-087)** | **Shayan Zameer**    **(SP21-BCS-088)** |
| 1. SDS Document (Word .docx file) 2. Activity-Diagram (Module 1, 3, 5, 7,9) 3. Sequence-Diagram (Module 1, 3, 5, 7,9) 4. Data Dictionary 5. User Interface Design (collaborated) 6. Class Diagram (collaborated) 7. ERD (collaborated) 8. Algorithm and Implementation (4 modules) | 1. SDS Presentation (PowerPoint .ppt file) 2. Activity Diagram (Module 2, 4, 6, 8) 3. Sequence Diagram (Modules 2,4,6,8) 4. JSON Schema 5. User Interface Design (collaborated) 6. Class Diagram (collaborated) 7. ERD (collaborated) 8. Algorithm and Implementation (4 modules) |