Digital Companion

An Engineering Project in Community Service

Phase - II Report

Submitted by

- 1. 19BAI10025 Jigyasa Bisht
- 2. 19BAI10054 Vanshika Sabhani
- 3. 19BAI10077 Rashmi Rawat
- 4. 19BCE10274 Sanjana Singh
- 5. 19BCE10283 Rachita Jha
- 6. 19BCE10456 Divya Mathew
- 7. 19BCG10075 Nandini Joshi
- 8. 19BOE10080 Anjali Patel

in partial fulfillment of the requirements for the degree of

Bachelor of Engineering and Technology



VIT Bhopal University

Bhopal

Madhya Pradesh

April, 2022



Bonafide Certificate

Certified that this project report titled "**Digital Companion**" is the bonafide work of "19BAI10025 Jigyasa Bisht, 19BAI10054 Vanshika Sabhani, 19BAI10077 Rashmi Rawat, 19BCE10274 Sanjana Singh, 19BCE100283 Rachita Jha, 19BCE10456 Divya Mathew, 19BCG10075 Nandini Joshi, 19BOE10080 Anjali Patel" who carried out the project work under my supervision.

This project report (Phase II) is submitted for the Project Viva-Voce examination held on 20th April, 2022.

Supervisor

Comments & Signature (Reviewer 1)

Comments & Signature (Reviewer 2)

LIST OF FIGURES

FIGURE NO.	TITLE	PAGE NO.
1.	System architecture of voice based email	8
2.	Flowchart of face recognition	10
3.	Process of Automatic Speech Recognition	11
4.	Process of text to speech recognition	12
5.	Sending mail process	13
6.	Reading mail Process	14
7.	Sending and Reading Combined mail process	14
8.	User screen	15
9.	Registration page	15
10.	Camera capturing the image of the user	15
11.	Giving instructions to perform send, read and exit functionality	16 17
12.	Using send functionality to send an email	18
13.	Displaying details to be sent by the user	
14.	Using read functionality to read an email	18
15.	Displaying details to be read by the user	19
16.	Using exit functionality to exit	19
17.	Screenshot of the email read by the email	20
4.0	assistant	20
18.	Screenshot of the email sent by the email assistant	21
19.	Plagiarism Report	24
		3 Pag

INDEX

CHAPTER NO.	TITLE	PAGE NO.
	List of figures	3
1.	Introduction	5-6
	1.1 Motivation	5-6
	1.2 Objective	6
2.	Existing work/ literature review	7
3.	Topic of work	8- 20
	3.1 System architecture	8
	3.2 Working principle	8-14
	3.2.1 Face recognition module 3.2.2 Speech recognition module 3.2.3 Email- sending and receiving module 3.3 Results and discussion	8-11 11- 13 13- 14
	3.4 Individual contribution by the members	21
4.	Conclusion and future enhancement	22
5.	References	22-23

1. Introduction

The internet is often regarded as the best innovation on the planet. Today's current era is digital, and it sees regular changes, as well as the implementation of those changes. These technologies are not available to people who are not physically perfect. Special efforts, such as voice-based email systems, benefit those with poor eyesight or who are visually handicapped. Emails have shown to be the most effective means of communication. There are other communication channels accessible, but email is the most commonly utilized in every profession, including education and business. Email services are not known to everyone on the earth. A person's vision must be clear in order to understand what is displayed on the screen. For visually impaired persons, using the internet is absolutely useless if they are unable to see the screen. There are many people who have visual impairments who wish to access all of the services that regular people do. To address these issues, we are building this project to make email easily accessible to blind people so that they can use it for free. The user who uses this project does not need to remember all of the important sites. Every action will be based on the interactions between commands. Every command will be provided by speech therefore the blind consumers don't need to worry about the email interface interaction. Existing technology, such as screen readers, automatic voice recognizers, speech-to-text, and text-to-speech, made things simpler for the visually handicapped, but only to a degree. As a result, we developed this voice-based email system for the blind, which enables visually impaired people to utilize email with ease. The most important consideration we made while building this system was to ensure the user's security by implementing face recognition that has been used for security and verification purposes. This project presents a desktop application based on Python that is specifically built for visually challenged persons. This application manages a voice-based emailing service that allows users to read and send emails without the need for assistance from their email accounts. Users can use certain keywords to do specific operations, such as Read, Send, Compose Mail, and so on. A visually impaired individual can utilize the voice mail system to account for mail services quickly and often. As a result, the visually impaired individual should not use third-party software to access emails since they are unsafe. This is a python-based desktop programme for visually impaired persons that uses speech to text and text to speech modules to allow everyone to operate their own accounts by voice inputs. Based on their matching commands, this system will always prompt the user to do desired functions. The fundamental motivation for creating this project is that because the keyboard is being phased out and mouse actions are becoming less common, many operations will rely solely on voice. Users of this system do not need to know where the keys are placed or what keyboard shortcuts they can use.

1.1 Motivation

This technique was intended to assist persons who are blind or partially impaired, as well as those who are illiterate. This is the system's primary and most significant goal. This project is an application for visually impaired or illiterate users who want to use email services in the same way that everyone else does. The fundamental motivation for this is that the user will no longer need to remember keyboard functions because all operations will be performed using voice commands. This system's purpose is to make the system more user-friendly and efficient. The user can effortlessly handle the system thanks to the voice-based technology, and will be able to

work with more speed and precision. The majority of internet pages are meant to be utilized and accessed by sighted individuals. One of the goals of this project is to present information to web users in a way that is both appealing and practical. However, for a blind or low-vision person, or someone who is illiterate, such web page building is not necessarily efficient or user friendly. The project's major goal is to address this problem. Nearly 285 million individuals worldwide are believed to be visually impaired, and the goal is to make an appropriate communication system available to them [1]. This was the motivating force for the creation of a certain system. The fact that all operations are predicated on mouse click events is one of the fundamental drawbacks of the current system. Our concept focuses on delivering basic functions such as create, send, and receive email, as well as advanced capabilities such as voice-based operation, mail search, and support for both voice and text-based email, all with additional convenience and simplicity.

1.2 Objective

The goal of Voice Based Email for Visually Impaired is to make it easier and more efficient to access emails. This application is built on the usage of speech-to-text and text-to-speech converters, allowing anyone to operate their email accounts solely with their voice and read, send, and perform all other necessary operations. The system will issue voice orders to the user to conduct particular things, and the user will answer. Speech-to-Text, commonly known as Automatic Speech Recognition, turns spoken speech into text, making email writing a breeze. The Text-to-Speech module generates an audio output of the message received, with the sender, topic, and body of the mail being read aloud by the system. The goal of this project is to create an application that will allow visually impaired people to utilize every service that a normal person uses to send and receive emails efficiently without needing a third party. This project attempts to design an application that uses the keyboard as little as possible and does not utilize the mouse at all. Text-to-speech and speech-to-text interactions are used to carry out all of the functions. The project is primarily concerned with sending and receiving emails. Every step is interactive, and it is completed by accepting a few actions, such as composing and receiving mail.

2. Existing Work / Literature Review

The most frequent postal services that we use on a daily basis are inaccessible to visually impaired people. This is due to the fact that they do not provide any means for the person in front of the screen to hear the content. They are unable to determine where to click in order to complete the essential actions since they are unable to visualize what is currently on screen. The current method is inaccessible to blind persons since it does not provide aural feedback while reading out text. There are now available email systems that only allow for voice recognition and text-to-speech systems to be accessed by remembering the keyboard. Speech recognition, interactive voice response, and the elimination of mouse click events are all used in the proposed system.

Literature review:

- According to a statistical report[2], email, 2014-2018, issued by the technology research firm, the market, Palo Alto, California, United States, and based on the, just 4.1 million email addresses were generated in 2014, but there are now over 5.2 billion users. It is one of the most regularly used communication methods because the invoice must be completed before the end of 2018. According to studies conducted by the Vision Loss Expert Group, all of the world's 253 million individuals are blind or have vision difficulties, which implies that roughly 253 million people can't go on the internet or send emails. In a study, they offer an email system that is simple to use for visually impaired people. Speech-to-text converters, Text-to-speech converters, and the Viterbi Algorithm are all considered. The algorithmic rule works by detecting the most suitable word as soon as the user spells it, and then matching the words that are guessed with the word that is pronounced.
- Saurabh Sawant and colleagues[3] presented a technique for visually impaired and illiterate people to improve their email connection. This system replaces the IVR system, which formerly relied on Screen Readers and a Braille Keyboard. Use a PHP feature called PHP mailer for the functionality. It's a library that allows you to send an email. For searching mail in boxes, the Knuth-Morris-Pratt Algorithm is utilized
- Payal Dudhbale et al. (2012)[3] developed a voice-based solution for blind persons on desktop and mobile devices. Blind persons can use the architecture to readily access the operating system's email and multimedia services.
- Visually challenged people can use email (2013). The system will urge the user to do the appropriate activities and will provide specific capabilities, according to the paper Voice based email system for blinds (2015)[4].
- The paper Voice Recognition System for the Visually Impaired-Virtual Cognitive Approach (2008) describes how a visually impaired individual can utilize a microphone to interact with the system[4].
- The speech-based email system (2018) employs voice recognition to transmit and read emails. Voice-based email system for blind people (2017) develops a desktop application that allows users to send and receive emails using voice commands[5].
- Voice recognition, TTS, and STT are used in the paper voice based email system for blind (2016)[6].

3. Topic of the work

A) System Architecture

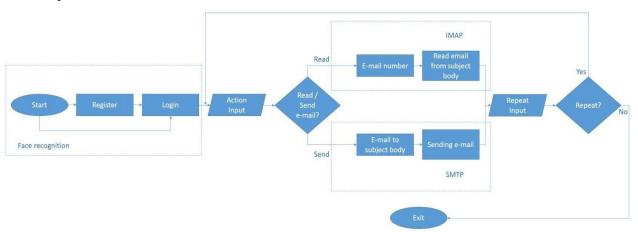


Fig. 1. System architecture of voice based email

As soon as the program starts, the email assistant asks the user whether he wants to register or login. If the user is new, then he will have to register first. On selecting 'register', the assistant asks the user his name and captures his face and saves it in a file named after the user. Then the assistant moves to the login functionality where the camera is on for a few seconds and the assistant recognizes the user's face with the file already saved. If the face matches, the user gets logged in. Then the assistant asks the user whether he wants to send the email, read the email or simply exit. If the user selects 'send', then the email assistant asks to whom he wants to send the email, followed by the subject and body of the mail. As per the input given by the user, the mail is sent to the required recipient. If the user selects 'read', then the assistant asks for the serial number of the email he wants to read from the latest. As per input given by the user, the assistant reads the required email. And if the user selects 'exit', the assistant simply says 'bye bye' and the program terminates.

B) Working Principles

The main objective of our project is to help the visually impaired people so that they can easily read, write or send their emails with the help of their voice. Keeping in mind the security of the user, this email service is available only if the user is registered, so that no third person can access others' email. And to do this work our project is mainly divided into three different modules, which will help the user to perform all these operations. Therefore, the three modules that are implemented in this project are as follow:

1. Face Recognition Module-

The face recognition module is mainly implemented for the security purpose of the user. The authentication of the user to login into the system to use the services is mainly done by facial recognition system. The basic working principle of this module is, the device

will capture the face of the user requesting to enter his/her name and store it into the provided location path of the computer. This module has two functionalities, first one is the register button which will be implemented by the new user who is registering into the system for the very first time and the second one is the login button which will be used by the existing user. When someone attempts to enter, the camera will capture the face of that person, perform various cv2 operations and compute the face encodings. If a correct match is found it will login the user into the system, or otherwise the person has to register in order to enter into the system to access the email service. Given below is a flowchart to explain the step by step working of the module in detail.

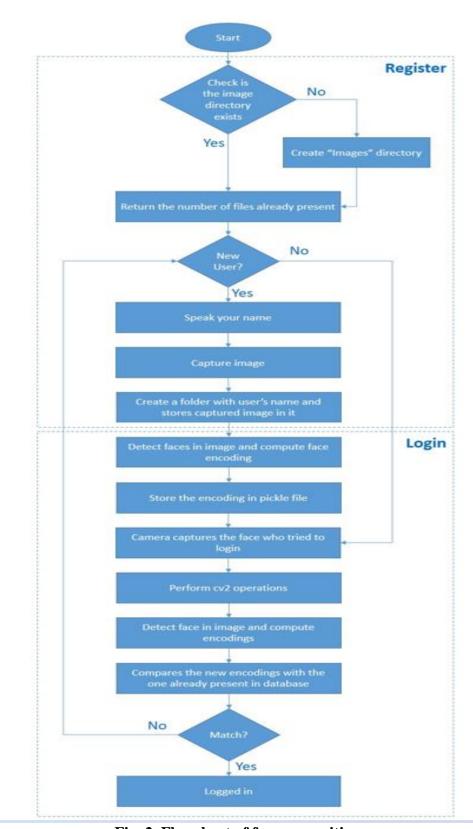


Fig. 2. Flowchart of face recognition

The major libraries that will be used in this module are as follows:

- DLIB It is used for facial landmark detection...
- Face Recognition It allows us to use various important functions for face matching.
- Open-CV It is an image processing library which allows us to perform image and video analysis.
- OS It allows us to use various functions to interact with the operating system.

The three main process that are being implemented in the facial recognition module will be-

- Face Detection- To recognise a face, it is first important that we detect the face first. We extract a human face and then move on to the next step.
- Feature extraction using face embedding The next step is to extract features from a face using a face embedding model. A face embedding is a vector that represents the features extracted from the given face and we use these vectors to recognise faces. That face embedding for the identical face may be really close in the vector space, whereas the face embeddings of two different faces may be really far away. We get a face embedding after passing the image through a face embedding model.
- Face recognition- We have face embedding for each face in the system. After
 entering a new face to the system, it calculates its face embedding and
 compares it with the ones that already exist. The face is recognised, if its face
 embedding closely matches any other face embedding in the database.

2. Speech-Recognition Module-

Speech-to-text-

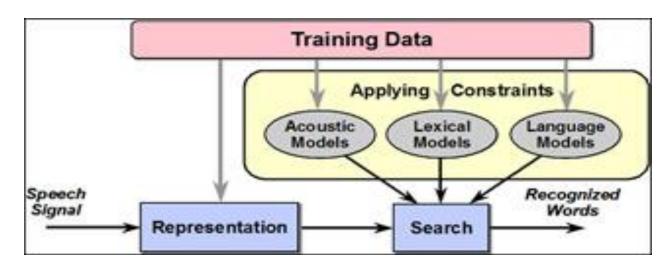


Fig. 3. Process of Automatic Speech Recognition[7]

The basic idea of automatic speech recognition is that any human's speech or words produce vibrations in the air, which are known as sound waves. These continuous or analogue waves are digitized, processed, and finally decoded into words and sentences that make sense.

- Voice can be thought of as an acoustic waveform, which is a signal that contains message information.
- A normal human can generate voice at an average rate of 10 tonnes per second. The average data rate is 50/60 bits per second that when converted to information accounts for roughly the same rate.
- The microphone converts this acoustic waveform into an analog electrical signal. These analog signals are further converted to a digital sample by an analog-to-digital converter.
- This converter accurately measures the shaft at discrete intervals.
- The digitized signal is a stream of periodic signals sampled at 16000 times per second and is unsuitable for actual speech recognition because the pattern is difficult to locate.
- To extract the actual information, the time domain signal is converted to the frequency domain signal. The FFT technique is used by the Digital Signal Processor to accomplish this.
- The component in the digital signal is analyzed every 1/100th of a second, and the frequency spectrum for each such component is computed. In other words, the digitized signal is divided into small frequency amplitude segments.

• Text-to-speech Input Text Preprocessing Unicode conversion Segmentation Segmentation Concatenation

Fig. 4. Process of text to speech recognition

The flow of the text to speech module is depicted in the above diagram:

- Language processing: Starting with the approach of language processing, it generates a written text of the scanned text on the side of prosody.
- Digital Signal module processing: It translates representative data from information science into audible and understandable speech.
- Pre- processing: In pre-processing the input text, text to speech conversion can be done. Abbreviations, acronyms, and numerals in the text are extended here.
- Unicode: Following that, the pre-processed text will be transformed to Unicode. Unicode was created with the specific goal of overcoming the constraints of traditional encodings.

- The pre-processed text is used to identify the fonts in the input text, which is then converted to Unicode. The encoded text has now been divided into syllables, and duplicates have been deleted.
- Syllable: The syllabled text is then mapped to the database's syllable sound files. For the final outputs, these syllables will be concatenated and smoothed.
- The level of the open system is determined by the language of the software's inflection factors (phrasing and accentuation), the model's amplitude, and the software's length, which includes the long sound, and, as a result, the length of the residues that determine the length of the language, units, time, and speaking.
- This results in a natural-sounding human voice. Variations can be made to the final product. The user can customize the voice, rate, and loudness of the output speech.

3. Email- Sending and Receiving Module-

It is the most important module of our project. The main functionality of sending and reading of all the mails by the user will be implemented in this module. This module will be divided into two parts-

• Sending email

SMTP is a protocol that would be used for sending email. When we write an email in gmail and hit the send button, the email travels from our computer to our email SMTP server using SMTP protocol. The SMTP server will send the message to the recipient's email server using SMTP. Then the email will stay on the recipient's email server until the recipient logs into the email account.

The major library used for this is Smtplib. The smtplib is a Python library for sending emails using the straightforward Mail Transfer Protocol (SMTP). We use the following command to install smtplib to our python interpreter: -!pip install smtplib.

The user first gives the command that he wants to 'SEND' the email. The email assistant asks the user to whom he wants to send the email, followed by the subject and body of the email. Further, the email is sent to the required recipient and the gui corresponding to it is displayed.

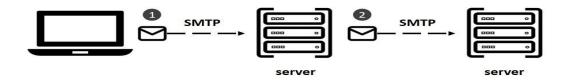


Fig. 5. Sending mail process

• Reading email

IMAP is a protocol used for reading email from an email accessible account. IMAP functions are the intermediary between the person's email server and email client. Reading email using IMAP means, reading email off the server. This works inside an

email client e.g. gmail.com, when the user signs into their gmail account, the client contacts the gmail server using IMAP. Now after connecting the subject of emails are displayed by the email client and IMAP reads the message for the user. IMAP helps letting users read email quickly.

The major libraries used for this are imaplib. The imaplib is a Python library for accessing emails using the Internet Mail Access Protocol (IMAP). We use the following command to install smtplib to our python interpreter: -!pip install imaptlib.

The user first gives the command that he wants to 'READ' the email. The email assistant asks the user the order of email to get read. Further, the email is read with every detail provided to the user with the gui displaying the corresponding email message.

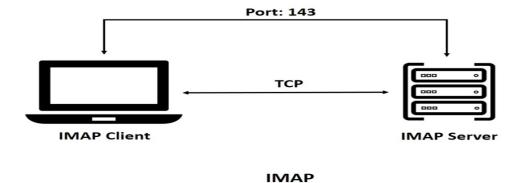


Fig. 6. Reading mail process

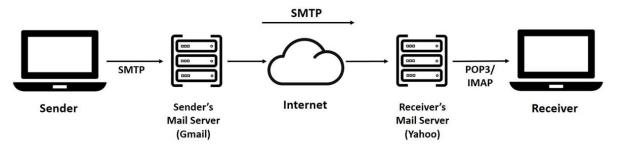


Fig. 7. Sending and Reading combined mail process

C) Results and Discussion

We have combined all the three above given modules and proposed a system which is specially designed for the visually impaired users that provides voice-based mailing service where they could read and send emails on their own, without any guidance. In this project we have worked with tkinter in pycharm to make the graphical user interface(GUI) to ease the use of the system for the user.

Userscreen- The user screen interface will appear where it will ask to choose between register or login. If it is a new user then they will speak register to register into the system and if already an existing user then they will speak login to login into the system.

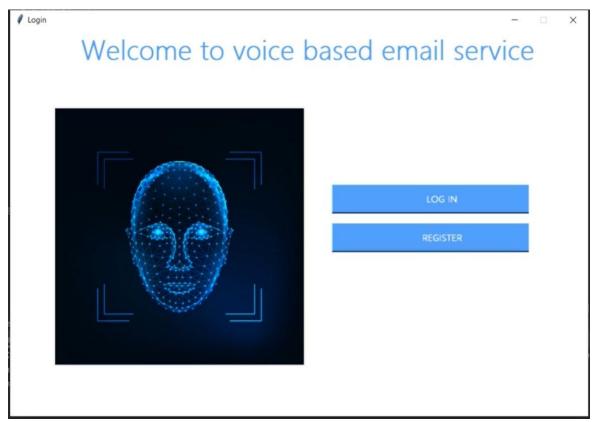


Fig. 8: User Screen

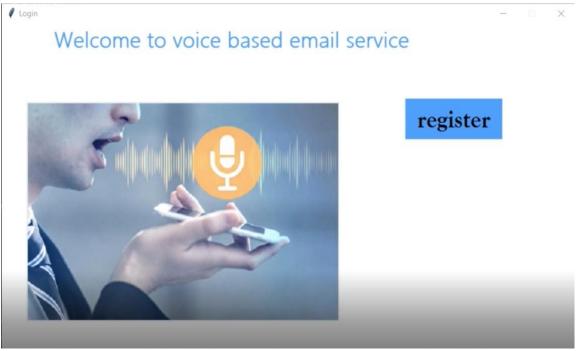


Fig. 9: Registration page

Camera capturing the user's image- Once the user speaks the chosen option, the camera of the device will open to capture the image if it is a new user. After capturing the image it notifies the user that they have successfully registered into the system.

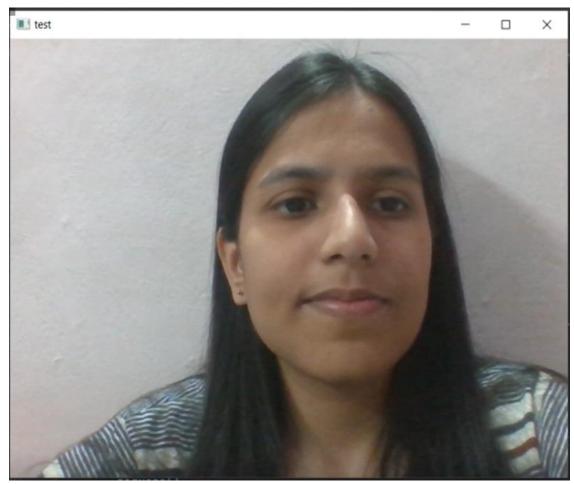


Fig. 10: Camera capturing the image of the user

Giving instructions to perform send, read and exit functionality- Once it successfully logged into the system it will again ask the user to to speak send to send mail, read to read inbox or exit to exit from the system.

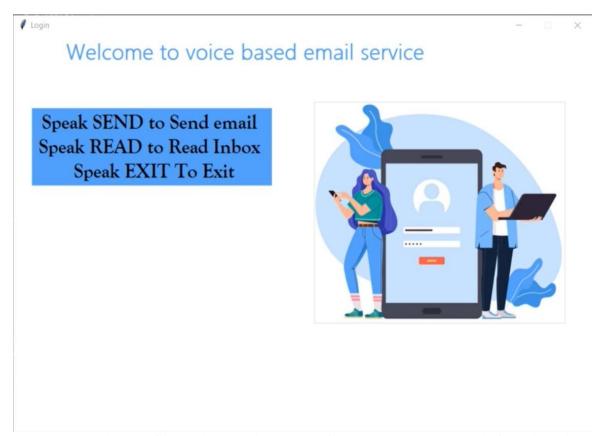


Fig. 11: Giving instructions to perform send, read and exit functionality

Using send functionality to send an email- If the user chooses to speak the send option to send the mail, the system will show the send interface. Then the system will ask for the details like to whom the user wants to send the mail, what will be the subject of the mail and the text to be sent in the mail. All these details would be displayed on the interface once the user speaks.

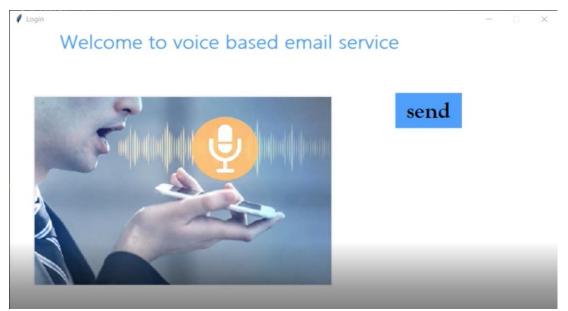


Fig.12: Using send functionality to send an email

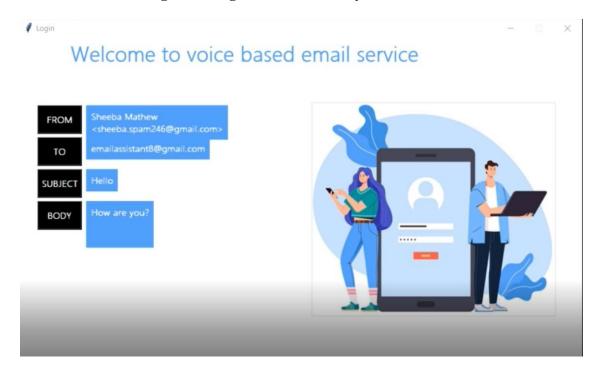


Fig.13: Displaying details to be sent by the user

Using read functionality to read an email- If the user chooses to speak the read option to read the mail, the system will show the read interface. All the details like who sent the mail, what is the subject of the mail and what is the text that the sender has sent would be displayed on the interface once the user speaks the read option.

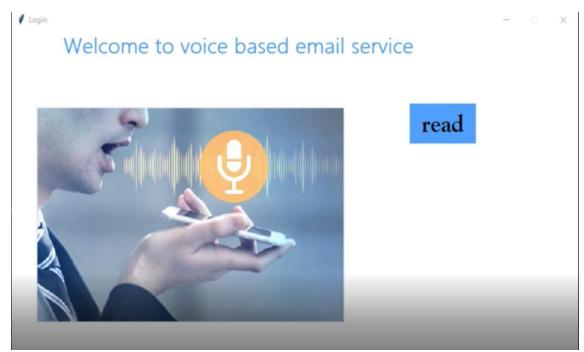


Fig. 14: Using read functionality to read an email

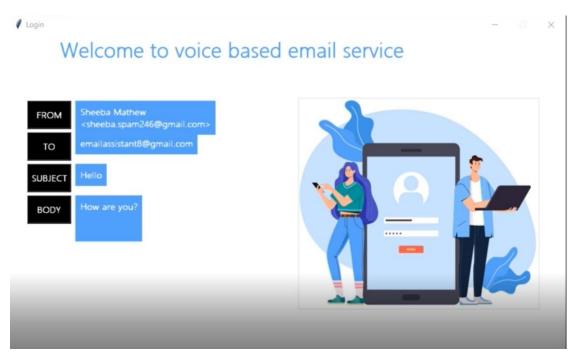


Fig. 15: Displaying details to be read by the user

Using exit functionality to exit- If the user chooses to speak the exit option it will show the exit interface and get exit from the system.

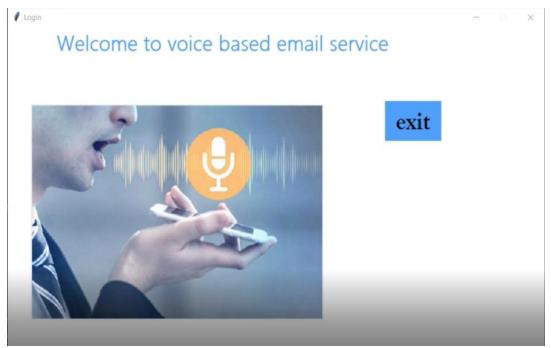


Fig. 16: Using exit functionality to exit

Below we have attached the screenshots of the email read and sent by the email assistant

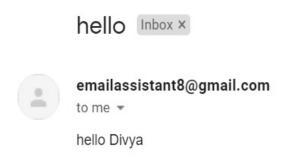


Fig. 17: Screenshot of the email read by the email assistant



Fig. 18: Screenshot of the email sent by the email assistant

D) Individual contribution by members

DSN3099 - Engineering Project in Community Service (EPICS) was a wonderful experience for me. As we were given 1 full year to develop this project, we had plenty of time to brainstorm, learn, evolve and rectify our mistakes. Starting from July when we were asked to form a team of 8 members, then selecting a topic that would cover all the domains such that each member of the group can contribute to finally bringing our ideas to action and execute the project helped me a lot to understand how challenging, at the same time, thought provoking it is to work in a team. You get to know ideas and opinions from people of varied domains. Around mid May, when we started executing our project, I faced a lot of problems while coding, but I knew that my team members were always there to help me out. I was specifically working on sending email functionality using SMTP. First I started with gathering information about how SMTP works. For this, I had to study computer networks thoroughly. I took help from Anjali and Nandini as they were in charge of research work. Then I studied about smtplib and its various commands and their functionality. smtplib is a Python library for sending emails using the Simple Mail Transfer Protocol (SMTP). Thankfully, I had prior knowledge about python so I was able to write the main code for sending email using SMTP. My major part was done here. Then, as I had done speech recognition in my previous project, I helped Rashmi and Rachita in their module of speech recognition. By the time of Review 2, our individual modules were ready. We were left with merging and gui designing. So, as soon as review 2 was approved by our guide, I started merging all the modules. It took me about 2-3 weeks when I was finally able to merge the code along with my colleague Jigyasa. Then we were only left with the gui design and it took me hardly a day or two to make the basic layout of gui. Then I gave the gui content to Vanshika for beautification. And this is how our entire project was created with each and everyone's contribution and dedication. I realised that a lot of bugs come along when we develop something new, but there are millions of ways to debug and good research can always ease things out. This project has brought into me the spirit of team work and has improved my technical skills, especially in python and computer networks.

4. Conclusion and future enhancement

We offered a proposal tailored to the needs of those who are blind or visually impaired. It provides a voice-based mailing service that allows them to read and send mail without assistance. Users must use specific terms to do specific tasks, such as Read, Send, and so on. A blind person can use this email system to access their correspondence quickly and simply. As a result, the visually impaired's dependency on others for mail-related activities can be decreased. All of these notions have been abolished, and the system now solves all of the challenges that the visually impaired confront. It makes advantage of a speech recognition programme to give an efficient voice input technique for blind people mailing gadgets. It is also beneficial to the crippled and illiterate. The project's major goal is to develop email communication for blind individuals using voice commands due to their incapacity to use the internet and its functionalities. We were successful in receiving unseen mails and generating a voice output with the sender's mail id, subject, and message. We were successful in developing text-to-speech and speech-to-text modules, as well as implementing a chatbot for effective user-system communication. We designed a register and login module for users' convenience, as well as security and privacy to some level utilizing face recognition.

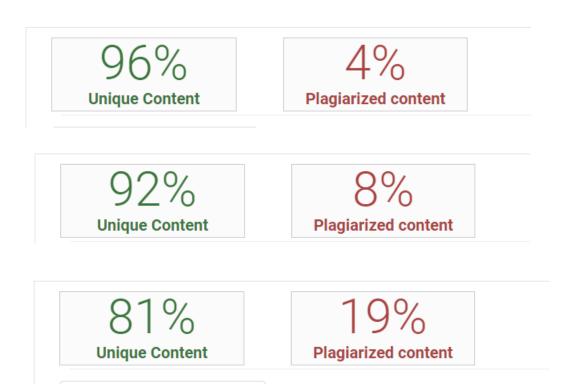
The application's key flaws could be used to improve this project in the future. This application has two significant flaws: accurate voice recognition and the ability to connect an image or document. As a result, we can add an image or document attachment for the sender in the future. In the future, we hope to make the system completely voice-based and without a keyboard. As a result, visually challenged people can easily access services. The system that has been designed so far only works on desktop computers. Because the use of mobile phones is becoming more popular, there is potential for this service to be included as a mobile phone application as well. Also, to make the system safer, security measures to be implemented during the login phase can be updated.

5. References

- 1.http://oaijse.com/VolumeArticles/FullTextPDF/360 5.VOICE BASED E-MAIL SYSTEM FOR BLIND PEOPLE.pdf
- 2. https://ijisrt.com/assets/upload/files/IJISRT21JUL1171.pdf
- 3.https://www.researchgate.net/publication/344296191_Voice_based_E-mail_for_the_Visually_Impaired
- 4. https://www.ijert.org/email-system-for-visually-impaired-people
- 5. https://ijisrt.com/assets/upload/files/IJISRT21JUL1171.pdf
- 6. https://www.irjet.net/archives/V7/i12/IRJET-V7I12430.pdf

- 7.https://ocw.aprende.org/courses/electrical-engineering-and-computer-science/6-345-automatic-speech-recognition-spring-2003/
- 8. https://www.ijircst.org/DOC/2-voice-based-email-for-blind.pdf
- 9. https://www.arcjournals.org/pdfs/ijrscse/v3-i1/5.pdf
- 10. https://www.geeksforgeeks.org/project-idea-voice-based-email-visually-challenged/
- 11.<u>https://www.seminarsonly.com/Engineering-Projects/Computer/voice-based-emailsystem.php</u>
- 12.https://www.researchgate.net/publication/324526106 A voice based text mail_system for visually_impaired
- 13. https://www.irjet.net/archives/V8/i12/IRJET-V8I12239.pdf
- 14. https://core.ac.uk/download/pdf/83592918.pdf
- 15. https://www.academia.edu/43061813/Voice Based E-mail For The Visually Impaired
- 16.http://oaijse.com/VolumeArticles/FullTextPDF/360_5.VOICE_BASED_E-MAIL_SYSTEM_FOR_BLIND_PEOPLE.pdf
- 17.https://www.researchgate.net/figure/Block-diagram-of-Text-to-Speech-System-Techniques-of-speech-synthesis-5-a_fig3_304601298
- 18. https://www.ijert.org/voice-based-email-system-for-blinds
- 19.http://ijariie.com/AdminUploadPdf/Voice Based E mail Application for Blind Visually I mpaired People ijariie5500.pdf
- 20. http://jit.ac.in/journal/cse/CSE_20-21/9.pdf

PLAGIARISM REPORT



 $((4+8+19)/300) \times 100 = 10.33 \%$ plagiarized content

