NATURAL LANGUAGE PROCESSING

Slot: G1+TG1

Project Title: Speech to Indian Sign Language Translator

Group Members:

Nimish K. Aggarwal – 19BCE0014

Vrushali Deshmukh – 19BCE0033

Tarang Garg – 19BCE0053

Aviral Goyal – 19BCE0883

Mohit Keyal – 19BCE2603

1. Abstract:

Be it communication, playing computer games, attending seminars or video conferences, deaf people always miss out the fun that a normal person does. Communication is the most important difficulty they face with normal people and also every normal person does not know the sign language. The aim of our project is to develop a communication system for the deaf people. It converts the audio message into the sign language. This system takes audio as input, converts this audio recording message into text and displays the relevant Indian Sign Language images or GIFs which are predefined. By using this system, the communication between normal and deaf people gets easier.

2. Introduction:

It is said that Sign language is the mother language of deaf people. This includes the combination of hand movements, arms or body and facial expressions. There are 135 types of sign languages all over the world. Some of them are American Sign Language (ASL), Indian Sign Language (ISL), British Sign Language (BSL), Australian Sign Language (Auslan) and many more. We are using Indian Sign Language in this project. This system allows the deaf community to enjoy all sort of things that normal people do from daily interaction to accessing the information.

3. Motivation:

According to an independent census, there are about 700,000 – 800,000 mute people in the world out of which thousands live in India. Quite often, these people also suffer from hearing impairments. Therefore, the only means of communication for them is the Sign Language which is not very popular. This makes life harder for the people who suffer from this disability.

Our aim is to integrate our knowledge of Computer Sciences with this issue and create a program that will take the audio of the speaker as input and show the hand signs of the Indian language according to the audio. This will help make communication easier, and it will do away with the need of us learning the sign language.

4. Related Works:

[1] As per Amit Kumar Shinde on his study of sign language to text and vice versa in Marathi Sign language recognition is one of the most important research and it is the most natural and common way of communication for the people with hearing problems. A hand gesture recognition system can help deaf persons to communicate with normal people in the absence of an interpreter. The system works both in offline mode and through web camera.

[2] Neha Poddar, Shrushti Rao, Shruti Sawant, Vrushali Somavanshi, Prof. Sumita Chandak in their paper discussed about the prevalence of deafness in India is fairly significant as it is the second most common cause of disability. A portable interpreting device which convert higher mathematics sign language into corresponding text and voice can be very useful for the deaf people and solve many difficulties.

[3] The glove based deaf-mute communication interpreter introduced by Anbarasi Rajamohan, Hemavathy R., Dhanalakshmi is a great research. The glove comprises of five flex sensors, tactile sensors and accelerometer. The controller matches the gesture with pre-stored outputs. The evaluation of interpreter was carried out for ten letters _A„ _B„ _C„ _D„ _F„ _I„ _L„ _O„ _M„ _N„ _T„ _S„ _W„.

[4] As per the Neha V. Tavari A. V. Deorankar Dr. P. N. Chatur in his report discuss that many physically impaired people rely on sign language translators to express their thoughts and to be in touch with rest of the world. The project introduces the image of the hand which is captured using a web camera. The image acquired is processed and features are extracted. Features are used as input to a classification algorithm for recognition. The recognized gesture is used to generate speech or text. In this system, flex sensor gives unstable analog output and also it requires many circuits and is thus very expensive.

5.Methodology:

5.1 Tools employed:

An application which takes in live speech or audio recording as input, converts it into textand displays the relevant Indian Sign Language images or GIFs.

- Front-end using EasyGui.
- Speech as input through microphone using PyAudio.
- Speech recognition using Google Speech API.
- Text Preprocessing using NLP.
- Dictionary based machine translation is done.
- Implementation in Visual Studio

5.2 Work Breakdown Structure:

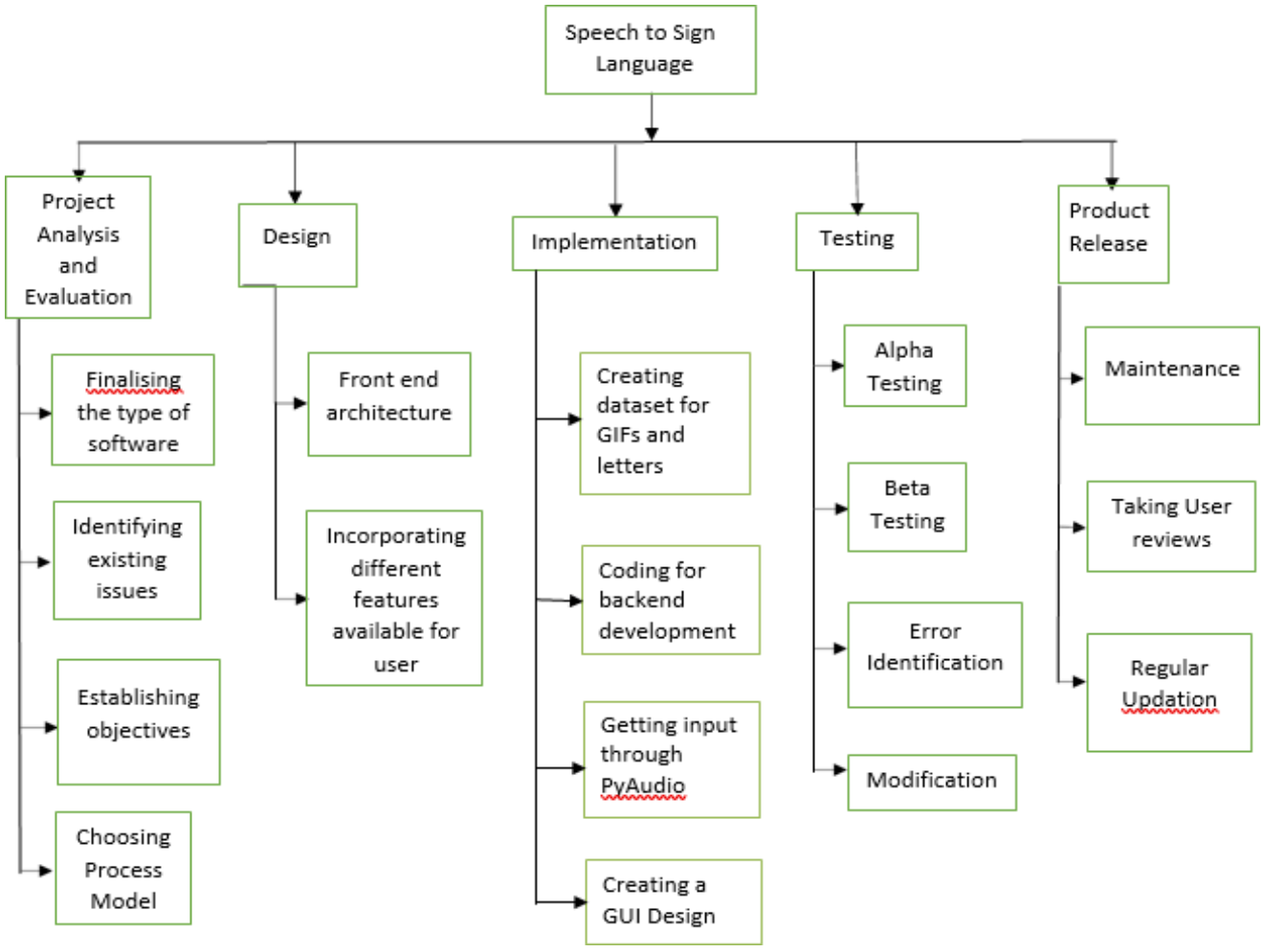


Figure 1. Work Breakdown Structure

5.3 Functional and Non-Functional Requirements:

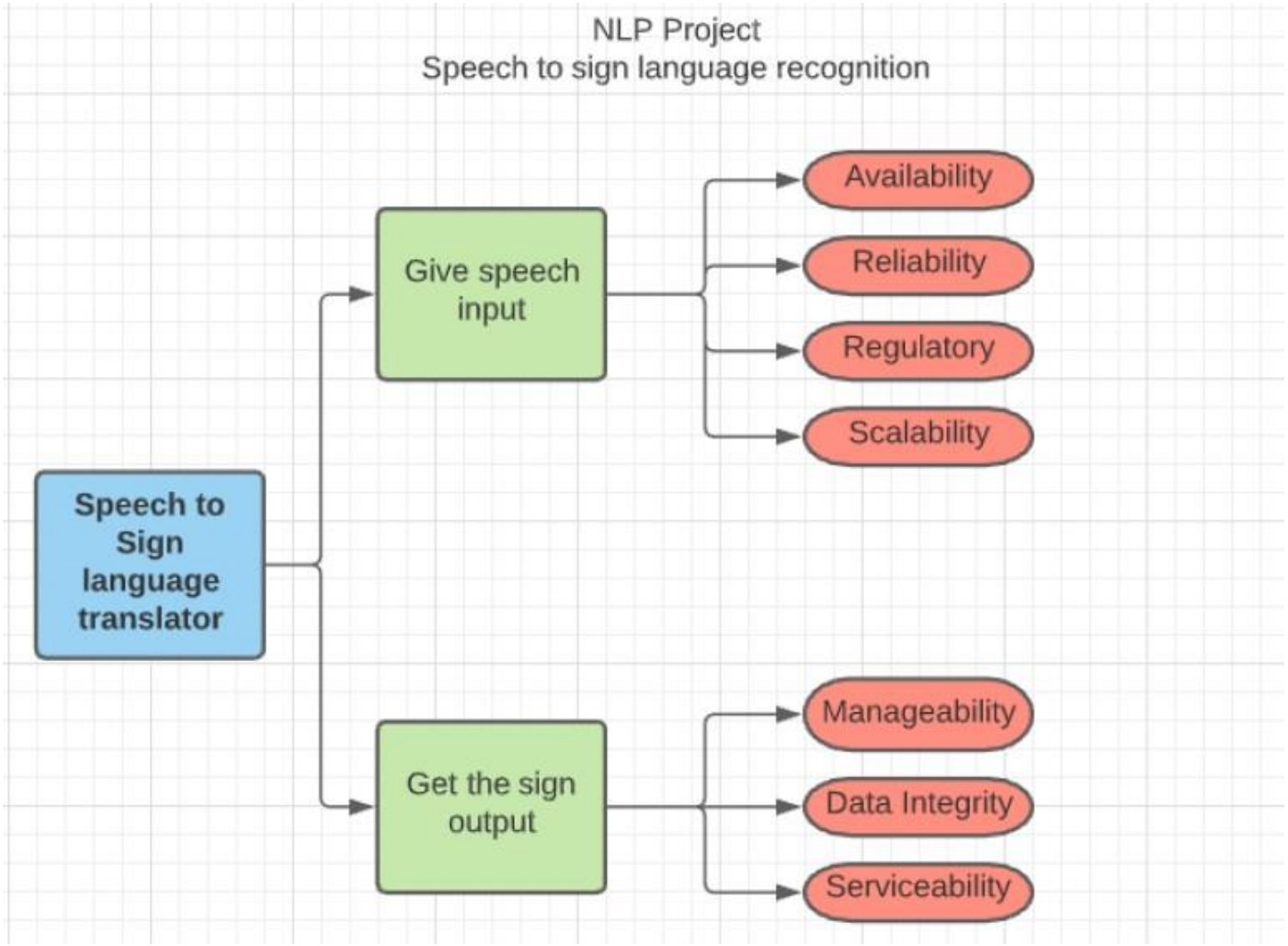


Figure 2. Functional and Non-Functional Requirements

5.4 Stakeholders and USE Case Diagram:

1. Project Managers
2. Project Sponsors
3. Coders
4. Executives
5. Programmers
6. Software Developers

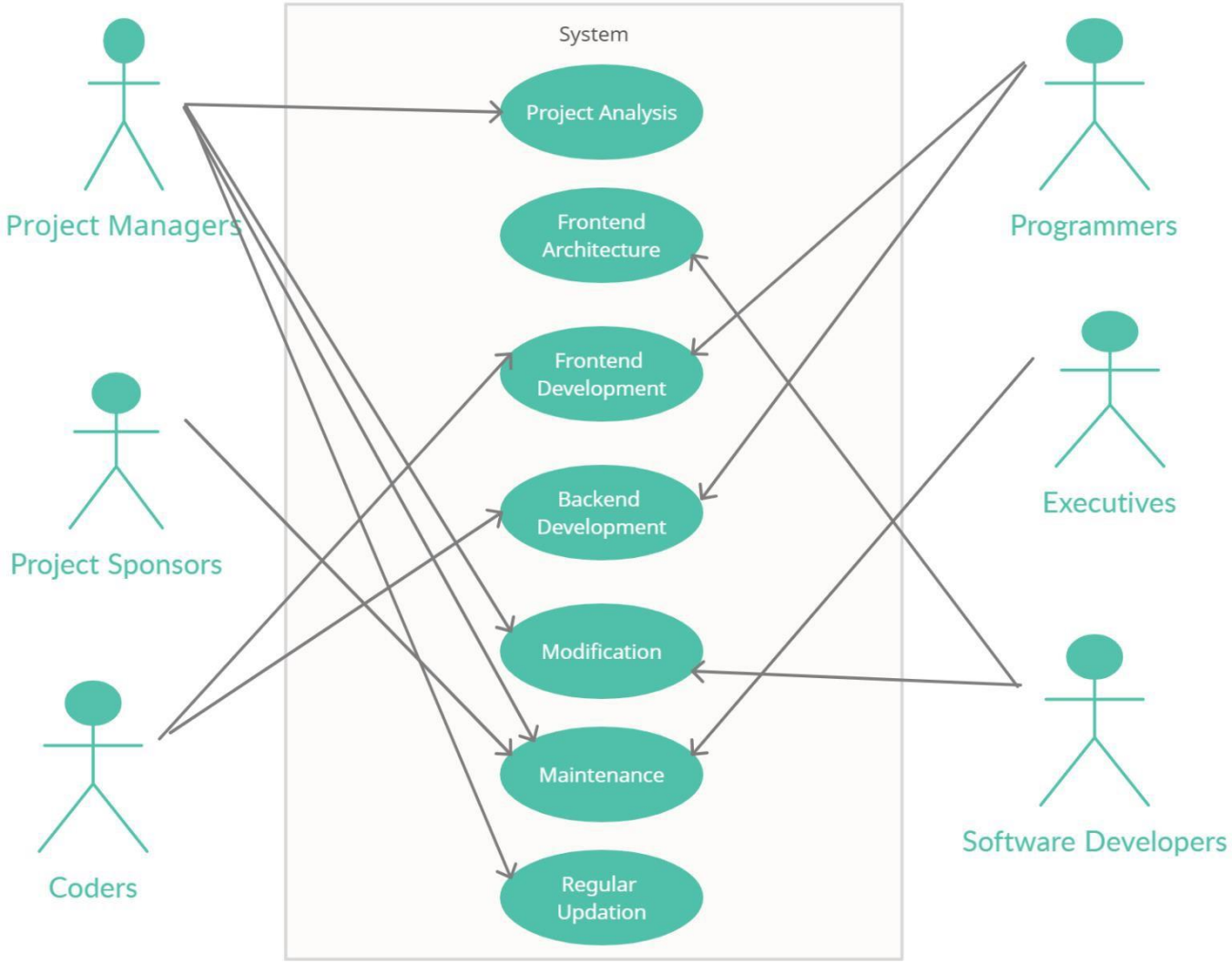


Figure 3. USE Case Diagram

5.5 Block Diagrams:

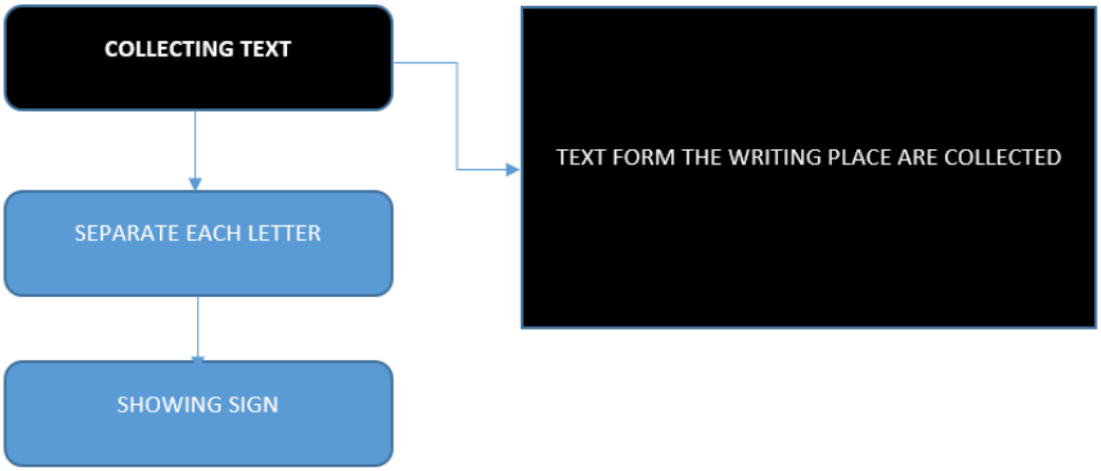


Figure 4. Text Collection

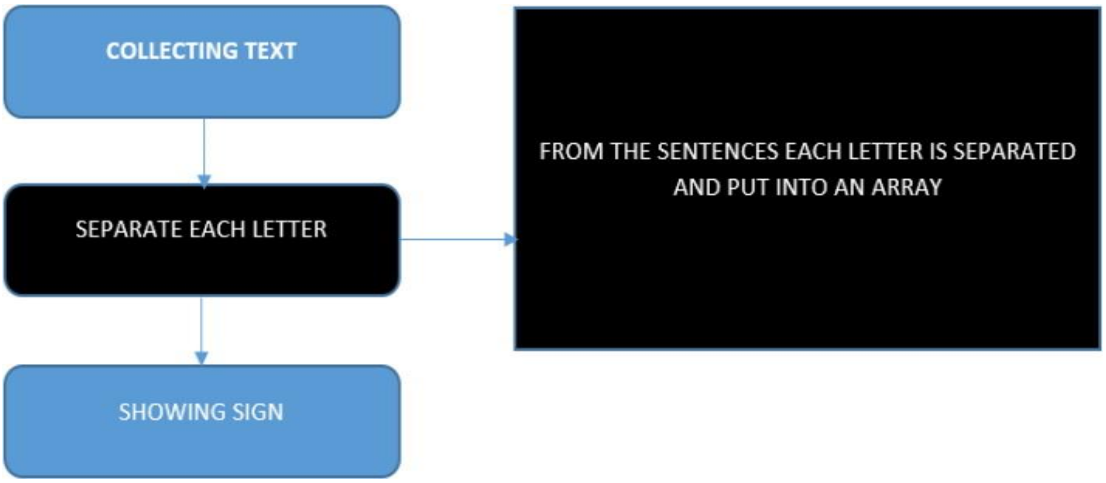


Figure 5. Text Separation

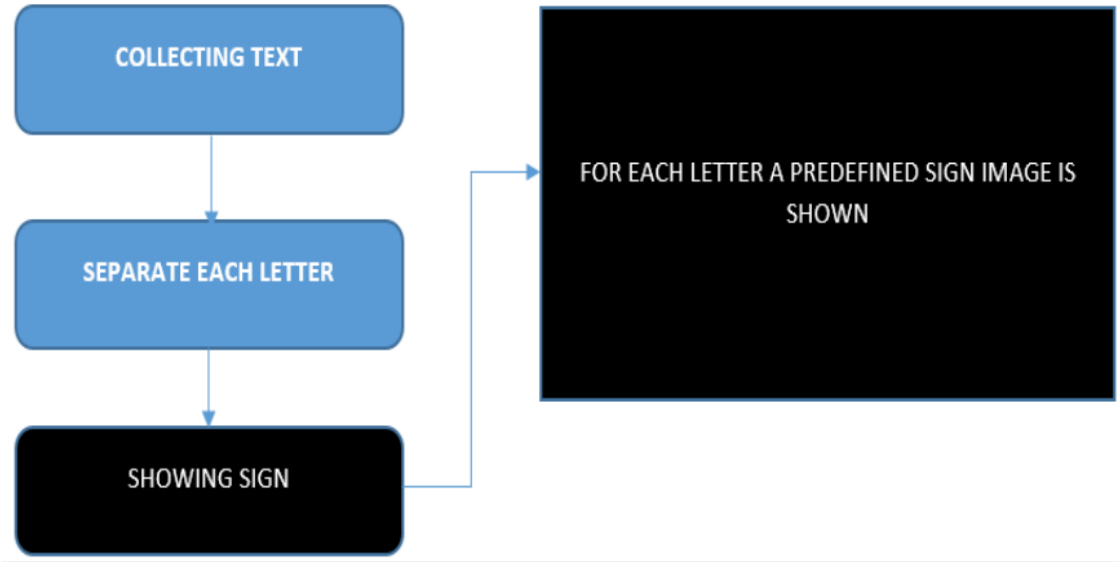


Figure 6. Text conversion to Sign Language

5.6 Predefined Gestures:

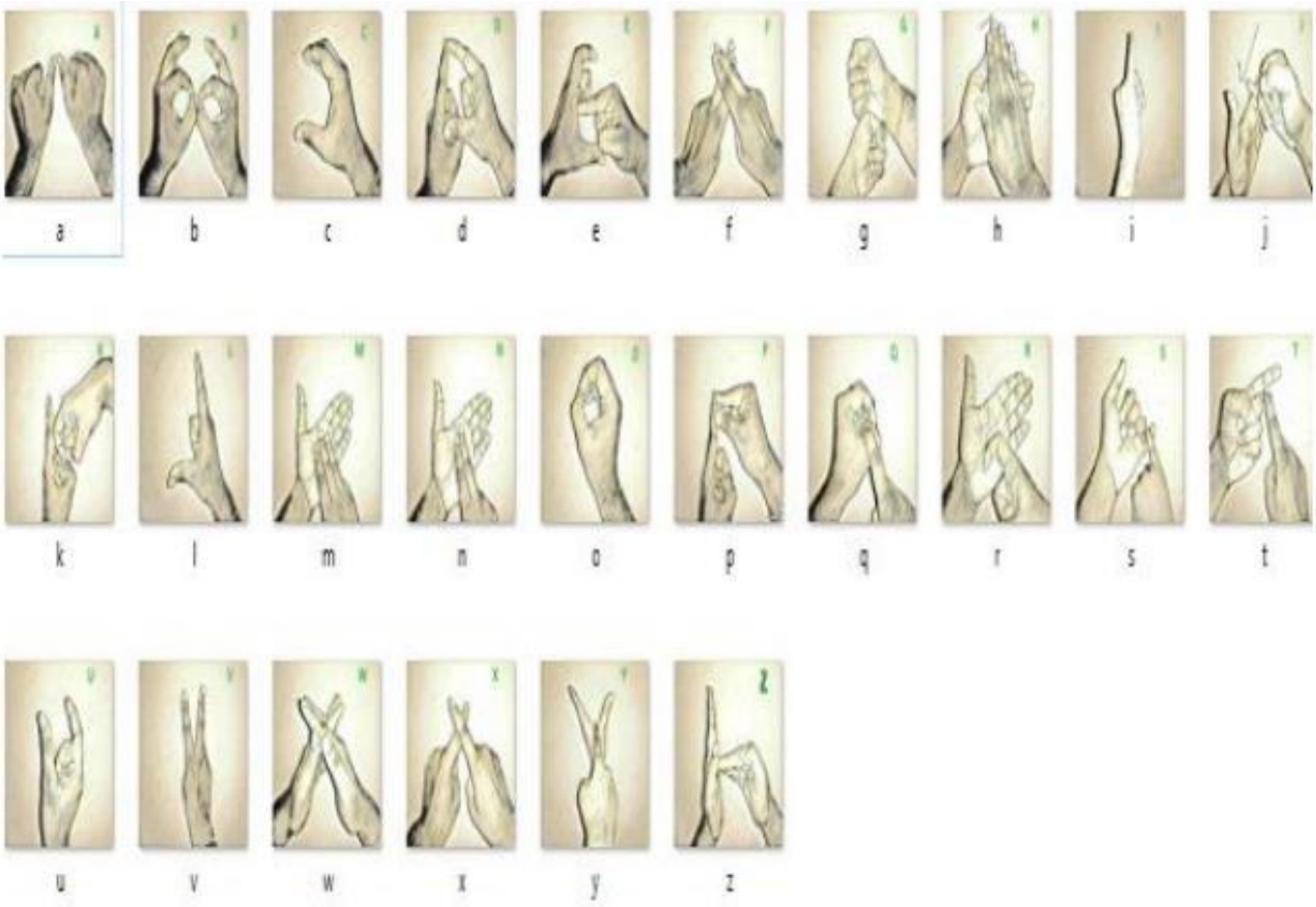


Figure 7. Hand Signs for each English alphabet

5.7 Procedure:

5.7.1 Audio to Text Conversion:

1. Audio input is taken using python PyAudio module.
2. Conversion of audio to text using microphone
3. Dependency parser is used for analyzing grammar of the sentence and obtaining relationship between words.

5.7.2 Text to Sign Language:

1. Speech recognition using Google Speech API.
2. Text Preprocessing using NLP.
3. Dictionary based Machine Translation.
4. ISL Generator: ISL of input sentence using ISL grammar rules.
5. Generation of Sign language with signing Avatar.

6.Implementation:

6.1 Code:

```
import speech_recognition as sr
import numpy as np
import matplotlib.pyplot as plt
import cv2
from easygui import *
import os
from PIL import Image, ImageTk
from itertools import count import
tkinter as tk
import string
#import selecting
# obtain audio from the microphone
def func():
    r = sr.Recognizer()
    isl_gif=['any questions', 'are you angry', 'are you busy', 'are you hungry', 'are you sick','be careful',
        'can we meet tomorrow', 'did you book tickets', 'did you finish homework', 'do yougo to office', 'do you have money',
        'do you want something to drink', 'do you want tea or coffee', 'do you watch TV', 'dont worry', 'flower is beautiful',
        'good afternoon', 'good evening', 'good morning', 'good night', 'good question', 'had your lunch', 'happy journey',
        'hello what is your name', 'how many people are there in your family', 'i am a clerk', 'i am bore doing nothing',
         'i am fine', 'i am sorry', 'i am thinking', 'i am tired', 'i dont understand anything', 'igo to a theatre', 'i love to shop',
        'i had to say something but i forgot', 'i have headache', 'i like pink colour', 'i live in nagpur', 'lets go for lunch', 'my mother is a homemaker',
        'my name is john', 'nice to meet you', 'no smoking please', 'open the door', 'please call me later',
        'please clean the room', 'please give me your pen', 'please use dustbin dont throw garbage', 'please wait for sometime', 'shall I help you',
```

'shall we go together tommorow', 'sign language interpreter', 'sit down', 'stand up', 'take care', 'there was traffic jam', 'wait I am thinking',

'what are you doing', 'what is the problem', 'what is todays date', 'what is your father do', 'what is your job',

'what is your mobile number', 'what is your name', 'whats up', 'when is your interview', 'when we will go', 'where do you stay',

'where is the bathroom', 'where is the police station', 'you are wrong','address','agra','ahemdabad', 'all', 'april', 'assam', 'august', 'australia', 'badoda', 'banana', 'banaras', 'banglore',

'bihar','bihar','bridge','cat', 'chandigarh', 'chennai', 'christmas', 'church', 'clinic', 'coconut', 'crocodile','dasara',

'deaf', 'december', 'deer', 'delhi', 'dollar', 'duck', 'febuary', 'friday', 'fruits', 'glass', 'grapes', 'gujrat', 'hello',

'hindu', 'hyderabad', 'india', 'january', 'jesus', 'job', 'july', 'july', 'karnataka', 'kerala', 'krishna','litre', 'mango',

'may', 'mile', 'monday', 'mumbai', 'museum', 'muslim', 'nagpur', 'october', 'orange', 'pakistan', 'pass', 'police station',

'post office', 'pune', 'punjab', 'rajasthan', 'ram', 'restaurant', 'saturday', 'september', 'shop','sleep', 'southafrica',

'story', 'sunday', 'tamil nadu', 'temperature', 'temple', 'thursday', 'toilet', 'tomato', 'town', 'tuesday', 'usa', 'village',

'voice', 'wednesday', 'weight','please wait for sometime','what is your mobile number','whatare you doing','are you busy']

```
arr=['a','b','c','d','e','f','g','h','i','j','k','l','m','n','o','p','q','r', 's','t','u','v','w','x','y','z']
with sr.Microphone() as source:#
    image      = "signlang.png"
    # msg="HEARING IMPAIRMENT ASSISTANT"
    # choices = ["Live Voice","All Done!"]
    # reply    = buttonbox(msg,image=image,choices=choices)
    r.adjust_for_ambient_noise(source)
    i=0
    while True:
        print("I am Listening") audio
        = r.listen(source)
```

```
                # recognize speech using Sphinx
                try:
                        a=r.recognize_google(audio)a
                        = a.lower()
                        print('You Said: ' + a.lower())

                        for c in string.punctuation:
                            a= a.replace(c,"")

                        if(a.lower()=='goodbye' or a.lower()=='good bye' or a.lower()=='bye'):
                                print("oops!Time To say good bye")
                                break

                        elif(a.lower() in isl_gif):


                            class ImageLabel(tk.Label):
                                    """a label that displays images, and plays them if they are
gifs"""
                                    def load(self, im):
                                        if isinstance(im, str):
                                            im = Image.open(im)
                                        self.loc = 0
                                        self.frames = []

                                        try:
                                            for i in count(1):
                                                self.frames.append(ImageTk.PhotoImage(im.copy()))
                                                im.seek(i)
                                        except EOFError:
                                            pass

                                        try:
```

```
                    self.delay = im.info['duration']
                except:
                    self.delay = 100

                if len(self.frames) == 1:
                    self.config(image=self.frames[0])
                else:
                    self.next_frame()

            def unload(self):
                self.config(image=None)
                self.frames = None

            def next_frame(self):
                if self.frames:
                    self.loc += 1
                    self.loc %= len(self.frames)
                    self.config(image=self.frames[self.loc])
                    self.after(self.delay, self.next_frame)
        root = tk.Tk()
        lbl = ImageLabel(root)
        lbl.pack()
        lbl.load(r'ISL_Gifs/{0}.gif'.format(a.lower()))
        root.mainloop()
    else:
        for i in range(len(a)):
                    if(a[i] in arr):

                        ImageAddress =  'letters/'+a[i]+'.jpg' ImageItself
                        = Image.open(ImageAddress)
                        ImageNumpyFormat = np.asarray(ImageItself)
                        plt.imshow(ImageNumpyFormat)
```

```
                                        plt.draw()
                                        plt.pause(0.8)
                                    else:
                                        continue


                    except:
                        print(" ")
                    plt.close()
while 1:

  image    = "signlang.png"
  msg="HEARING IMPAIRMENT ASSISTANT"
  choices = ["Live Voice","All Done!"]
  reply    = buttonbox(msg,image=image,choices=choices)if
  reply ==choices[0]:
      func()
  if reply == choices[1]:
      quit()
```

**Separate dataset consisting of all GIFs has been manually created for this project. Along with this, another dataset with predefined gestures corresponding to every alphabet has also been made.

**A picture from external sources was loaded into the software for GUI Design.

6.2 GUI Design:

Figure 8. GUI Design for front end by EasyGui

6.3 Speech input:

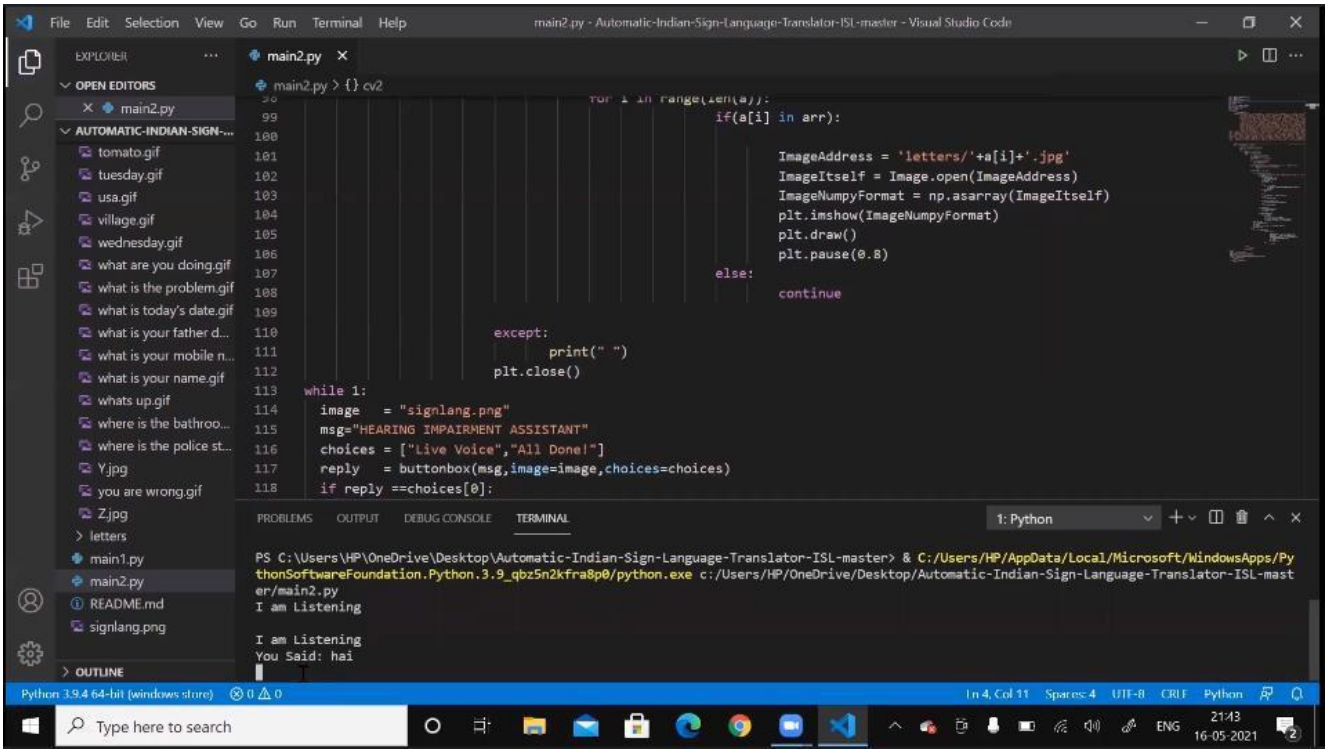


Figure 9. Speech Input taken through microphone using PyAudio package

6.4 Conversion from Audio to text:

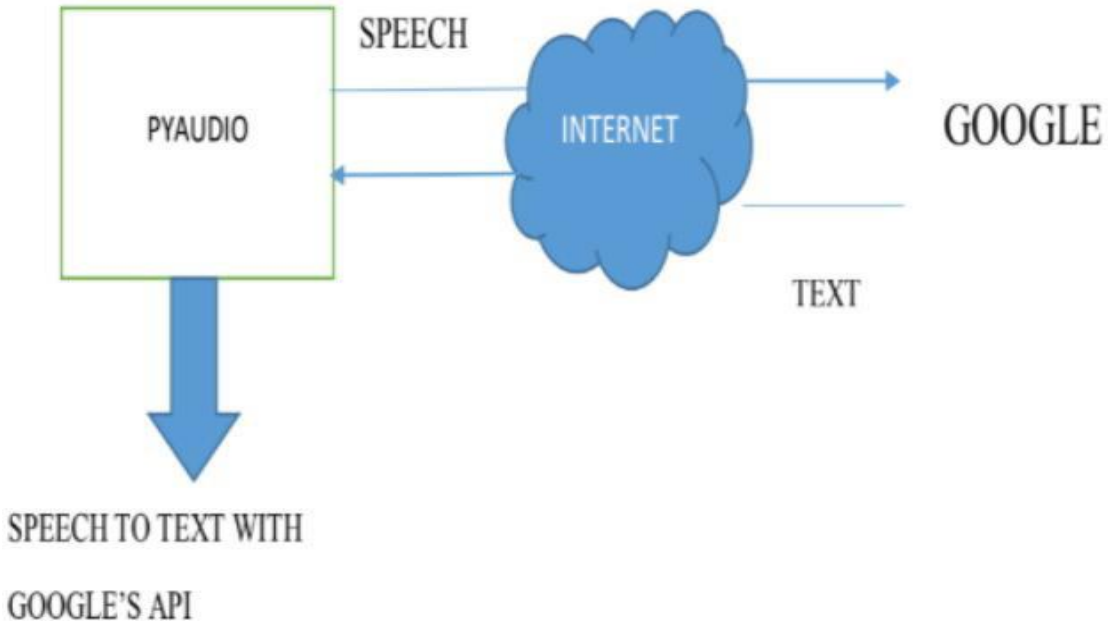


Figure 10. Google Speech-to-Text (Converts audio to text by applying neural network models in an easy-to-use API)

6.5 Preprocessing of text using NLP:

Machine can understand binary language (i.e., 0 and 1) only. To make the machine understand human language, NLP was introduced. Natural Language Processing is the ability of the machine where it processes the text that was said and structures it. NLP creates an algorithm that translates text into word by labelling them based on the position and function of the words in the sentences. Human language is converted meaningfully into a numerical form. This allows computers to understand the nuances implicitly encoded into our language. Text preprocessing consists of three things - Tokenization, Normalization and Noiseremoval.

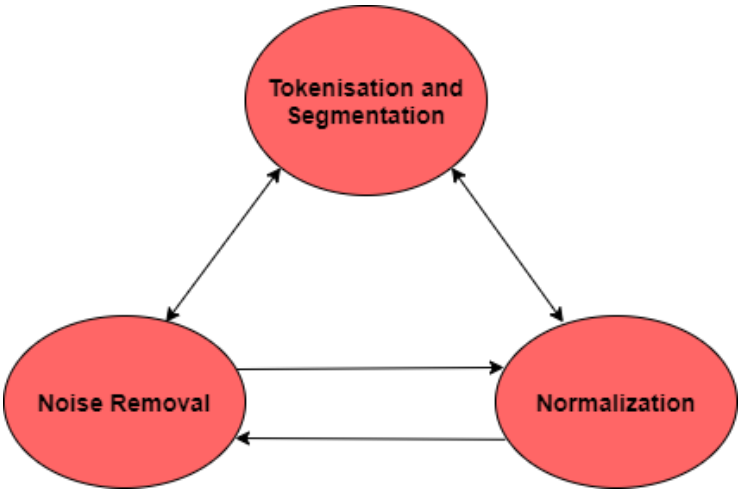


Figure 11. Text Preprocessing

NLP helps the machine in understanding human language through the following steps:

1. We give audio as input to the machine.
2. The machine records that audio input.
3. Then machine translates the audio into text and displays it on the screen.
4. The NLP system parses the text into components; understand the context of the conversation and the intention of the person.
5. The machine decides which command to be executed, based on the results of NLP.

7.Results:

7.1 Letter-by-letter output:

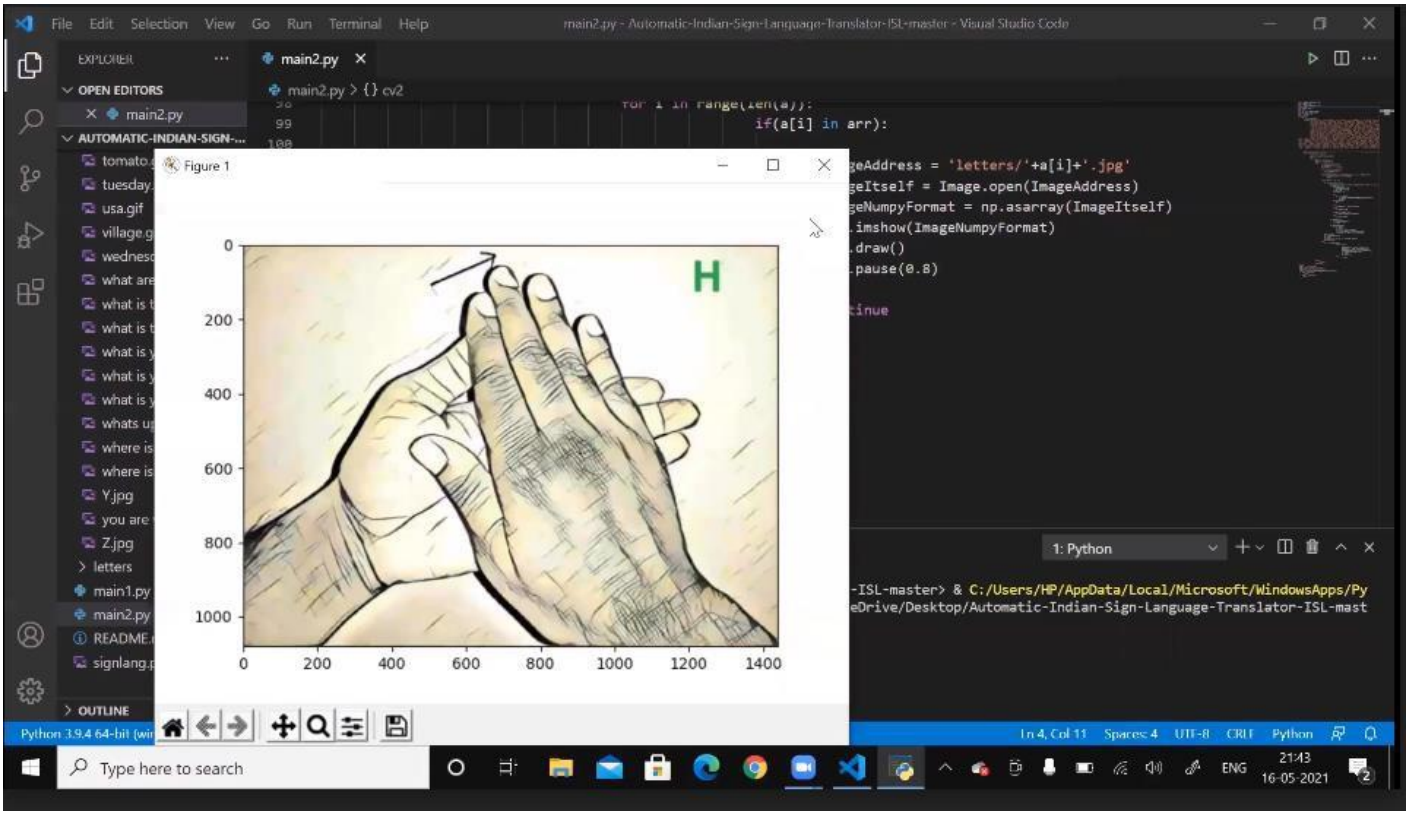


Figure 12. Word by word output (Letter “H”)

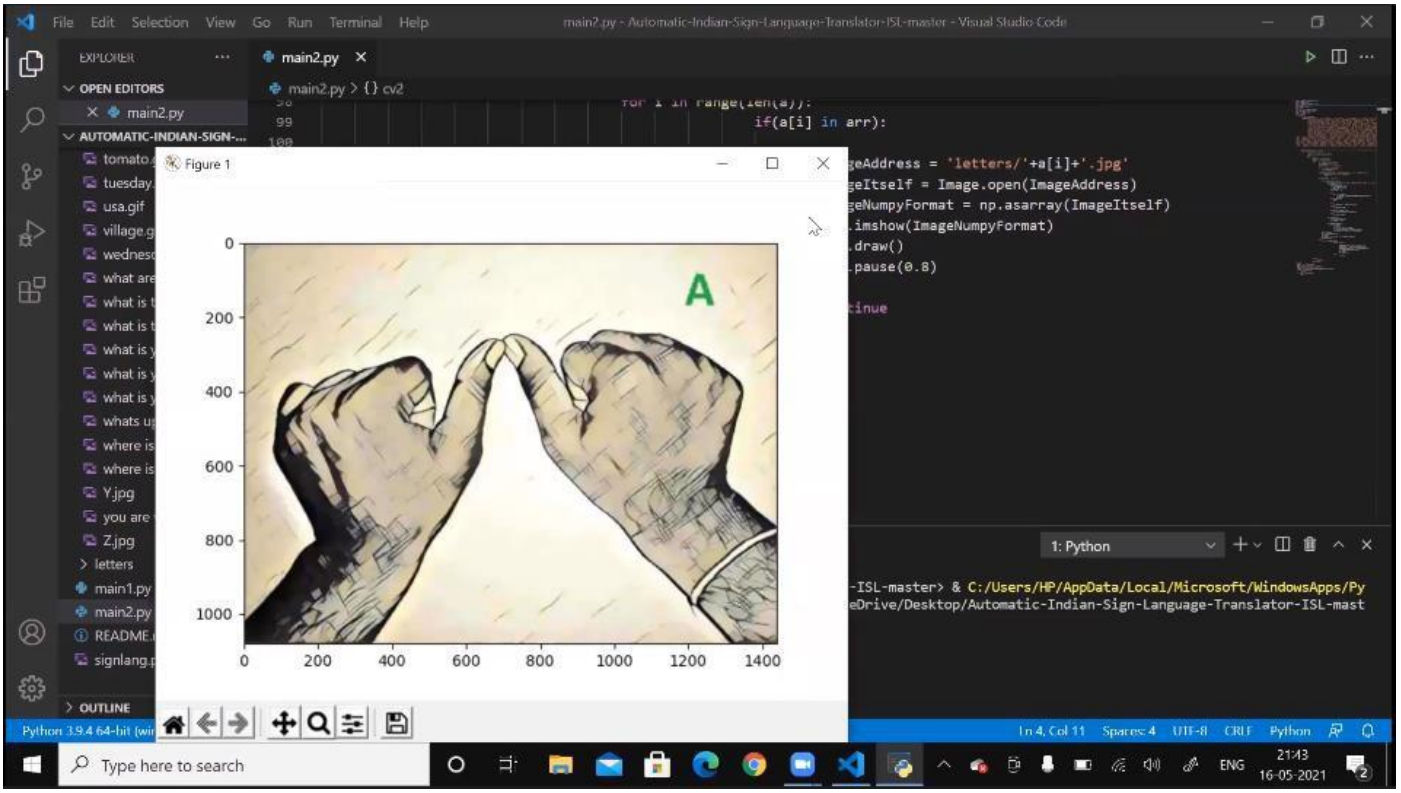


Figure 13. Word by word output (Letter “A”)

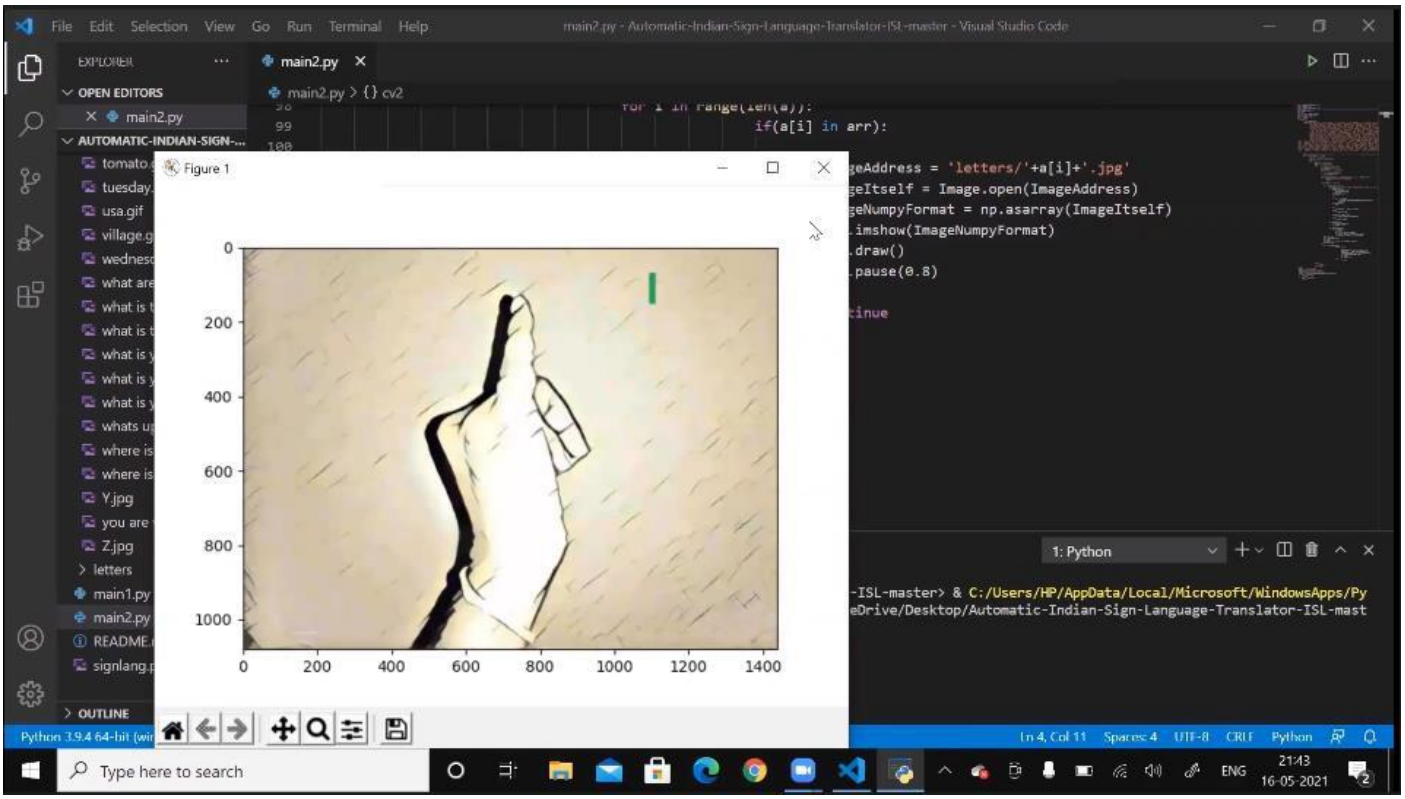


Figure 14. Word by word output (Letter “I”)

7.2 GIF Output:

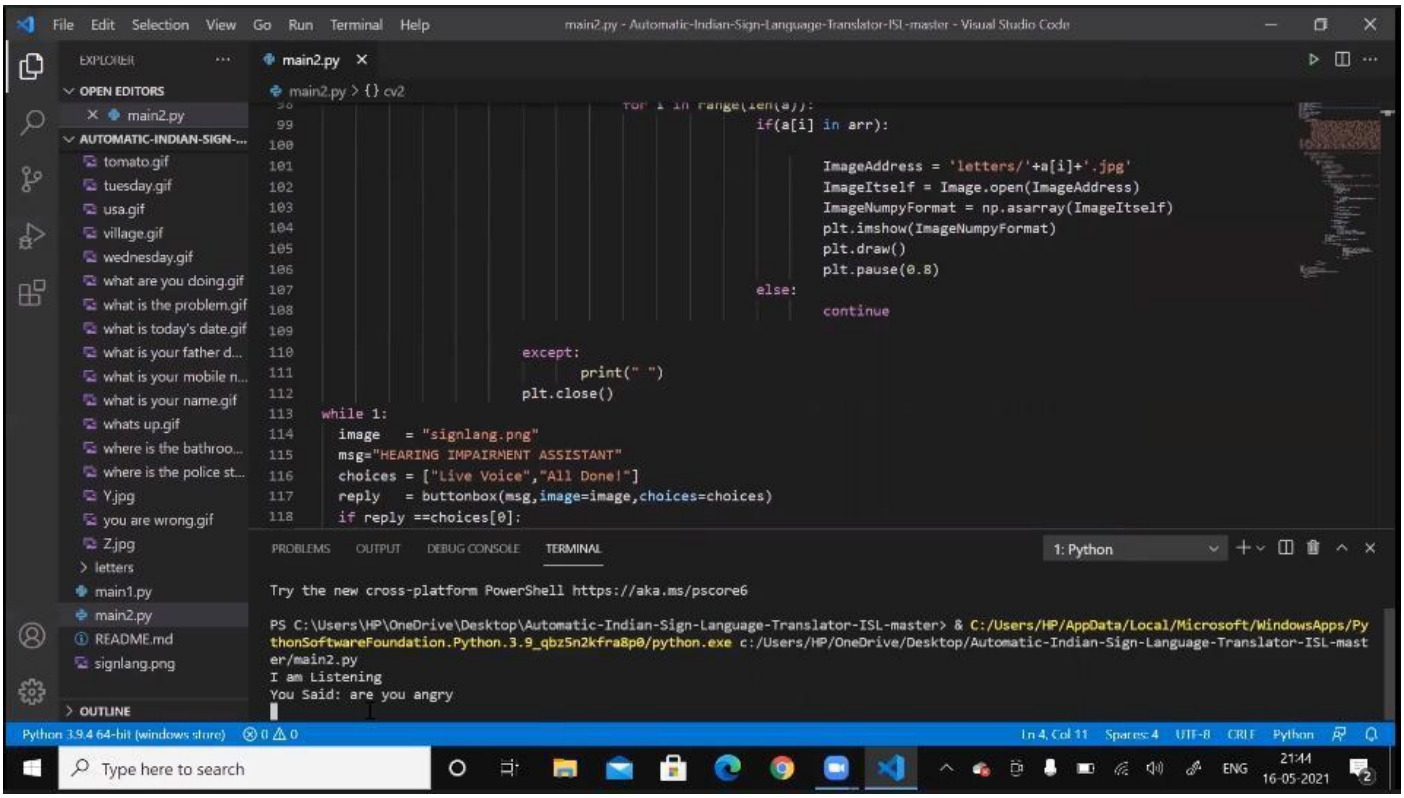


Figure 15. Speech input of an existing phrase

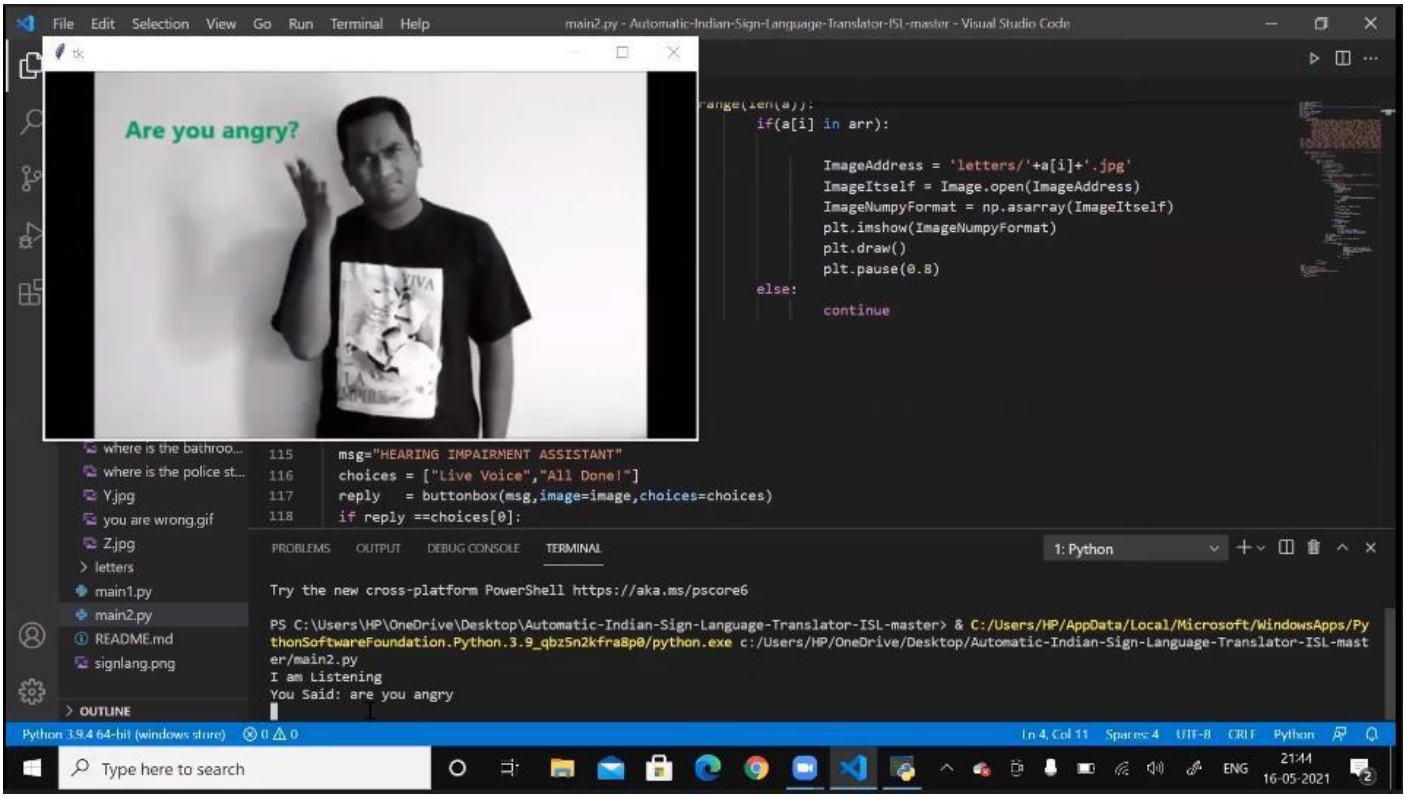
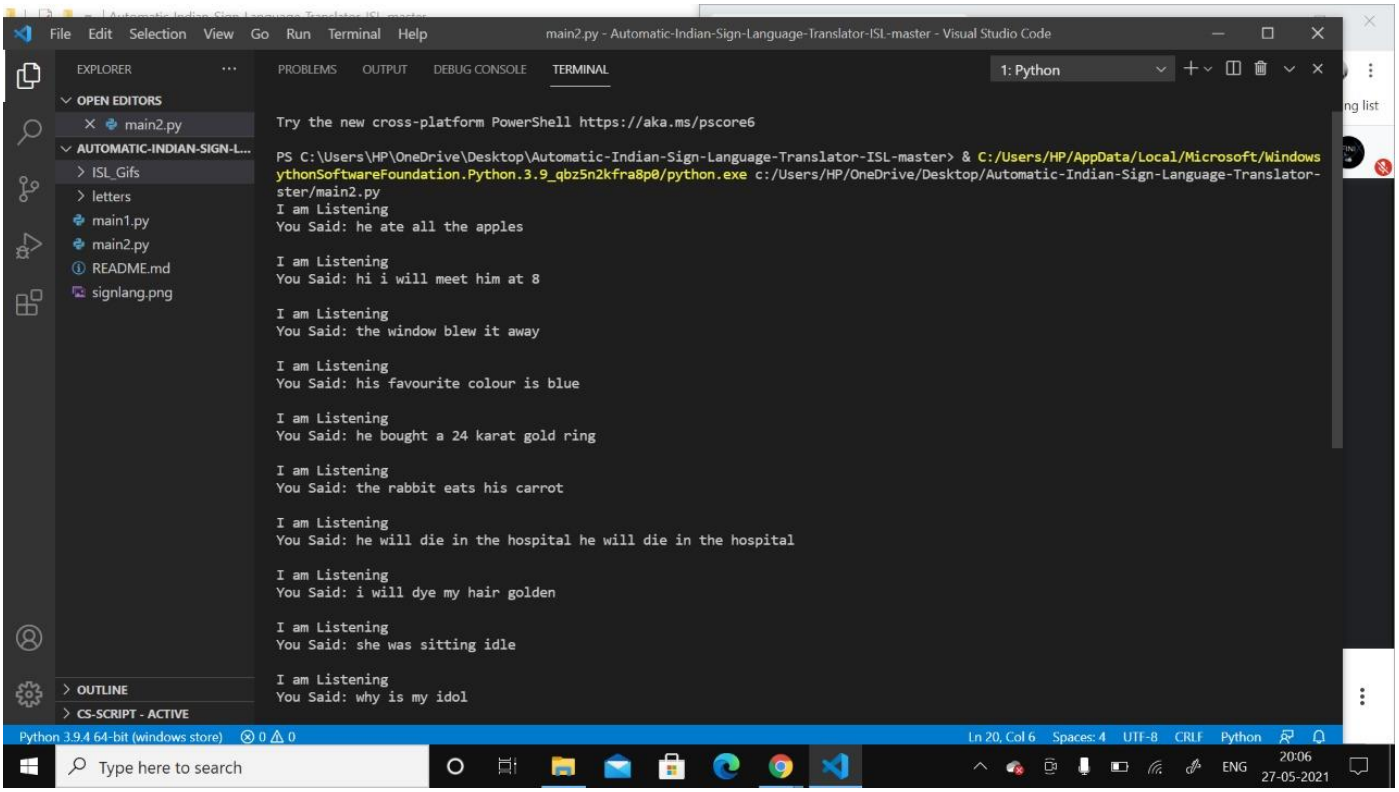


Figure 16. GIF Output

7.3 Test cases for Homophones:



Test Case ID	Test Objective	Test Data	Expected Result	Actual Result	Pass/Fail
1001	Checking speech recognition for homophones	Sentence 1	ate	ate	Pass
1002	Checking speech recognition for homophones	Sentence 2	8	8	Pass
1003	Checking speech recognition for homophones	Sentence 3	blew	blew	Pass
1004	Checking speech recognition for homophones	Sentence 4	blue	blue	Pass
1005	Checking speech recognition for homophones	Sentence 5	karat	karat	Pass
1006	Checking speech recognition for homophones	Sentence 6	carrot	carrot	Pass
1007	Checking speech recognition for homophones	Sentence 7	die	die	Pass
1008	Checking speech recognition for homophones	Sentence 8	dye	dye	Pass
1009	Checking speech recognition for homophones	Sentence 9	idle	idle	Pass
1010	Checking speech recognition for homophones	Sentence 10	idol	idol	Pass

We see a 100% accuracy when it comes to predicting the spelling of words that sound the same but have different meanings and are spelt differently. This is because Google Speech API uses n-gram model to predict words while processing. For e.g., in the sentence “He ate all his apples”, it doesn’t make sense to use “8” instead of “ate”. Therefore, because of n-gram model, Google Speech API predicts the correct spelling by checking the words preceding it.

The link of the paper discussing the use of n-gram models in Google Speech Recognition: [43819.pdf (googleusercontent.com)]

8.Utility of the application:

Form link: https://forms.gle/mS9sDSNmovapiHRE7

Total response: 126

Results:

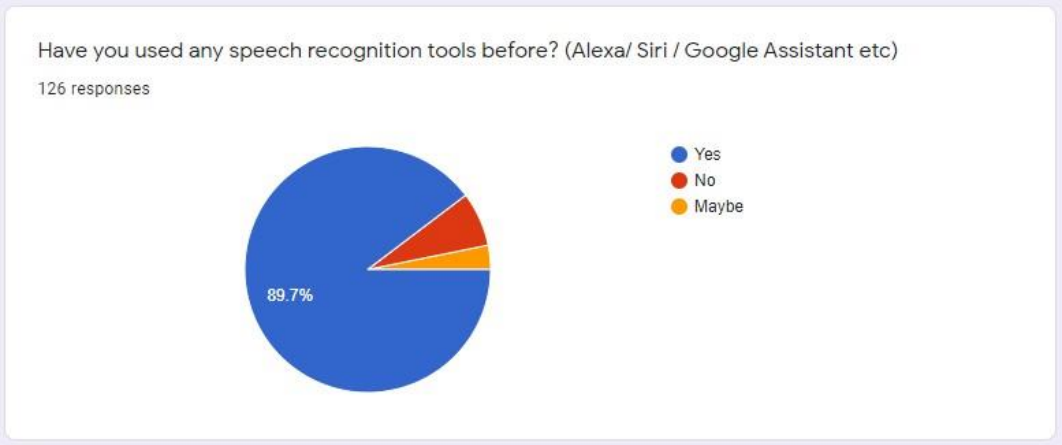


Figure 17

Insight: Most of the people have used a speech recognition tool before

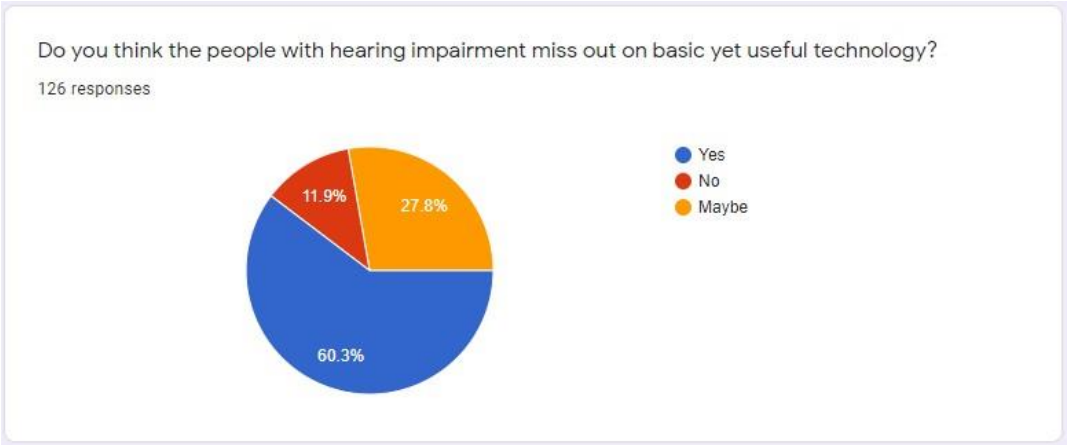


Figure 18

Insight: 81.1% of the people believe that people with hearing impairment miss out on basic but useful technology in some way or the other

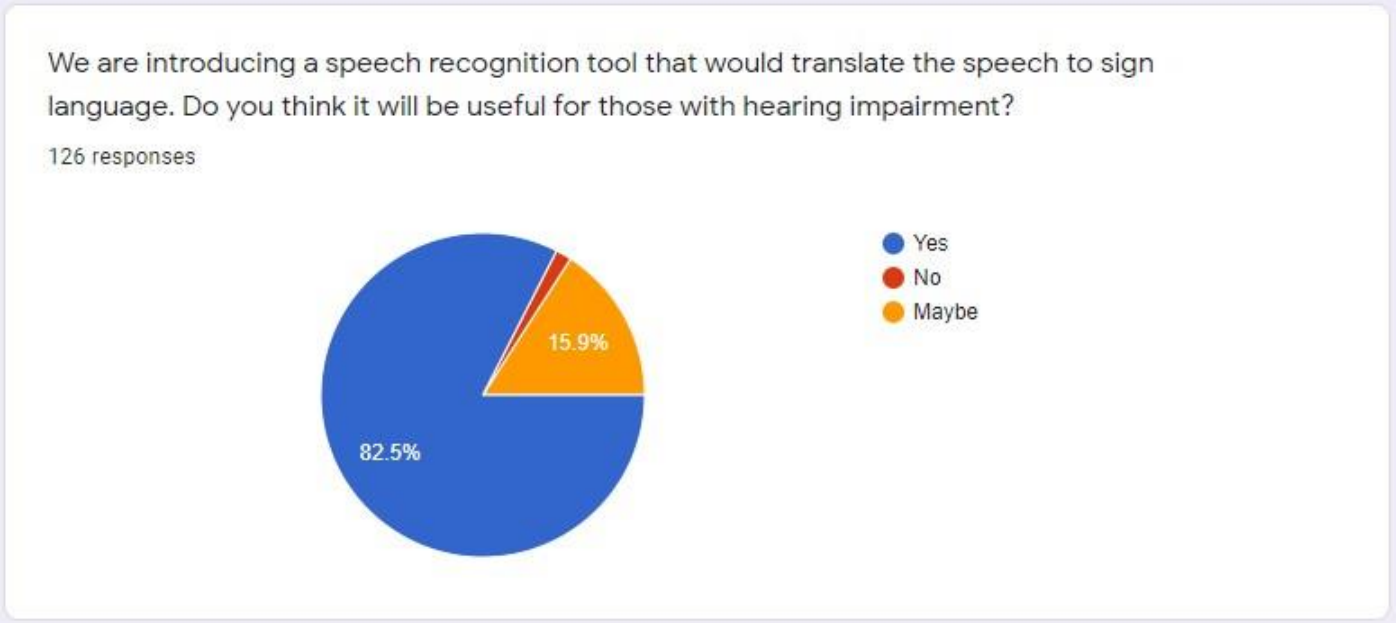


Figure 19

Insight: 98.4% of the people think that our tool will help people with hearing disability by making their lives easier through effective communication.

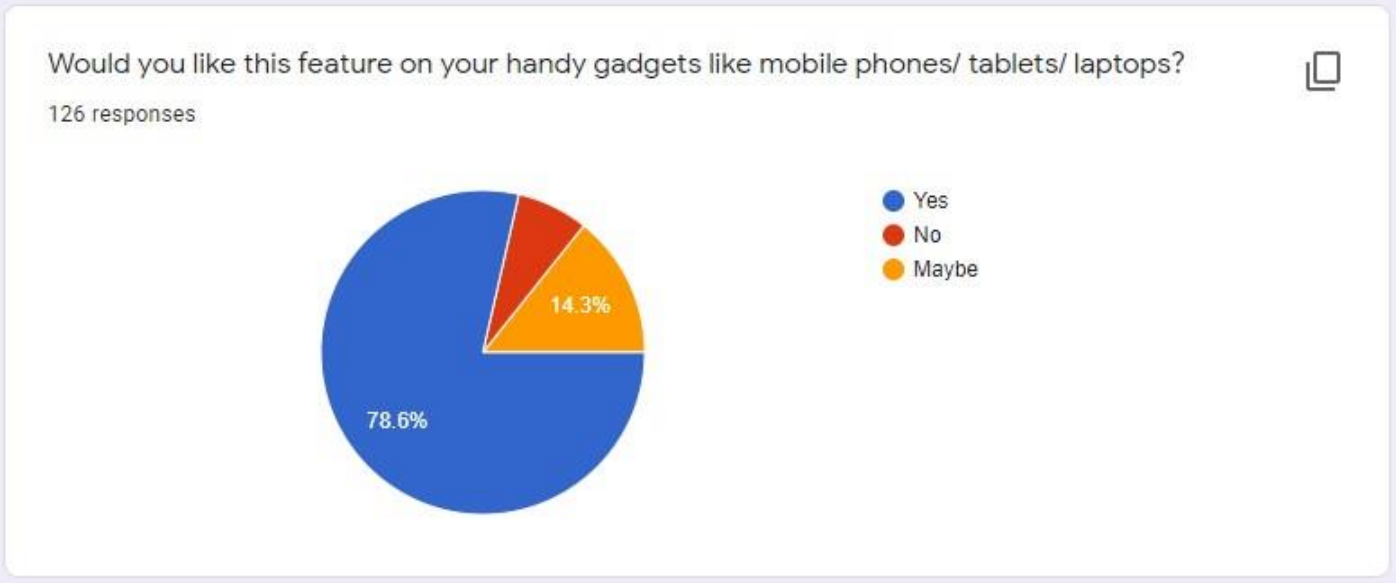


Figure 20

Insight: People largely believe that having such an app in phones will be helpful. It is the need of the hour so that people with hearing disabilities can cope up and live their lives as normal people with no communication barriers do.

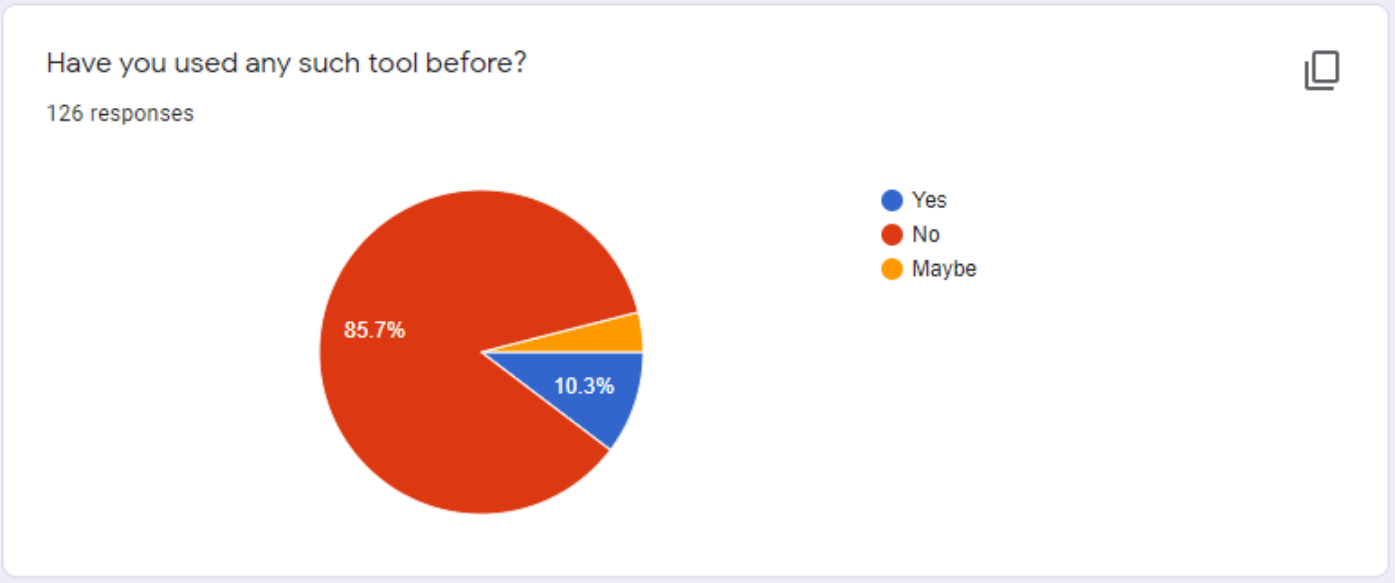


Figure 21

Insight: Mostly, people have never used such a tool which explains how necessary and relevant it is. It is our aim to make this tool available to everyone including normal people as well as those with hearing disability to bridge the communication gap between them.

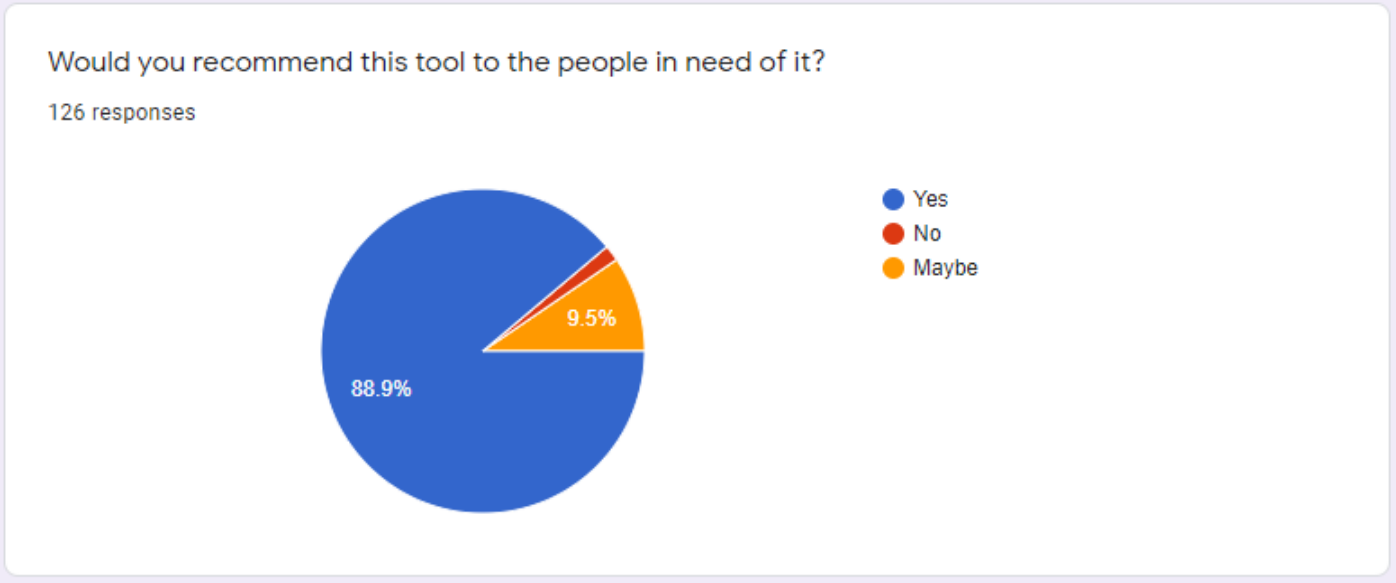


Figure 22

Insight: People largely recommend this tool to those with hearing issues, which shows it a general notion that a tool like this could be extremely important in the lives of deaf people.

9. Future Works:

Sign language translator is extremely useful in various areas. In schools, colleges, hospitals, universities, airports, courts, etc., anywhere anyone can use this system for understanding the sign language to communicate. It makes communication between a normal hearing person and a person with hearing disability easier.

The future work is to develop an application that can be used by news channel. A small screen at corner of the screen could display sign-language visuals so that a deaf person is also able to listen to the news. At present, only DD news is using this kind of presentation technique, but they show a person manually translating the newsreader and conveying the information through hand signals. Therefore, by using this application, we can convey news faster and without error.

We also look forward to expanding the project by including facial expressions into the system. This would make the application more user-friendly.

10. References:

[1] Amit Kumar Shinde and Ramesh Khagalkar “sign language to text and vice versa recognition using computer vision in Marathi” International journal of computer Application (0975-8887) National conference on advanced on computing (NCAC 2015).

[2] Neha Poddar, Shrushti Rao, Shruti Sawant, Vrushali Somavanshi, Prof.Sumita Chandak "Study of Sign Language Translation using Gesture Recognition" International Journal of Advanced Research in Computer and Communication Engineering Vol. 4, Issue 2, February 2015.

[3] Deaf Mute Communication Interpreter Anbarasi Rajamohan, Hemavathy R., Dhanalakshmi M. (Department of B.M.E., Sri Sivasubramania Nadar College of Engineering, Chennai, Tamil Nadu, India). International Journal of Scientific Engineering and Technology (ISSN: 2277-1581) Volume 2 Issue 5, pp: 336-341 1 May 2013.

[4] Vajjarapu Lavanya, Akulapravin, And M.S., Madhan Mohan “Hand Gesture Recognition and Voice Conversion System Using Sign Language Transcription System" ISSN: 2230-7109 (Online) | ISSN: 2230-9543 (Print) IJECT Vol. 5, Oct - Dec 2014.

[5] Natural Language Processing: State of The Art, Current Trends and Challenges Diksha Khurana1, Aditya Koli1, Kiran Khatter1,2 and Sukhdev Singh1,2 1Department of Computer Science and Engineering Manav Rachna International University, Faridabad-121004, India 2Accendere Knowledge Management Services Pvt. Ltd., India.

[6] https://www.slideshare.net/mobile/madhuriyellapu/signlan guage-translator-ieee-power-point

11. Questions:

Q1. What are the crucial features of your project?

The Application has several crucial features:

1. GUI Design for front end
2. Speech input
3. Speech to text conversion
4. Letter by letter translation to the sign language in case of an unfamiliar phrase
5. GIF output for a familiar phrase.

Q2. What are main NLP steps (very specific) used?

Audio to Text Conversion:

1. Audio input is taken using python PyAudio module.
2. Conversion of audio to text using microphone.
3. Dependency parser is used for analyzing grammar of the sentence and obtaining relationship between words.

Text to Sign Language:

1. Speech recognition using Google Speech API.
2. Text Preprocessing using NLP.
3. Dictionary based Machine Translation.
4. ISL Generator: ISL of input sentence using ISL grammar rules.
5. Generation of Sign language with signing Avatar.

Refer to figure 4, 5, 6 on page 7 and figure 10, 11 on page 15 for more.

Q3. What is the relevance of this project (Motivation)?

According to a study by the World Health Organization in 2014, 5% of the world's population, or about 360 million people suffer from hearing disability. Due to this, these people miss out on basic activities and often have to stay dependent on someone else for survival. They also struggle to communicate with others that do not understand sign language.

The motivation behind this project was to utilize our knowledge of Computer Science and build an application that would make communication for these disabled people much easier. This application can especially be used by people with no hearing disability to convey their thoughts and communicate with deaf people easily.

Q4. Which are the main tools used?

The following tools have bees used in this project:

1. Front-end using EasyGui
2. Speech as input through microphone using PyAudio
3. Speech recognition using Google Speech API
4. Text Preprocessing using NLP
5. Dictionary based Speech to Sign language Translation
6. Implementation in Visual Studio

Q5. Which are the similar tools available?

Alternate tools are as follows:

1. For building GUI for front end, Enaml, Tkinter, Uwrid. Flexx, etc. could be used.
2. For taking speech input and recording audio, python-sounddevice is also a good choice.
3. Instead of Visual Studio, NetBeans could also be used for implementation.

Q6. Why did you prefer those tools?

1. EasyGUI is a better choice as compared to these other tools because it is a modern display-based embedded system that will save precious development time, because much of the time-consuming development work on visual components are moved from traditional direct coding to an efficient visual PC application.
2. PyAudio is a Python library for audio feature extraction, classification, segmentation and applications. This is better than its counterparts because it has been in the market or a longer time and has great utility and reliability.
3. Visual Studio is a state-of-the-art tools and services that can be used to create apps for devices, the cloud, and everything in between. It is a suite of component-based software development tools and other technologies for building powerful, high-performance applications. IntelliSense and UI are also the reasons behind choosing Visual Studio. Intuit**,** Starbucks, and Yahoo! are some of the popular companies that already use Visual Studio. Visual Studio also has a broader approval, being mentioned in 676 company stacks & 1009 developers stacks; compared to NetBeans IDE, which is listed in 62 company stacks and 46 developer stacks.

Q7. What are challenges you faced during project implementation?

The biggest challenge that we faced was collecting the GIFs of common phrases in order to create a dataset that has a sufficient number of daily-use sentences. Lesser number of sentences would render the GIF feature useless. Other than this, the coding part was also tricky because in a project as big as this, even a single mistake here and there would mess up the entire project.

Q8. How did you resolve/ overcome those challenges?

We overcame these challenges by compartmentalizing and distributing work equally so that no single person gets burdened with the project. For the GIF part, we searched on google for common phrases and found out their translations into sign languages as GIFs. It took a lot of time but we finally managed to have sufficient number of GIFs. As far as the coding part is concerned, we broke the tasks and assigned them to each and every person. Everyone came up with their snippet of code and we finally put it all together. We definitely had to make some edits and correct some errors, but finally we were able to come up with a fully functioning python code.

Q9. What are the social impacts of your project?

Our project has the potential to impact the society positively. As mentioned earlier, millions of people around the world suffer from hearing disabilities. In such a case, having an application like this would be instrumental in making communication easier for them. Apart from this, our project would also increase sensitization towards those with hearing disabilities and spread awareness among the masses about their plight. This way, even the deaf people would be able to make their voices heard and become an active part of the society instead of just being neglected.

Q10. What is future enhancement?

The future enhancements include adding a feature to the application that would recognize expressions as well as the hand signals of the sign language and give a text and speech output. At the moment, only those with no hearing disability can use this app to communicate with those who do. But, in order to have a back-and-forth communication without any problem, we need an app that recognizes hand signals as well. Other than this, we also wish to enhance the ability of the application and provide it to news channels so that deaf people can also tune in to news channels and get updated. Currently, only DD news has this feature.

Q11. Contributions of each Member?

1. Collection of dataset and relevant libraries – Aviral Goyal and Mohit Keyal
2. Implementation of sign language images – Tarang Garg and Aviral Goyal
3. Implementation of sign language gifs – Nimish K Aggarwal and Tarang Garg
4. Implementation of GUI – Vrushali Deshmukh and Nimish K Aggarwal
5. Conducting test for homophones and utility survey – Vrushali Deshmukh and Mohit Keyal

12. Videos:

PPT Presentation –

https://drive.google.com/file/d/19yTINAhAs9sHyruT_Y210fxudiLzmeG6/view?usp=sharing

Document Presentation (21 minutes) –

https://drive.google.com/file/d/1dtg-7a5UroYdaaP2tXrtm8dCFZ5WPckK/view?usp=sharing

Implementation (6.5 minutes) –

https://drive.google.com/file/d/1m2EzFieBDmc0PqRz_NywuiWlYim0_VzU/view?usp=sharing