

# BUNDELKHAND INSTITUTE OF ENGINEERING AND TECHNOLOGY



**ELECTRONICS AND COMMUNICATION ENGINEERING**

**PROJECT REPORT**

**TOPIC:SPEECH RECOGNISION USING PYTHON**

**Submitted by**

NIKHIL KUMAR( 2004331033)

**Submitted to**

Mr.MUDRESH TRIPATHI

## **ACKNOWLEDGEMENT**

The Satisfaction and euphoria that accompany the successful completion of this seminar report on “**SPEECH RECOGNITION USING PYTHON** “ would be incomplete without the mention of the people who made it possible and under whose constant guide the seminar report was completed.

We profusely thank **Mr. Mudersh Tripathi** for his encouragement and support throughout the project.

I am greatly indebted to **Prof. Deepak Nagaria , Head of Department of Electronics and Communication Engineering** for providing us unhindered encouragement in carrying out the seminar report successfully.

Finally our sincere thanks to all other members who were helpful directly and indirectly in the completion of our project.

## **Declaration**

We hereby declare that the dissertation entitled **“SPEECH RECOGNITION USING PYTHON”** has been submitted by us in partial fulfilment of the requirement for the award of **BACHELOR OF TECHNOLOGY** in **ELECTRONICS And COMMUNICATION ENGINEERING** from Bundelkhand Institute of engineering and Technology affiliated to Abdul Kalam Technical University and is record of bonafied work carried out by us under the guidance of **Mr. Mudresh Tripathi**.

# **Introduction**

Speech is the most common means of communication around the world. Most of the population in the world relies on speech to communicate with each other. Suppose we are building a model and instead of a written approach we want our system to respond to speech, it becomes fairly difficult and requires a lot of data to be processed. A speech recognition system overcomes this barrier by translating speech to text. we will go through the speech recognition.

# How Speech Recognition Works?

Speech recognition system basically translates the spoken utterances to text. There are various real life examples of speech recognition system. For example- siri, which takes the speech as input and translates it into text.

The advantage of using a speech recognition system is that it overcomes the barrier of literacy. A speech recognition model can serve both literate and illiterate audience as well, since it focuses on spoken utterances.

We can also make an inventory of all the endangered languages around the world using a speech recognition system. While it looks pretty intriguing and not complex at all, a speech recognition system faces a lot of challenges in the making.

# How To Install SpeechRecognition In Python?

To install SpeechRecognition package in python, run the following command in the terminal and it will be installed on your system.

```
$ pip install SpeechRecognition
```

Another approach to this, can be adding the package from the project interpreter if you are using **pycharm**.

The package has a Recognizer class which is basically where the magic happens. It is basically a class which is used to recognize the speech. Following are seven methods which can read various audio sources using different APIs.

- recognize\_bing( )
- recognize\_google( )
- recognize\_google\_cloud( )
- recognize\_houndify( )
- recognize\_ibm( )
- recognize\_wit( )
- recognize\_sphinx( )

Now, recognize\_sphinx can be used to run the speech recognition system offline as well. It requires the installation of Pocketsphinx.

```
1import speechrecognition as sr
2
3#instance of recognizer class
4r = sr.Recognizer()
```

## Taking Input From Microphones

To use the microphones, we will have to install pyaudio module as well. We use the microphone class to get the input speech from the microphone instead of any other input method like an audio file.

For most of the projects, we can use the default microphones. But if you do not wish to use the default microphone, you can get the list of microphone names using the `list_microphone_names` method.

To capture the input from the microphone we use the `listen` method.

```
1  
2  
3import speechrecognition as sr  
4r = sr. Recognizer ()  
5with sr. Microphone () as source:  
6    audio = sr.listen (source)
```

## How To Install Pyaudio In Python?

To install Pyaudio in python, run the following command in the terminal or if you are using pycharm add the package from the project interpreter in the settings.

```
$ pip install PyAudio
```

## Use Case

We will make a program using the speechrecognition module in python to recognize speech and execute the following:

1. convert the speech to text
2. open a URL using webbrowser module
3. pass a query using speech recognition to make a search in the url

```
import speech_recognition as sr  
r= sr.Recognizer()  
with sr.Microphone() as source:  
    print('speak anything : ')  
    audio = r.listen(source)  
    query = r.recognize_google(audio)  
    print(query)
```

## **Use Case**

We will make a program using the speechrecognition module in python to recognize speech and execute the following:

1. convert the speech to text
2. open a URL using webbrowser module
3. pass a query using speech recognition to make a search in the url.



Following is the program for the above problem statement:

```
import pytsx3 #pip install pytsx3
import speech_recognition as sr #pip install speechRecognition
import datetime
import wikipedia #pip install wikipedia
import webbrowser
import os
import smtplib

engine = pytsx3.init('sapi5')
voices = engine.getProperty('voices')
# print(voices[1].id)
engine.setProperty('voice', voices[0].id)

def speak(audio):
    engine.say(audio)
    engine.runAndWait()

def wishMe():
    hour = int(datetime.datetime.now().hour)
    if hour>=0 and hour<12:
        speak("Good Morning!")

    elif hour>=12 and hour<18:
        speak("Good Afternoon!")

    else:
        speak("Good Evening!")

    speak(" Hlw sir i am your assistant Sir. Please tell me how may I help you")

def takeCommand():
    #It takes microphone input from the user and returns string output

    r = sr.Recognizer()
    with sr.Microphone() as source:
```

```

    print("Listening...")
    r.pause_threshold = 1
    audio = r.listen(source)

    try:
        print("Recognizing...")
        query = r.recognize_google(audio, language='en-in')
        print(f"User said: {query}\n")

    except Exception as e:
        # print(e)
        print("Say that again please...")
        return "None"
    return query

def sendEmail(to, content):
    server = smtplib.SMTP('smtp.gmail.com', 587)
    server.ehlo()
    server.starttls()
    server.login('youremail@gmail.com', 'your-password')
    server.sendmail('youremail@gmail.com', to, content)
    server.close()

if __name__ == "__main__":
    wishMe()
    while True:
        # if 1:
            query = takeCommand().lower()

            # Logic for executing tasks based on query
            if 'wikipedia' in query:
                speak('Searching Wikipedia...')
                query = query.replace("wikipedia", "")
                results = wikipedia.summary(query, sentences=6)
                speak("According to Wikipedia")
                print(results)
                speak(results)

            elif 'open youtube' in query:
                webbrowser.open("https://www.youtube.com/watch?v=803l9Wz_XFY&list=")

```

```

elif 'open google' in query:
    webbrowser.open( query + "http://www.w3.org/2000/svg/")
    speak("this is google main page sir")
elif ' site' in query:
    webbrowser.open( site + " .com/")

elif 'open map' in query:
    webbrowser.open( query +
"https://www.google.com/maps/@27.6194928,78.2076104,13.83z" )
    speak("your location is this")
elif 'open stackoverflow' in query:
    webbrowser.open("stackoverflow.com")

elif 'play music' in query:
    music_dir = 'F:\\lokesh\\Alarms\\music'
    songs = os.listdir(music_dir)
    print(songs)
    os.startfile(os.path.join(music_dir, songs[0]))
    speak("this is your favorite song sir")

elif 'the time' in query:
    strTime = datetime.datetime.now().strftime("%H:%M:%S")
    speak(f"Sir, the time is {strTime}")

elif 'open code' in query:
    codePath = "C:\\Users\\Lenovo\\AppData\\Local\\Programs\\Microsoft VS
Code\\Code.exe"
    os.startfile(codePath)

elif 'email to lokesh' in query:
    try:
        speak("What should I say?")
        content = takeCommand()
        to = "lokesh12022002@gmail.com"
        sendEmail(to, content)
        speak("Email has been sent!")
    except Exception as e:
        print(e)
        speak("Sorry my friend lokesh sengar")

```

```
elif 'exit' in query:  
    exit()
```

## Packages available for speech recognition in python

- apiai
- SpeechRecognition
- Google\_speech\_cloud
- assemblyai
- Pocketsphinx
- Watson\_developer\_cloud
- wit

We will go through the details of Speech Recognition package in this file, lets also take a look down the memory lane to understand how speech recognition systems have evolved over the years.

The very first prototype of the speech recognition was in fact a toy, named **radio rex** which came around 1920's. It had a dog sitting in a dog house which would pop out as soon as someone uttered the word rex.

The only problem with the model was that the spring was attached to an electromagnet which was sensitive to energy ranging around 500hz. Being purely a frequency detector, it could be remotely termed as a speech recognition model.

After the introduction of deep neural networks, most of the speech recognition models work on the neural networks. The possibilities are unimaginable with the neural networks, the vocabulary can go up to 10k words and more.

## **Challenges Faced By A Speech Recognition System**

A speech recognition system becomes difficult to make because we have so many sources of variability when it comes to speech.

### **Style of speaking**

Every individual person has a varied style of speaking, including accents as well. As we all know, we have different accents for speaking English too. There is American English, British English and so many other accents when it comes to speaking the most common language in the world. Pronunciation also makes it difficult for a speech recognition system to translate the speech altogether.

### **Environment**

Environment adds a lot of background noise to the system as well. An isolated room compared to an auditorium will have a lot of variability in background noises. Even echo can add a lot of noise in the system as well.

### **Speaker characteristics**

An old person's voice may not be the same as that of an infant. The characteristics of a person's speech depends on many factors including the harshness and clarity as well.

### **Language constraints**

Some spoken utterances may not have a viable meaning when it comes to translation.

# Conclusion

**In this Speech Recognition in Python you first understood what speech recognition is and how it works. You then looked at various speech recognition packages and their uses and installation steps. You then used Speech Recognition, a python package to convert speech to text using the microphone feature.**