**Final Report**

**Voice/*Speech* Recognition *System***

**Artificial Intelligence**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sr.No | Registration  Number | Name of student | Roll  Number | Marks Obtained | Signature |
| 1. | 11805111 | P Ajay Kumar Reddy | RK18PTB41 |  |  |
| 2. | 11815813 | Shykat Roy | RK18PTB42 |  |  |
| 3. | 11815075 | Supantha Sarker | RK18PTB43 |  |  |
| 4. | 11911307 | Harish Soni | RK18PTB44 |  |  |

***Submitted To***

**Nikita Kaushik, Assistant Professor**

**Lovely Professional University**

**Jalandhar, Punjab, India.**

******

***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. In this project, we will go through the speech recognition module in python.

Speech recognition is the process of converting spoken words to text. Python supports many speech recognition engines and APIs, including Google Speech Engine, Google Cloud Speech API,  
Microsoft Bing Voice Recognition and IBM Speech to Text.

In this system user can commend through their speech like natural language. According to user commend system will perform multiple activities . For example if the commend like “**open google**”, system will react on that command and open google for browsing.

***Literature Review***

• Programming Collective Intelligence

• Recommender Systems: Voice Recognition

• Building Recommendation Engines

***Proposed Methodology***

A library that helps is named “SpeechRecognition”. We install it with pyenv, pipenv or virtualenv. We can also install it system wide:

|  |
| --- |
| pip install SpeechRecognition |

The SpeechRecognition module depends on pyaudio, we have installed them from package manager.  
On Manjaro Linux these packages are called “python-pyaudio” and “python2-pyaudio”, they may have another name in different system.

**Speech Recognition demo**

Anyone can test the speech recognition module, with the command:

python -m speech\_recognition

Results will show in terminal

**Speech Recognition with Google**

Google Speech Recognition engine, which we’ve tested for the English language.

For testing purposes, it uses the default API key.  
To use another API key, use

|  |
| --- |
| `r.recognize\_google(audio, key="GOOGLE\_SPEECH\_RECOGNITION\_API\_KEY")` |

**How Speech Recognition Works? / Algorithm**

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.

**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 a 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 the 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.

***Result and Discussion***

The result of the project came as I was able to learn a lot of new things which we were not familiar earlier .It gives the best outcome of the any speech which is given by any user .We always think that how the modern and latest voice recognition works like goole assistant , siri etc. After completing this project we are able to understand that how it exactly work and which modules are used in building voice recognition system.

***Conclusion***

we would like to conclude that after this project we learnt the working of modern speech recognition system which we have experienced before while using google assistant, siri YouTube and many other platforms. We came across lots of new things , module and websites to learn about machine learning and Artificial Intelligence.

***References***

• Programming Collective Intelligence

• Recommender Systems: The Voice Recognition

• Building Recommendation Engines

• Website: [https://pythonprogramminglanguage.com/speech- recognition/](https://pythonprogramminglanguage.com/speech-%2520%2520recognition/)

• YouTube videos

• Wikipedia

***Work Distribution***

1. P Ajay Kumar Reddy : collect info
2. Harish Soni : whole project compilation and execution
3. Supantha Sarker : collect info
4. Shykat Roy : Report