

# INT404 ARTIFICIAL INTELLIGENCE

**PROJECT TITLE: SPEECH RECOGNITION**

## Name : K. Sai Vinesh

**Roll No : 62**

**Reg No : 11803034**

**Section : K18JC**

**Group : 2**

**Submitted to : Sagar Pande sir**

Table of Contents

1. ABSTRACT 2
2. INTRODUCTION 3 – 5
   1. Aim
   2. Motivation
   3. Context
3. Conceptual Framework 5 - 6
   1. Libraries used and explanation
   2. Advantages of Speech Recognition
4. Scope of the Project 6 -7
5. Screenshots 7 - 8
6. Working 8 - 10
7. References 10
8. Conclusion 10

**SPEECH RECOGNITION**

**ABSTRACT:-**

Speech recognition technology is one from the fast growing engineering technologies. It has a number of applications in different areas and provides potential benefits. Nearly 20% people of the world are suffering from various disabilities; many of them are blind or unable to use their hands effectively. The speech recognition system in those particular cases provide a significant help to them, so that they can share information with people by operating computer through voice input.

This project is designed and developed keeping that factor into mind, and a little effort is made to achieve this aim. The project is capable to recognize the speech and convert the input audio into text.

**ACKNOWLEDGEMENT:-**

I would like to thank my mentor Prof. Sagar Pande for his advice and inputs on this project. May thanks to my friends as well, who spent their valuable time to listen and provide feedback.

**INTRODUCTION:**

Aim:

To create a simple AI Speech Recognition which is used to interact with the user on basis of user inputs.

Motivation:

I chose to do this project because I’m in very much excited and interested that how basically the back end of the Speech Recognition works as we know on the basis user inputs that how front end will gives the reply to the user. So, I wanted to work upon the backend and to create a simple AI Speech Recognition. Since I’m beginner I wanted to be it simple.

Context:

Speech Recognition (SR) is the ability to translate a dictation or spoken word to text.

Speech Recognition known as “automatic speech recognition“(ASR),or speech to text(STT)

Speech recognition is the process of converting an acoustic signal, captured by a microphone or any peripherals, to a set of words.

To achieve speech understanding we can use linguistic processing

The recognized words can be an end in themselves, as for applications such as commands & control data entry and document preparation.

In the society every one either human or animals wish to interact with each other and tries to convey own message to others. The receiver for messages may get the exact and full idea of the senders, or may get the partial idea or sometimes cannot understand anything out of it. In some cases may happen when there is some lacking in communication (i.e when a child convey message, the mother can understand easily while others cannot).

**Types of speech recognition**

Speech recognition systems can be divided into the number of classes based on their ability to recognize that words and list of words they have. A few classes of speech recognition are classified as under:

**Isolated Speech**

Isolated words usually involve a pause between two utterances; it doesn’t mean that it only accepts a single word but instead it requires one utterance at a time.

**Connected Speech**

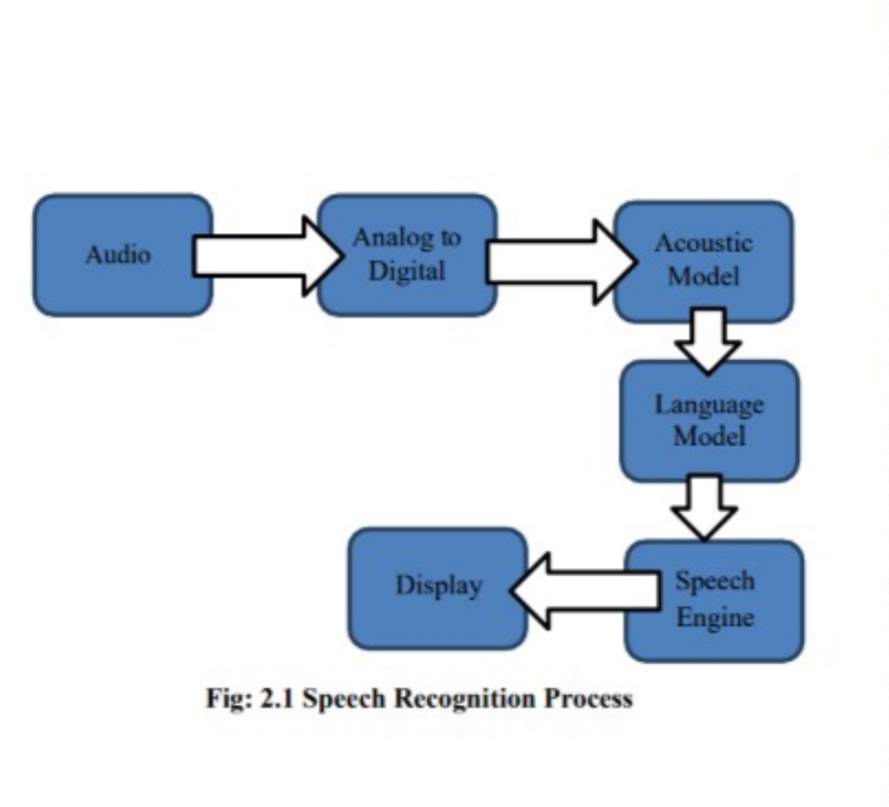
Connected words or connected speech is similar to isolated speech but allow separate utterances with minimal pause between them.

**Continuous speech**

Continuous speech allow the user to speak almost naturally, it is also called the computer dictation.

**Spontaneous Speech**

At a basic level, it can be thought of as speech that is natural sounding and not rehearsed. An ASR system with spontaneous speech ability should be able to handle a variety of natural speech features such as words being run together, “ums” and “ahs”, and even slight stutters.



**Conceptual framework:**

Library used:

1. Speech Recognition

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.

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

The Speech Recognition module depends on pyaudio, you can install them from your package manager.  
On Linux these packages are called “python-pyaudio” and “python2-pyaudio”, they may have another name in your system.

1. PyAudio

Pip install pyaudio

PyAudio provides [Python](http://www.python.org/) bindings for [Port Audio](http://www.portaudio.com/), the cross-platform audio I/O library. With PyAudio, you can easily use Python to play and record audio on a variety of platforms, such as GNU/Linux, Microsoft Windows, and Apple Mac OS X / mac OS.

**Advantages for using Speech Recognition**:

### 1. Solves Inefficiencies and Reduces Wasted Time

In a Yale Medicine study, clinicians utilized a speech recognition solution. With this software, doctors could use their voice to complete and close encounters far more quickly than before. With the widespread adoption of speech recognition systems in recent years, there have been many reports of physicians improving efficiencies and saving time.

A 2018 report by KLAS showed that overall, productivity increased due to speech recognition adoption. Clinicians were able to spend less time completing reports while using speech recognition tools allowing them to focus more on their patients.

### 2. Clinics and Hospitals Can Save Money

Many organizations find that speech recognition systems can help them reduce costs. With speech recognition, it may be possible to reduce overtime hours for transcriptionists and/or stop outsourcing clinician dictation to outside firms. It is more efficient for everyone, saving money in the process.

### 3. Clinician Satisfaction

When a clinician is able to complete their work in a shorter amount of time, they’ll feel more satisfied with their job overall. They will spend less time sitting at their desk completing documentation while spending more time with their patients. They can also focus on their life outside of work because they won’t be finishing up documentation during evenings or weekends.

According to a 2018 Health Day News report, EHRs can be made more efficient and user-friendly by adding in speech recognition features.

**Speech Recognition weakness:**

Besides all these advantages and benefits, yet a hundred percent perfect speech recognition system is unable to be developed. There are number of factors that can reduces the accuracy and performance of a speech recognition program. Speech recognition process is easy for a human but it is a difficult talk for a machine , comparing with a human mind speech recognition programs seems less intelligent, this is due to that fact that a human mind is God gifted thing and the capability of thinking, understanding and reacting is natural, while for a computer program it is a complicated task, first it need to understand the spoken words with respect to their meanings, and it has to create a sufficient balance between the words, noise and spaces. A human has a built in capability of filtering the noise form a 10 speech while a machine requires training, computer requires help for separating the speech sound from the other sounds.

**Scope of the Project**

**This project has the speech recognizing though it is not a complete replacement of what we call a notepad but still a good text editor to be used through voice.**

Milestones:

13/02/2020 -- Learning about Speech Recognition.

15/02/2020 – Applications, Advantage and dis-advantges of Speech Recognition.

20/02/2020 – Research about the libraries required to implement for Speech Recognition.

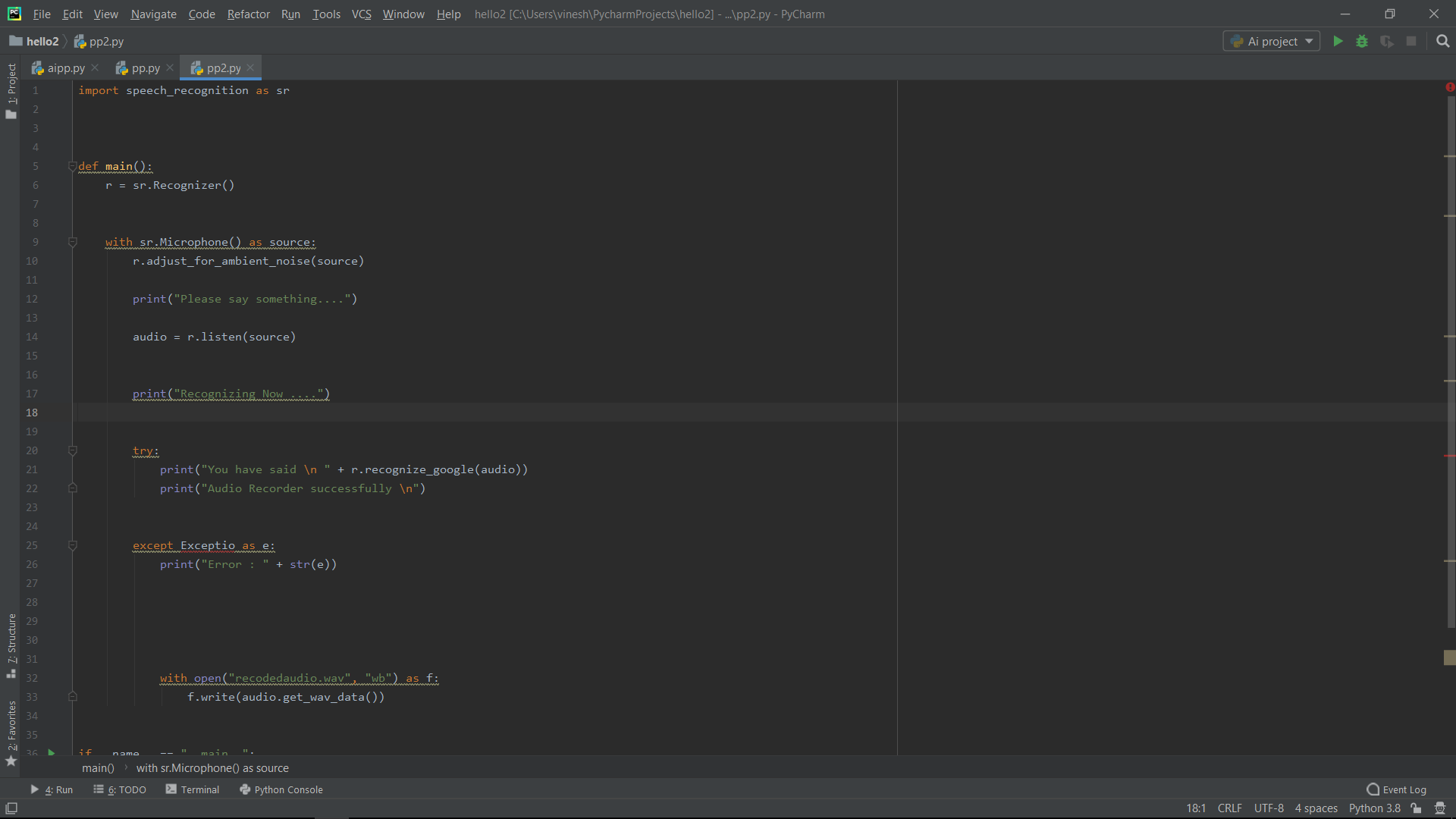
05/03/2020 – Setting the code.

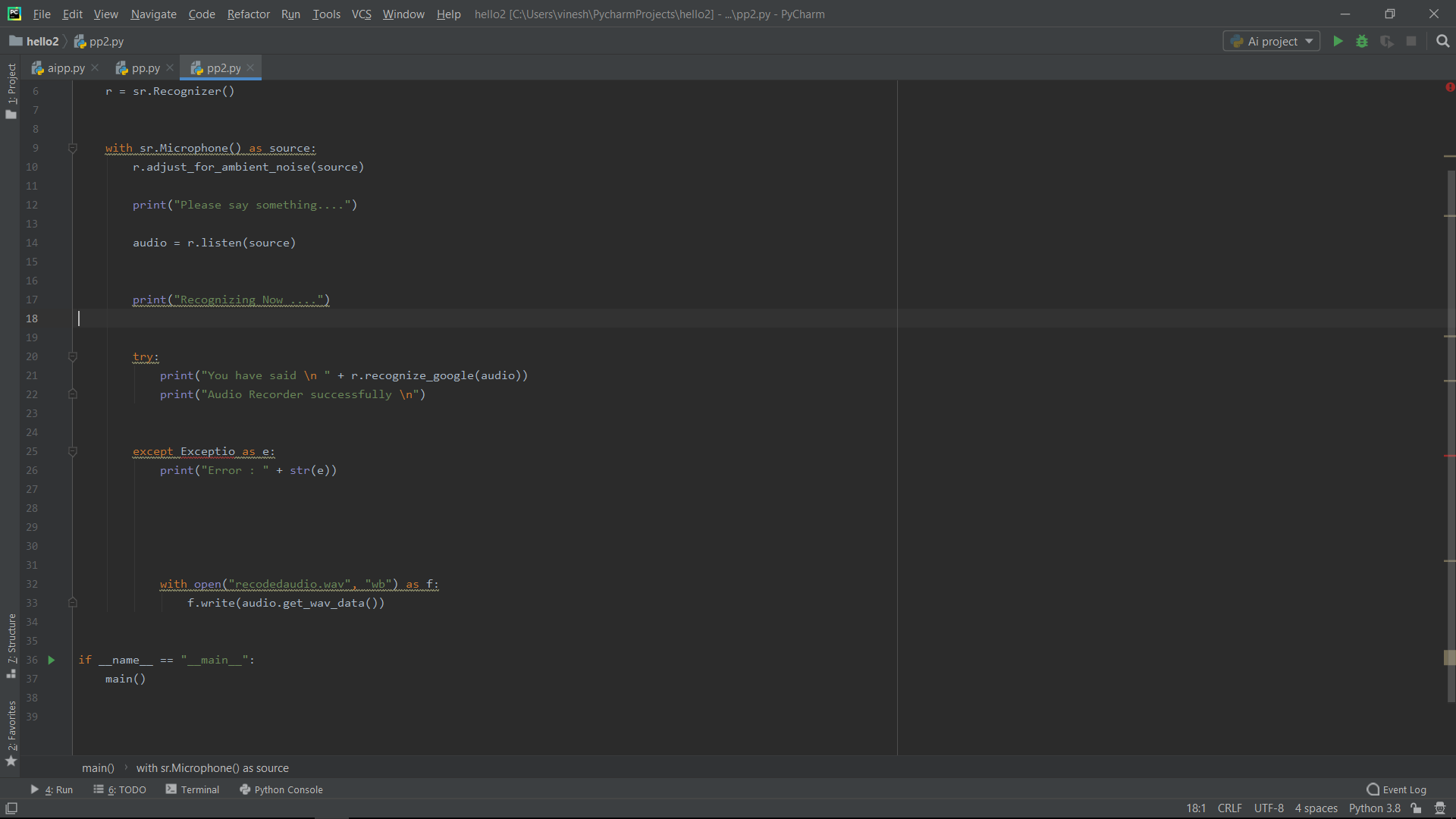
11/03/2020 – Testing the code.

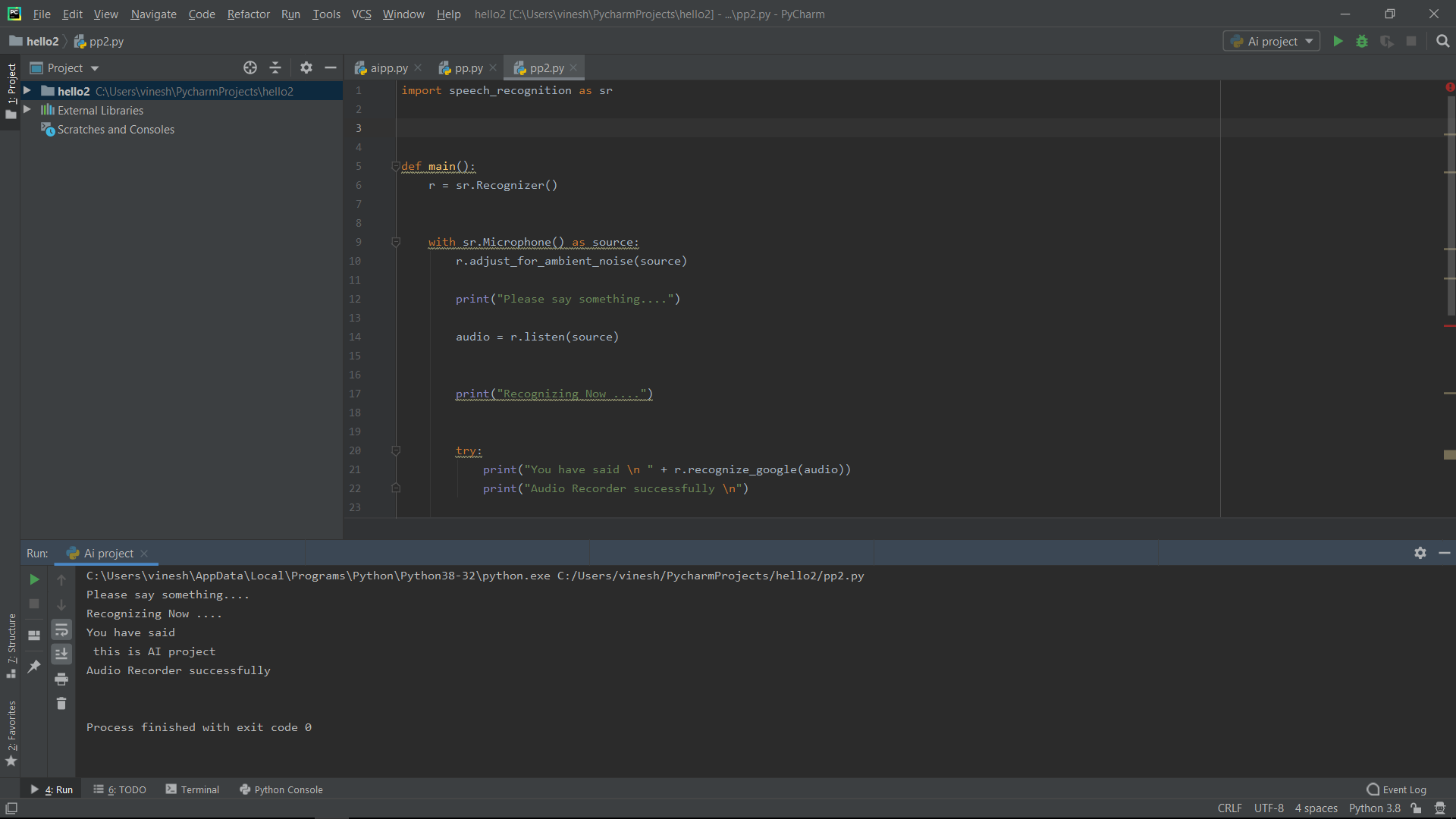
19/03/2020 – Rechecking and adding details to the code.

21/03/2020 – Code is completed(code is fully functional).

**Screenshots:-**







**Working:**

Speech recognition software works by breaking down the audio of a speech recording into individual sounds, analyzing each sound, using algorithms to find the most probable word fit in that language, and transcribing those sounds into text.

Speech recognition software uses natural language processing (NLP) and deep learning neural networks. “NLP is a way for computers to analyze, understand, and derive meaning from human language in a smart and useful way,” according to the Algorithm blog. This means that the software breaks the speech down into bits it can interpret, converts it into a digital format, and analyzes the pieces of content.

From there, the software makes determinations based on programming and speech patterns, making hypotheses about what the user is actually saying. After determining what the users most likely said, the software transcribes the conversation into text.

This all sounds simple enough, but the advances in technology mean these multiple, intricate processes are happening at lightning speed. Machines can actually transcribe human speech more accurately, correctly, and quickly than humans can.

**References:**

* I have researched about Speech Recognition from Google.
* I researched and learnt the libraries required for it from You Tube.

**Conclusion:**

This project work of speech recognition started with a brief introduction of the application and the technology in the computer (desktop applications)This project able to write the text through both keyboard and voice ,the speech recognition .

One challenge is to develop ways in which our knowledge of the speech signal, and of speech production and perception, can be incorporated more effectively into recognition methods. For example, the fact that speakers have different vocal tract lengths could be used to develop more compact models for improved speaker-independent recognition.