



## Converting Audio Signal into Text in Python Using Speech Recognition API

<sup>1</sup>Amit Kumar Rai, <sup>2</sup>Samapti Rudra, <sup>3</sup>Sourav Karmakar, <sup>4</sup>Sourav Bhattacharyya, <sup>5</sup>Mousumi Mitra, <sup>6</sup>Mousumi Roy, <sup>7</sup>Tomaly Nayak

<sup>1</sup>Assistant Professor, Electronics and Communication Engineering, Asansol Engineering College, (MAKAUT) Asansol, India

<sup>2,3,4,5,6,7</sup>Electronics and Communication Engineering, Asansol Engineering College, (MAKAUT) Asansol, India

<sup>1</sup>amitrai2212@gmail.com, <sup>2</sup>samaptirudra2001@gmail.com, <sup>3</sup>souravkarmakar927@gmail.com,

<sup>4</sup>bhattacharyyasourav35@gmail.com, <sup>5</sup>mousumimit12@gmail.com, <sup>6</sup>mousumiroy429@gmail.com,

<sup>7</sup>tomalynayak303@gmail.com

### ABSTRACT

Flexible As a personal voice assistant, this project is built with Python, Qt designer and speech recognition API. Autonomous devices like voice assistant are becoming smarter to interact with human as well as computer. The main task is to reduce the use of input devices. We chose Python over other programming languages as it is understandable and easy to write the script while building our voice assistant. We have also deployed an attractive GUI with the help of Qt designer so that the user can have a favorable experience when working with it. With the help of speech recognition API which allows us to convert speech into text has helped us to make our assistant more interactive. Our voice assistant can have the capability to work with and without internet. Without internet it can perform the task such as opening of notepad, CMD, taking screenshot, can do some basic mathematical calculations, restarting the system and turning off the system too. With the help of internet its functions expand. Almost with 100% accuracy it can fetch weather report of user's location, can work simultaneously in Google, YouTube. It is also capable of opening social media platforms like Instagram, Facebook. For entertainment and time-saving purpose it can also be used to play music and read PDF respectively. To sum up, the above task can function excellently with just voice commands. The upgrades are immeasurable. In this paper AI can be used to minimize human efforts with same efficiency

**Key words:** Speech bRecognition, Python, Speech to Text

### INTRODUCTION

Our world is becoming digitalized day by day with the innovation of new technologies. With the development of these technologies our personal as well as professional life has become very easy and handy. Intelligent Personal Assistant is one of the most vital innovations in terms of easing our lives. Now we decided to build a PC personal Assistant that works on voice commands and executes our query. Our project is mainly built using Python which is one of the most properly used, user-friendly, interpreted language used by the professionals. This software uses device's microphone to accept the voice and the output takes place at the systems speaker. Our project is a mixture of various technologies like voice recognition, voice analysis and language processing. When the user will ask the voice assistant to perform the task, the natural audio signal is converted into digital data which is then analyzed by the software. Keywords are reserved

words, they perform some particular tasks, if those keywords are present then that particular task is performed. for example, 'translate' keyword is used for the translation of the word from one language to another. So mainly keywords perform particular task if that is present in your audio actual task is going to be performed. Our voice assistant can perform task like opening and closing of, notepad, command prompt, searching information from Wikipedia, fetching top news and reading the news, telling the present weather of particular city, translating word to sentence, taking screenshot etc. In python there is an API called Speech Recognition which converts speech to text. This is a built-in library function in python. In our project we have used this library Speech Recognition for conversion of speech to text. We have also used libraries like Wikipedia to fetch particular information from wikipedia, pyttsx3 library is used to perform text to speech and many others. So mainly our project will convert audio signal to text signal using Speech Recognition API. This is the workflow of our project.

### **EXISTING SYSTEM**

We all aware of the voice assistants like Google Siri, Cortana etc. which process our language and voice recognition. The working principle of such voice assistant is that they listen the commands from the user and perform that specific function. They provide accurate and efficient result because they use AI. This assistant is basically developed to reduce human efforts. They are very time consuming as they almost removed the concept of typing. They are no less than human assistant. Apart from humans, they work way more fast. Voice assistant from Google and Microsoft basically need an account and internet connectivity to run.

### **PROPOSED SYSTEM**

Inspired from Google our team decided to deploy our own voice assistant which made our day-to-day daily activities on desktop very easier. It became easier to search on Google without opening the browser and perform any other daily task like playing music opening your favorite IDE with the help of a single voice command. Our voice assistant does not need any account to run. It can perform with and without internet as well. We did this project with the help of python and written our code in VS code. All the necessary packages and libraries were easily installable in this IDE. Our voice assistant can perform any task with same effectiveness. With the help of this project, we have realized to the basic concept of AI which results in decreasing human efforts.

### **BASIC WORK DESCRIPTION**

The project aims to develop a personal-assistant for desktop. Our voice assistant draws its inspiration from virtual assistants like google assistant for android, and Siri for iOS. It has been designed to provide a user-friendly interface for carrying out a variety of tasks by employing certain well-defined commands.

Users can interact with the assistant either through voice commands or using keyboard input. We have made some components which we have broke them into functions. We have made a function called task Execution () which will execute most of the functions of our voice assistant. This function will open notepad, cmd, play music, etc. There is a takeCommand () function which will ask user to give their commands. As a personal assistant, our voice assistant assists the end-user with day-to-day activities like general human conversation, searching queries in google, searching for videos, sending messages, taking screenshot, live weather conditions, etc.

### **SOFTWARE USED**

#### **Python: -**

- Python is a general purpose, high level, interpreted language with easy syntax
- created by Guido Van Rossum in 1989
- Python is an open - source language
- python can be used to make desktop, web and mobile apps.

#### **Qt: -**

- Qt is a cross-platform application and UI framework.
- Using Qt, you can write applications once and deploy them across desktop, mobile and embedded operating systems without rewriting the source code.

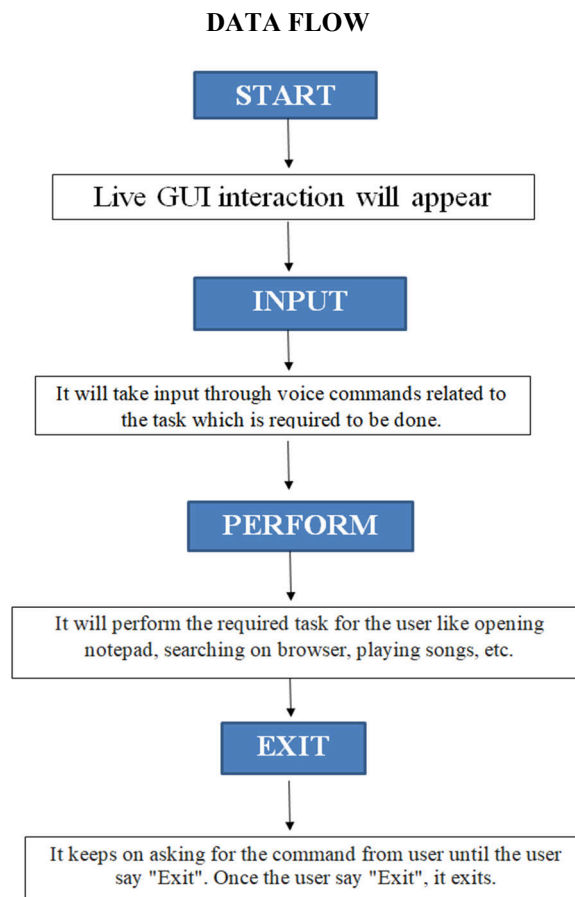
#### **Speech recognition: -**

- Breaks down audio into individual sounds
- Converts these sounds into a digital format

- Uses algorithms and models to find the most probable word fit in that language.
- Example - Google Translate, Siri, Alexa

**VS code: -**

- Visual Studio Code
- VS is one of the most powerful IDE that works with all major programming.
- It is maintained by Microsoft Corporation and this IDE is loaded with rich features and functionalities that one can possibly imagine.



**Fig. 1** Data Flow

#### **LIBRARIES USED**

pyttsx3: Conversion of speeches into texts.

Speech Recognition: Python module which converts human speech into text in our voice assistant.

Datetime: This library provides us the current date and time.

Wikipedia: Module used to fetch results through Wikipedia.

pyPDF2: Module used to read PDFs.

Pyjokes: It is a python libraries which stores humorous jokes.

Webbrowser: Creates an interface to display documents in web browser.

Pyautogui: Used for graphics purposes.

os: It represents Operating System related functionality.

sys: Providing access to all the variables and functions within the code results in strong interaction with interpreter.

TEST RESULTS

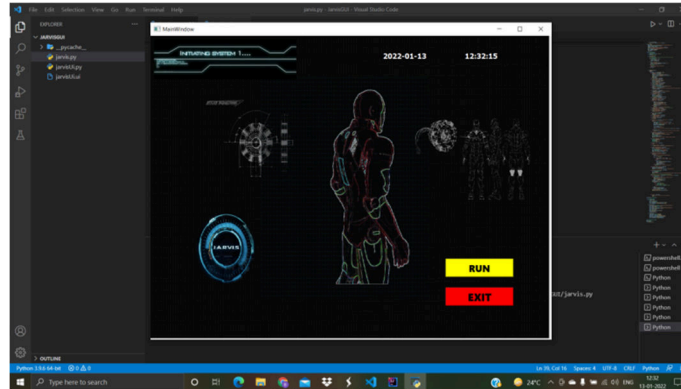


Fig. 2 After executing the program GUI window appears

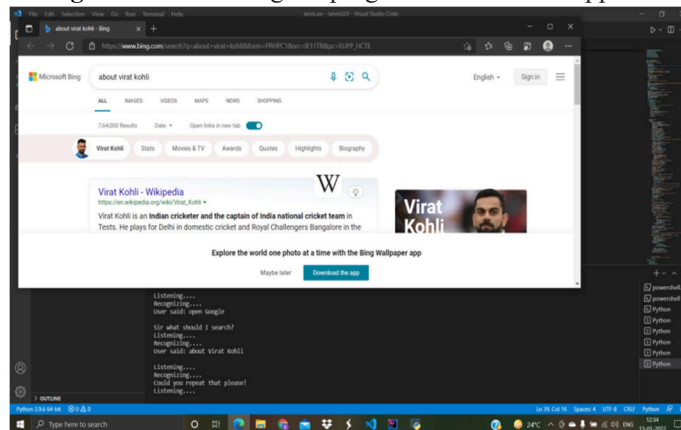


Fig. 3 According to user command "Open google and search about Virat Kohli"

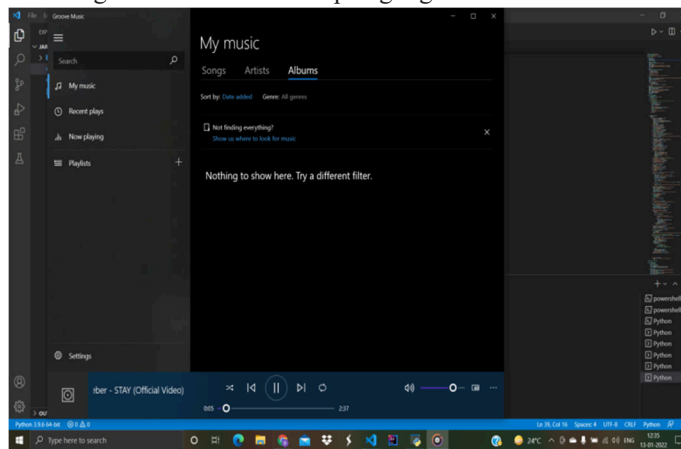


Fig. 4 According to user command "play music"

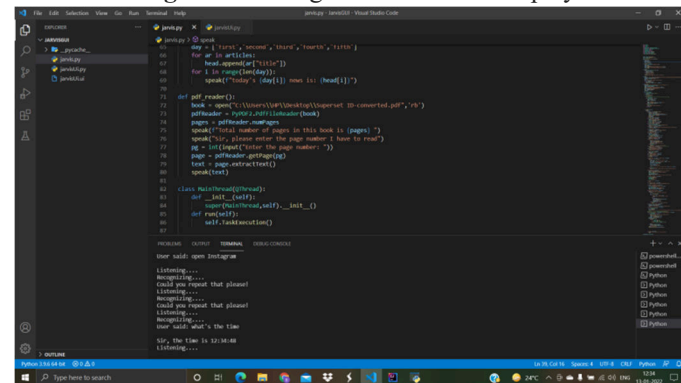


Fig. 5 According to user command "What's the time?"

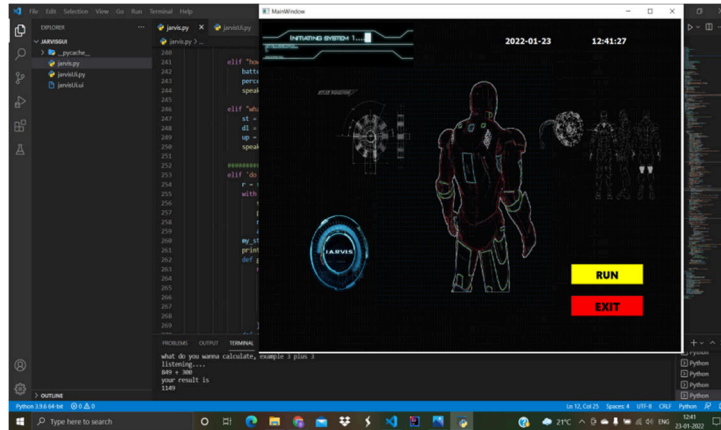


Fig. 6 According to user command "Do some calculation"

### FUTURE SCOPE

- Voice commands can be encrypted to maintain security.
- Make JARVIS to learn more on its own and develop a new skill in it.
- JARVIS android app can also be developed.
- Make more Jarvis voice terminals.
- Make our voice assistant connect with multiple devices.
- Streaming videos on smart TV through this voice assistant.

### CONCLUSION

In this paper, we discussed the design and implementation of a Voice Assistance. Though it is a modular in nature so we can easily change or add any functions in this project without affecting the current system. This project will help for visually impaired and physically challenge people. Further, in the long run, our Voice Assistant is planned to feature auto deployment supporting some daily functions in our system, and all operations which a general server administrator does.

### REFERENCES

- [1]. Subhas mani Kaushal, megha Mishra (IRJET) volume:07 Issue:09|sep 2020
- [2]. B. Marr, The Amazing Ways Google Uses Deep Learning AI.
- [3]. M.A.M.S.M.E.P.D.J.M.G.S.G.S. (2021, March 3). Creating Desktop Speech Recognition Using Python Programming. Ijarcce.Com. Retrieved January 24, 2022, from <https://ijarcce.com/wp-content/uploads/2021/04/IJARCCCE.2021.10325.pdf>
- [4]. Rishabh Shah, Siddhant Lahoti, Prof. Lavanya. K, "An Intelligent Chatbot using Natural Language Processing". International Journal of Engineering Research, Vol. 6, pp. 281-286, 1 May 2017
- [5]. Deepak Shende, RiaUmahiya, Monika Raghorte, AishwaryaBhisikar, AnupBhange, "AI Based Voice Assistant Using Python", Journal of Emerging Technologies and Innovative Research (JETIR), February 2019, Volume 6, Issue 2.
- [6]. <https://doc.qt.io/qt-5/qt designer-manual.html#:~:text=Qt%20Designer%20is%20the%20Qt,using%20different%20styles%20and%20resolutions.>
- [7]. <https://blogs.mulesoft.com/learn-apis/api-led-connectivity/what-are-apis-how-do-apis-work/>