

A
Project Report
On
AI Voice Assistant

SUBMITTED BY

Name : Kokane Deepak Balaji

Roll No. : N20111020

Degree Project II

M.Sc.(Computer Science) – SEM IV

Department of Computer Science,

Savitribai Phule Pune University,

Pune-411007



Index

Sr. No.	TITLE	PAGE NO.
1.	INTRODUCTION	3
2.	PROJECT AIM/ABSTRACT	4
3.	PROBLEM STATEMENT AND SOLUTION	5 - 6
4.	OBJECTIVE	7
5.	WORKING	8 – 9
6.	METHODOLOGY OF VIRTUAL ASSISTANT USING PYTHON	10 - 12
7.	ADVANTAGES OF VOICE ASSISTANT	13
8.	CONCLUSION	14
9.	REFERENCES	15
10.	GitHub Link :	15

INTRODUCTION

Today the development of artificial intelligence (AI) systems that can organize a natural human-machine interaction (through voice, communication, gestures, facial expressions, etc.) are gaining in popularity. One of the most studied and popular was the direction of interaction, based on the understanding of the machine by the machine of the natural human language. It is no longer a human who learns to communicate with a machine, but a machine learns to communicate with a human, exploring his actions, habits, behaviour and trying to become his personalized assistant.

Virtual assistants are software programs that help you ease your day to day tasks, such as showing weather reports, creating reminders, making shopping lists etc. They can take commands via text (online chatbots) or by voice. Voice-based intelligent assistants need an invoking word or wake word to activate the listener, followed by the command. We have so many virtual assistants, such as Apple's Siri, Amazon's Alexa, and Microsoft's Cortana.

This system is designed to be used efficiently on desktops. Personal assistants' software improves user productivity by managing routine tasks of the user and by providing information from an online source to the user.

This project was started on the premise that there is a sufficient amount of openly available data and information on the web that can be utilized to build a virtual assistant (Ai Voice Assistant) that has access to making intelligent decisions for routine user activities. we are going to build a voice-based application which is capable of doing all the above-mentioned tasks.

PROJECT AIM/ABSTRACT

In this modern era, day to day life became smarter and interlinked with technology. We already know some voice assistance like google, Siri. etc. Now in our voice assistance system, it can act as a basic medical prescriber, daily schedule reminder, note writer, perform operation on applications, calculator, and a search tool. This project works on voice input and give output through voice and displays the text on the screen. The main agenda of our voice assistance makes people smart and give instant and computed results. The voice assistance takes the voice input through our microphone (Bluetooth and wired microphone) and it converts our voice into computer understandable language gives the required solutions and answers which are asked by the user. This assistance connects with the world wide web to provide results that the user has questioned. Natural Language Processing algorithm helps computer machines to engage in communication using natural human language in many forms.

PROBLEM STATEMENT AND SOLUTION

Problem Statement :

The skills problem is something that you're not faced with on desktop computers, laptops and smartphones. That's because those devices have a display and a graphical user interface (GUI) which clearly defines the capabilities and boundaries of each application. When you fire up a Windows or Mac computer, you can quickly see the list of applications that has been installed on them and get a general sense of the tasks you can perform with them. But in that we can see multiple application and we need to locate for particular application for perform same task and same time it will hard to locate. That task can be perform over voice using Voice assistant.

In this modern era, day to day life became smarter and interlinked with technology. We already know some voice assistance like google, Siri. etc. Now in our voice assistance system, it can act as a basic tasks, and a search tool.

Solution :

This project works on voice input and give output through voice and displays the text on the screen. The main agenda of our voice assistance makes people smart and give instant and computed results. The voice assistance takes the voice input through our microphone (Bluetooth and wired microphone) and it converts our voice into computer understandable language gives the required solutions and answers which are asked by the user. This assistance connects with the world wide web to provide results that the user has questioned. Natural Language Processing algorithm helps computer machines to engage in communication using natural human language in many forms.

Technology: Python 3.10

Design Details :

- **Always show user query**

Users should never ask questions like "Did the system understand what I said?" when they interact with AI assistants. They should be able to see that the system understands their request correctly.

- **Designing the conversation**

Interaction with AI assistants should not feel awkward or break patterns that have evolved over the years in human-to-human conversation. You need to capture all possible user intentions (what the user will be trying to do with the help of AI) and frame them into the model of your AI assistant.

- **Create scripts for the dialog between the user and the assistant:**

Practice it. Focus both on the interactions where everything works well, and the user gets what they want and when the situation goes not as planned. By practicing interaction with AI assistant, you will identify situations you may not have already accounted for.

- **Voice-first interface**

It's expected that AI assistant is the voice-first interface. The true power of AI is that it allows users to avoid direct manipulation with UI and solve the tasks much faster.

Dependencies and requirements :

System requirements: Python 3.9, Visual Studio Code

Install all these python libraries:

pip install SpeechRecognition

pip install wikipedia

pip install pyautogui

pip install webbrowser

pip install os

pip install winshell

pip install pyttsx3

pip install datetime

OBJECTIVE (Future of Voice Assistant (Related Work))

Each company developer of the intelligent assistant applies his own specific methods and approaches for development, which in turn affects the final product. One assistant can synthesize speech more qualitatively, another can more accurately and without additional explanations and corrections perform tasks, others can perform a narrower range of tasks, but most accurately and as the user wants. Obviously, there is no universal assistant who would perform all tasks equally well. The set of characteristics that an assistant has depends entirely on which area the developer has paid more attention to. Since all systems are based on machine learning methods and use for their creation huge amounts of data collected from various sources and then trained on them, an important role is played by the source of this data, be it search systems, various information sources or social networks. The amount of information from different sources determines the nature of the assistant, which can result as a result. Despite the different approaches to learning, different algorithms and techniques, the principle of building such systems remain approximately the same. the technologies that are used to create intelligent systems of interaction with a human by his natural language. The main technologies are voice activation, automatic speech recognition, Teach-To-Speech, voice biometrics, dialogue manager, natural language understanding and named entity recognition. following figure shows that future of Voice Assistants Technologies :-

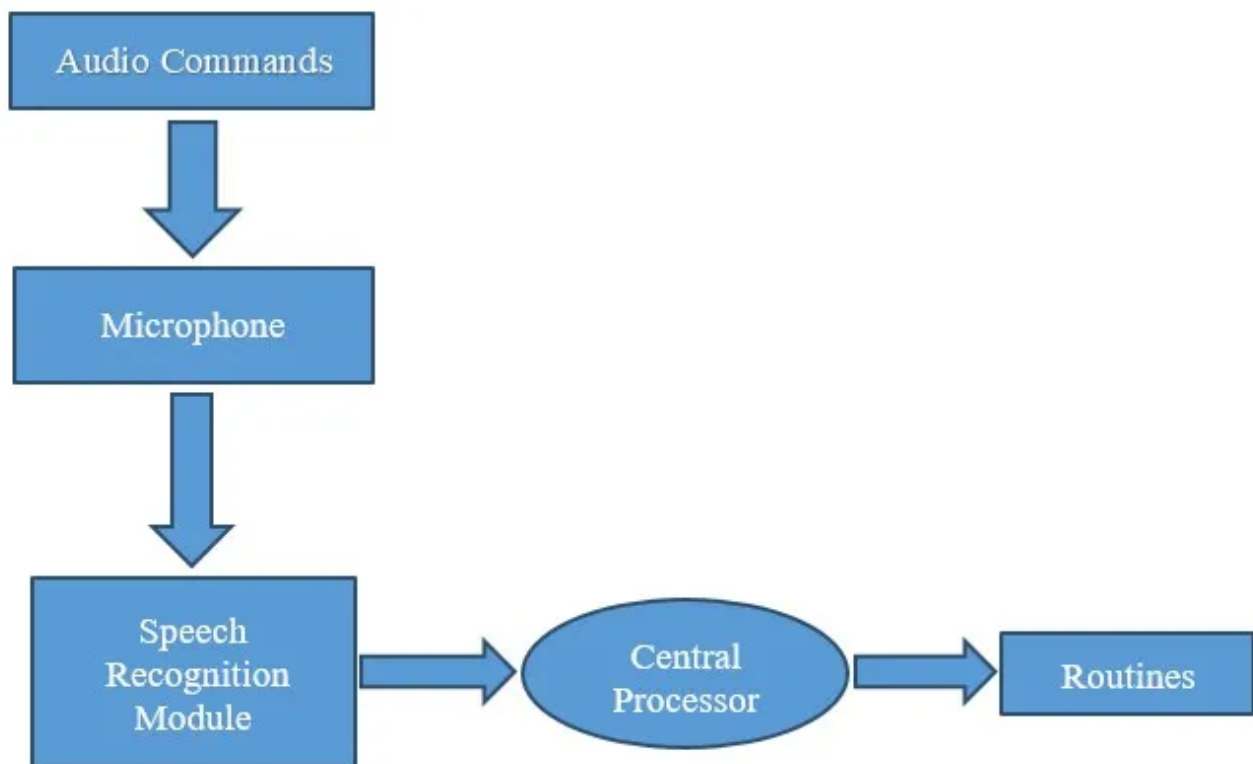
Voice Technology	Brain Technology
Voice Activation	Voice Bio-metrics
Automatic Speech Recognition (ASR)	Dialog Management
(Teach-To-Speech (TTS)	Natural Language Understanding (NLU)
	Named Entity Recognition NER)

Figure(1): Technologies for constructing intelligent systems of interaction with a human by natural language

WORKING

Proposed Plan Of Work

The work started with analysing the audio commands given by the user through the microphone. This can be anything like getting any information, operating a computer's internal files, etc. This is an empirical qualitative study, based on reading above mentioned literature and testing their examples. Tests are made by programming according to books and online resources, with the explicit goal to find best practices and a more advanced understanding of Voice Assistant.

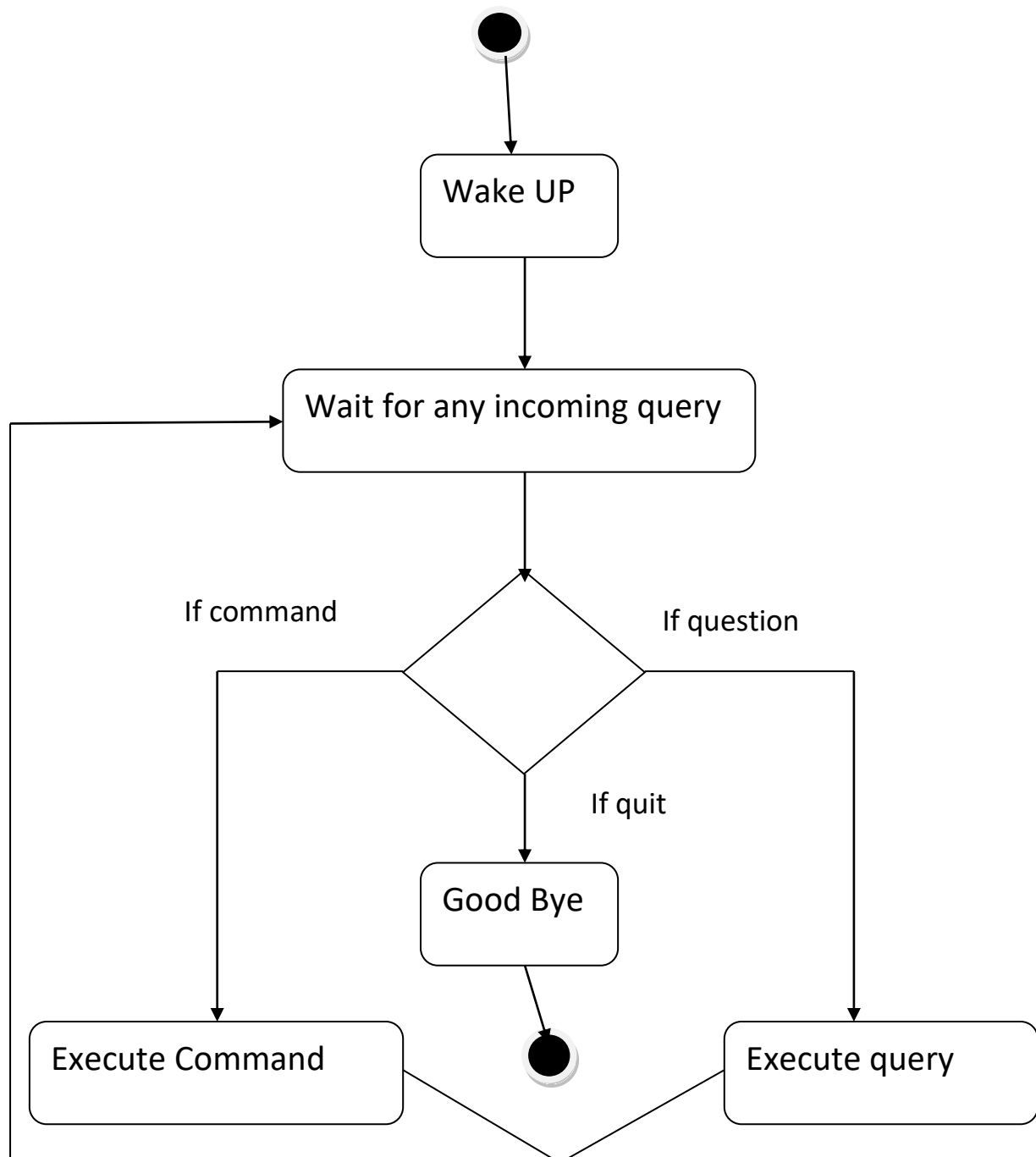


Figure(2): Basic Workflow of Ai Voice Assistant

Figure(2) shows the workflow of the basic process of the voice assistant. Speech recognition is used to convert the speech input to text. This text is then fed to the central processor which determines the nature of the command and calls the relevant script for execution. But, the complexities don't stop there. Even with hundreds of hours of input, other factors can play a huge role in whether or not the software can understand you. Background noise can easily throw a speech

recognition device off track. This is because it does not inherently have the ability to distinguish the ambient sounds it “hears” of a dog barking or a helicopter flying overhead, from your voice. Engineers have to program that ability into the device; they conduct data collection of these ambient sounds and “tell” the device to filter them out. Another factor is the way humans naturally shift the pitch of their voice to accommodate for noisy environments speech recognition systems can be sensitive to these pitch changes.

ACTIVITY DIAGRAM



Methodology of Virtual Assistant Using Python

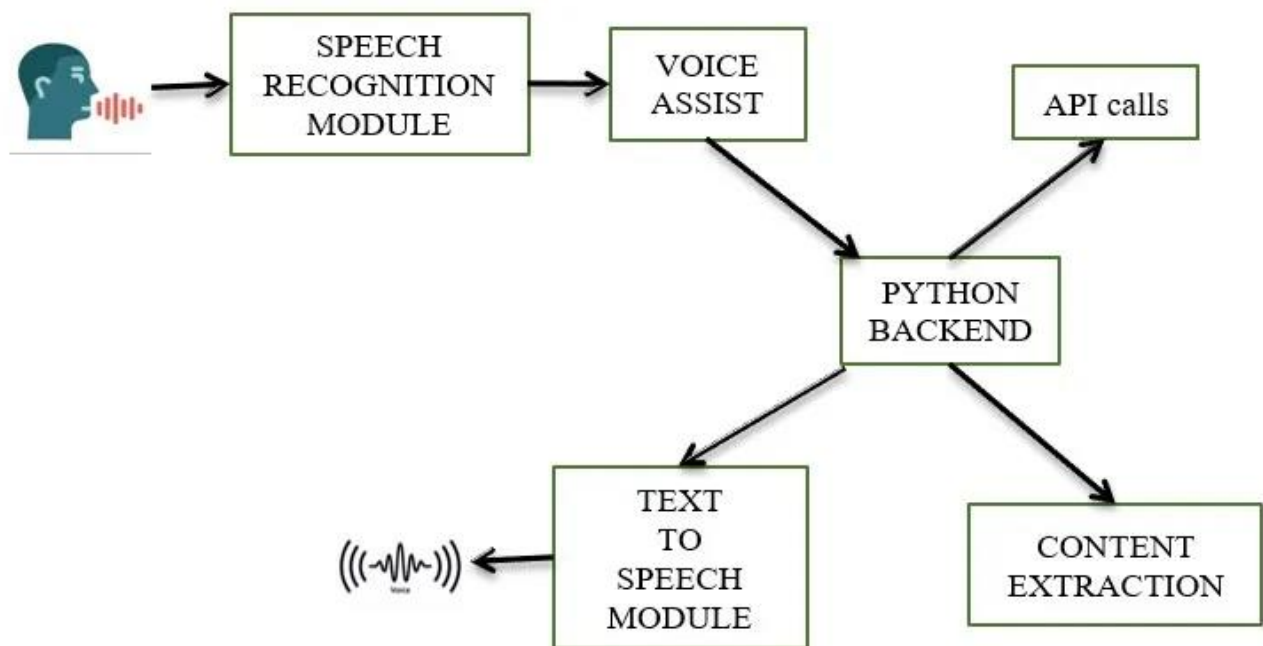


Fig 3 Detailed workflow

Speech Recognition module

The system uses Google's online speech recognition system for converting speech input to text. The speech input Users can obtain texts from the special corpora organized on the computer network server at the information centre from the microphone is temporarily stored in the system which is then sent to Google cloud for speech recognition. The equivalent text is then received and fed to the central processor.

Python Backend:

The python backend gets the output from the speech recognition module and then identifies whether the command or the speech output is an API Call and Context Extraction. The output is then sent back to the python backend to give the required output to the user.

API calls

API stands for Application Programming Interface. An API is a software intermediary that allows two applications to talk to each other. In other words, an API is a messenger that delivers your request to the provider that you're requesting it from and then delivers the response back to you.

Text-to-speech module

Text-to-Speech (TTS) refers to the ability of computers to read text aloud. A TTS Engine converts written text to a phonemic representation, then converts the phonemic representation to waveforms that can be output as sound. TTS engines with different languages, dialects and specialized vocabularies are available through third-party publishers.

Speech Recognition module

Speech recognition is a machine's ability to listen to spoken words and identify them. You can then use speech recognition in Python to convert the spoken words into text, make a query or give a reply. You can even program some devices to respond to these spoken words.

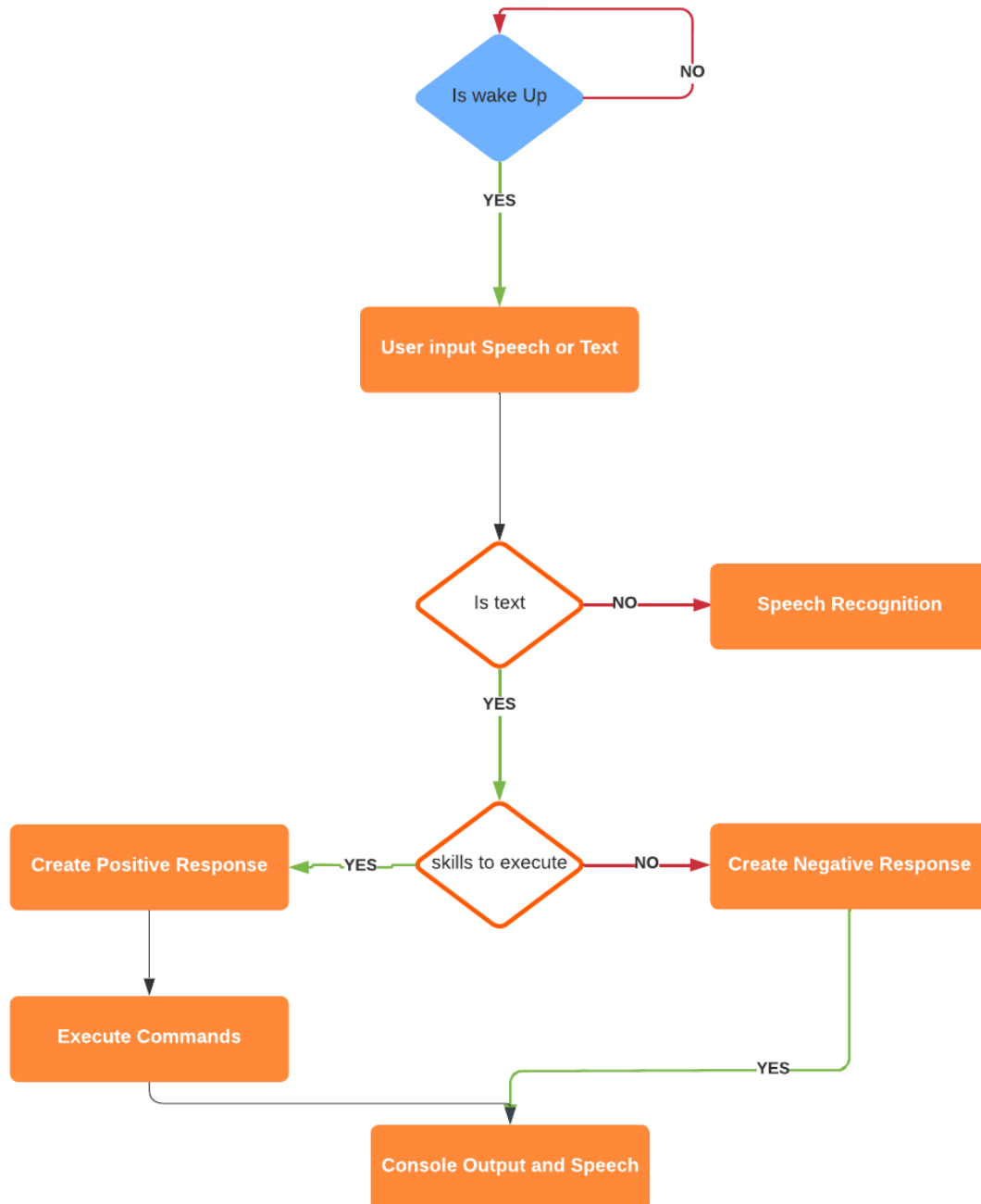
Webbrowser module

The webbrowser module provides a high-level interface to allow displaying web-based documents to users.

Winshell module

The winshell module is a light wrapper around the Windows shell functionality. It includes convenience functions for accessing special folders, for using the shell's file copy, rename & delete functionality, and a certain amount of support for structured storage.

Decision model



ADVANTAGES OF VOICE ASSISTANT

Saves Time by Automating Repetitive Tasks

Automating repeated tasks to a voice-activated personal assistant frees up the human time and resources. Also, it can efficiently perform these mundane tasks with no errors, which often leads to an improvement in customer satisfaction. While voice assistants are left to deal with routine tasks, humans can dedicate more time to duties where human intervention is required for successful business solutions and services.

Aids Hand-free Operation

Voice talking gives consumers hands-free access to many functions because you only need the voice to activate them. So it makes it easier and faster to do certain things.

Features :

Assistant Skills :

- Tells the internet availability (e.g 'is the internet connection ok?')
- Tells the daily news (e.g 'tell me today news')
- Opens windows applications (e.g 'open chrome/firefox')
- Tells everything it can do
- Tells the current Time
- Assistant Features
- Asynchronous command execution & speech recognition and interpretation
- Answers in general questions (via call Wolfram API)
- Easy voice-command customization
- Configurable assistant name ,etc

:

CONCLUSION

In this project “Virtual Assistant Using Python” we discussed the design and implementation of Ai Voice Assistant. The project is built using open source software modules with PyCharm community backing which can accommodate any updates shortly. The modular nature of this project makes it more flexible and easy to add additional features without disturbing current system functionalities.

It not only works on human commands but also give responses to the user based on the query being asked or the words spoken by the user such as opening tasks and operations. It is greeting the user the way the user feels more comfortable and feels free to interact with the voice assistant. The application should also eliminate any kind of unnecessary manual work required in the user life of performing every task. The entire system works on the verbal input.

REFERENCES

[Books] :

- [1] D. R. S. Caon, T. Simonnet, P. Sendorek, J. Boudy, and G. Chollet, “vAssist: *The Virtual Interactive Assistant for Daily Homer-Care*,” in Proceedings of pHealth, 2011.

[Links] :

Project Link : <https://github.com/dkokane0/Ai-Voice-Assistant>

Project Report Link :

<https://drive.google.com/file/d/11hzjSoEa6AENOR1zCH74cusIWzMucFr7/view?usp=sharing>