International Journal of Innovations in Engineering and Science, www.ijies.net

An Overview of Voice Enabled Personal Assistant Using Python

Rohit K. Roy[1], Megha D. Deshmukh[2], Rajat K. Biswas[3], Milind V. Lande[4]

[1,2,3]Students , [4] I/C HOD**,** *Government Polytechnic, Gadchiroli, Maharashtra, India*

Email - r*ohitroy3072@gmail.com ,deshmukhmegha776@gmail.com, rajatbiswas360@gmail.com,*
milind.jdiet@gmail.com

Abstract – *People now connect with computers in novel ways thanks to personal assistants, conversational interfaces, and chat bots. A personal virtual assistant may even perform certain basic duties like launching apps, reading out news, taking notes, etc. with just a voice command. Users can ask enquiries to them in the same way they would to a real person. Personal assistants like Siri, Google Assistant, and Alexa operate on text-to-speech technology.*

Python is being used to create a voice assistant that will enable users to complete any activity without needing a keyboard. The purpose of this Project is to examine the intelligent behaviour of voice assistants and how they might be applied to both academic and daily tasks.[18]
.

I- INTRODUCTION

Personal assistants are increasingly an essential part of our life, whether we realize it or not. It is as a result of all the features and simplicity of usage they offer. Personal assistants can also automate some routine duties so that a user can concentrate on what is most important to them.[20]

Personal assistants provide functions including making calls, sending messages, taking pictures, storing to-do lists on the go, accessing the internet, etc. Therefore, making use of these virtual assistant capabilities will enable one to save a lot of time and work. It's critical to concentrate more on what matters most to a person, whether it be personal or professional job. People frequently take longer to complete mundane chores. so that a user can concentrate on what is most important to them.[21]

Personal assistants provide functions including making calls, sending messages, taking pictures, storing to-do lists on the go, accessing the internet, etc. Therefore, using these qualities of a virtual assistant will enable someone to save a great deal of time and work. It's critical to concentrate more on what matters most to a person, whether it be personal or professional job. [18]

These kinds of personal assistants can automate mundane tasks that people frequently spend more time on. When working in an unfamiliar setting, people frequently struggle to discover the software they require, such as a browser, an IDE, or any other application. The majority of the time, people will waste hours looking for the application. [16]

This wastes time that could be used elsewhere. A voice-activated personal assistant will therefore aid in automating this process. The user is only required to issue a verbal command; the assistant will do the rest. The study discusses the use of a voice-activated personal assistant, which can help people get things done and save a lot of time by allowing them to issue speech instructions. Future-oriented technologies are reshaping digital experiences and altering how people

International Journal of Innovations in Engineering and Science,   www.ijies.net

interact with the world, including virtual reality, augmented reality, voice interaction, and IOT. One significant leap in human-machine interaction made possible by the development of artificial intelligence is voice control. Today, every major corporation uses voice assistant technology so that customers can speak to a machine for assistance. Therefore, with the Voice Assistant, we are advancing to the stage where we may speak to our machine.[14]

II- PROBLEM DEFINING

The amount of experience required to understand and use today's tools, software, and hardware is difficult for the older generation to stay up with. Gadgets and technology are being developed on a daily basis. Not only the older generations struggle to keep up with the rate of technological advancement, but also millennials, who believe it is preferable to have comparable systems for doing daily tasks. There are APIs (Application Programming Interfaces) available that can recognize and convert human speech into text and vice versa due to the growth of natural language processing. Due of this, even well-known entertainment industries have depicted a future when there will be plenty of voice-controlled robots that can help humans with almost anything on Earth. The ultimate luxury for the majority of us would be a personal assistant who is always on the lookout for your calls, anticipates your every need, and acts when necessary. Artificial intelligence assistants, often known as voice assistants, have made that luxury possible. When given a wake word or order, they can carry out a number of tasks while being quite little. They are able to turn on lights, respond to inquiries, play music, make online purchases, etc. Additionally, they have discovered that IoT (Internet of Things) connected gadgets are becoming more compatible. Voice assistants should not be confused with virtual assistants, who are independent contractors capable of doing a wide range of duties. Instead, voice assistants rely on technology. Voice assistants' usefulness in both the personal and professional spheres will increase as they become more capable.

III -PROJECT PURPOSE

With a personalized, cross-platform touch, this initiative seeks to offer voice-controlled automation services. The application uses speech recognition to control the machine and carry out simple tasks. This project presents a solution that streamlines the functionality of

modern assistants in light of the world's rising desire for personal assistants. In comparison to text- and image-based automation, this technology's competence has been demonstrated by its consistent expansion in a variety of industries. This project is a stand-alone programme that anyone who is unfamiliar with system operations can use. It can do things like launch a different application (like settings or the calculator), start the clock or alarms, respond to user greetings, adjust brightness or volume, etc. The commands can be spoken aloud to carry out these actions. To finish the given task, the programme interacts with a server. It can synthesize and analyses speech, and in this case, English will be the supported language. Users may interact with their gadgets hands-free thanks to the voice assistant. It can be a blessing for users who are unsure on how to use a system. It can open other applications, show websites, and reply to greetings. The task at hand will then be processed when the spoken words have been heard, recognised, and converted to text. The machine executable commands will communicate with other system apps and give the user the necessary feedback so they know the task has been completed successfully. This project can be expanded to cover a wide range of applications. It offers cross-platform interoperability by utilising the Electron framework, so it may be used on personal computers running any operating system.

IV-   LITREATURE SURVEY

These days, personal assistants and virtual assistants play a crucial role in our daily lives. Every organization, or person, is converting to these technologies since they make it easier for them to complete their tasks. Based on a desktop application, this system. [15] This system comprises a virtual assistant that can accept user input, comprehend it, analyse it, and carry out duties as necessary. As a result, consumers can save a tonne of time. Speech recognition has a lengthy history and has undergone several significant innovation waves. On smartphones and wearable technology, speech recognition for dictation, search, and voice commands has become a standard feature. The most effective form of human communication, according to V. Radha, C. Vimala, et al., is speech. Speech recognition is the best method of recognition, therefore it makes people identical and makes it simpler for computers to identify them. [18]This has a great reputation and aids with autonomous speech recognition. Dynamic Time Warping (DTW), HMM, and other voice recognition methods are some of the most used. It provides a

collection of typical speech waveform vectors for feature mining of Mel Frequency Cepstrum Coefficients (MFCC). Studies have shown that MFCC is more accurate and realistic than other voice recognition mining techniques. According to T. Schultz and A. Waiel et al., as speech technology products proliferate across the globe, their inflexibility with regard to unfamiliar target languages has become a significant issue.[17] The research is significant because it focuses on the issue of how to quickly and efficiently migrate large vocabulary continuous speech recognition (LVCSR) systems. In the context of the Global Phone project, which investigates LVCSR techniques in 15 languages, the research needs to evaluate acoustic models for a novel destination language using speech information from various source languages, but only restricted data from the destination language identification outcomes using language-dependent, independent, and language-adaptive acoustic models are described and discussed.[20]

V -METHODOLOGY

Speech Recognition Library contains several built-in functionalities in this concept's practical method of constructing a personal voice assistant that will enable the assistant to comprehend Using Text to Speech capabilities, the user's request and the response will both be voice-returned to the user.[21] The algorithms used in the background will translate the user's spoken instruction into text when the assistant hears it. Additionally, the assistant will carry out the appropriate action in accordance with the keywords present in the text (based on the user's order). Thanks to the functions found in many libraries, this is achievable.[15] Additionally, the assistant used some APIs to do all of the capabilities. These APIs were utilised by us for a variety of tasks, including computations and the extraction of news from online sources.Thanks to the functions found in many libraries, this is achievable. Additionally, the assistant used some APIs to do all of the capabilities. These APIs were utilised by us for a variety of tasks, including computations and the extraction of news from online sources. So, to parse the JSON Data coming from websites into string format, we utilised a library named JSON. This allows us to gather news from online sources and pass it as input to another function for later use. Additionally, we have a variety of libraries, each corresponding to a distinct technology, including libraries like Random. We implemented operating system-related features, such as shutting down a system, using the library OS. or resetting a device. A

library called pyautogui has been developed for features like screenshotting.A library called psutil is used for features like checking the battery's state.

Fig 1: Working of Personal Assistant

Natural language processing (NLP) is a technique used by virtual assistants to translate user text or voice input into actionable commands. Natural language audio signals are translated into executable commands or digital data that may be processed by the software when a user requests their personal assistant to complete a task. To determine an acceptable response, this data is then compared with software data. Machines can be operated using your own commands by using a virtual assistant. We employ a variety of Python installer packages, such as Speech recognition, gTTS, pipwin, etc., to create virtual assistants. In speech recognition, audio is transformed into text. This is frequently employed by voice assistants like Siri, Alexa, and others. Python offers a SpeechRecognition API that enables us to translate voice or audio commands into text for later processing. According to the flowchart above, users must first provide a command to interaction entities like laptops and PCs before those entities may listen to and understand it. In order to further analyse the process, compare this command to the cloud where our data is already stored.In the event that the request matches, the output is produced in both text and speech form a cloud's data Look up the procedure or piece of logic that will be used to process the request, then transmit the result of the backend operation as a response.
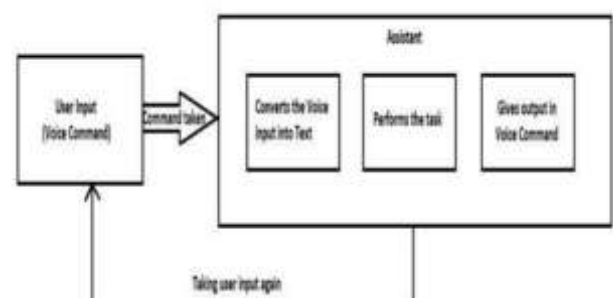

Fig 2: Block Diagram

International Journal of Innovations in Engineering and Science,   www.ijies.net

VI -TECHNOLOGIES USED

Python was chosen as the programming language for this project because of its adaptability and accessibility to a large number of libraries. We used Python programming language-supporting Microsoft Visual Studio Code (IDE) to create the Virtual Assistant. Python has a speech recognition package that includes certain built-in functions. We will first define a function that will turn the text into speech. We employ the pyttsx3 library for that. We'll set the library instance's initial value to a variable. We employ the say() method and supply the text as an argument; the result is a vocal response. Another function has been defined to identify the user's voice command. We provide the microscope source in that function, use the appropriate functions within its bounds, and save the output in a variable. We can use a variety of services for the entire process, including the Microsoft Bing and Google Speech Recognition engines as well as goods made by other major corporations like IBM, Humidify, etc. We select Google's Speech Recognition Engine for this project since it will translate each analogue voice command into a digital text format. The Assistant will look for the keyword after receiving that text as input. The relevant function will be called and conduct the action accordingly, such as telling the time, date, or battery status, taking a screenshot, saving a quick note, and many more actions, if the input command contains a word that matches the relevant term. The main benefit of this personal virtual assistant is that it saves a lot of time and can even respond to questions from people with various accents. There is no requirement that one must issue a precise command to initiate a certain action. The user has the freedom to offer commands in their own language. Python 3.8.1 was used to create this voice-activated personal assistant for the PC. And Microsoft Visual Code was the IDE (Integrated Development Environment) that we employed.

VII -WORKING MODEL

There are three modules in this Assistant. The first is that the assistant will accept user voice input. The user input is then analyzed and mapped to the corresponding intent and function. The third option is the assistance speaking the user's outcome throughout. The assistant will initially begin taking user input. The assistant will translate the analogue voice input into digital text after receiving it. The assistant will start asking the user for input again if it was unable to translate the voice into text. If converted, it will begin to analyses the input and map it

to a certain function. Later, the user will receive the output by voice command. The goal of this study is to provide users with the simplest and most efficient way to interact with the system via voice instructions delivered in natural languages. This project gets rid of the exhausting setups and configurations, as well as system overheating, which eventually reduces performance. The user feels as though he is conversing with another human when speaking in normal language to the system, and the system's answer causes him to forget that a machine is actually doing all the work.
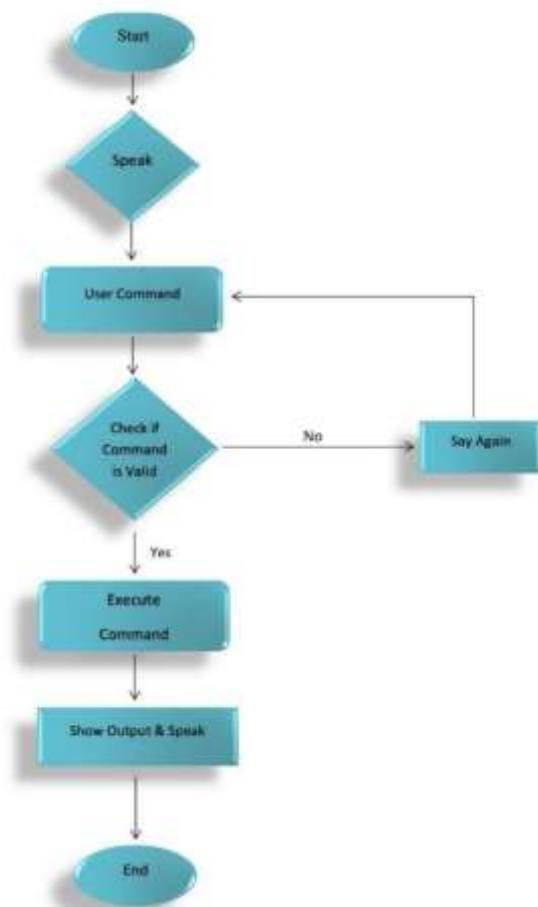

Fig 3: Flowchart for working model

HARDWARE AND SOFTWARE REQUIREMENTS.

1. Pc/Laptop
2. 8GB RAM or More
3. i3 Proccessor or more
4. Windows 11 or more
5. 256GB ROM
6. Visual Studio Code
7. Python 3.10.5
8. Internet Connection

International Journal of Innovations in Engineering and Science,   www.ijies.net

Modules Imported In the Project

A. Subprocess:- This module is used for getting system subprocess details which are used in various commands i.e Shutdown, Sleep, etc. This module comes built-in with Python.

B. WolframAlpha:- It is used to compute expert-level answers of any command using Wolfram's algorithms, knowledgebase and AI technology.

C. Pyttsx3:- pyttsx3 is a text-to-speech conversion library in Python. It works offline and is compatible with python like Python 2 and 3 works without internet connection or delay. The text-to-speech

VIII -CONCLUSION

This report provides a thorough description of the design and implementation of a voice-enabled personal computer assistant in Python. In comparison to earlier times, this voice-activated personal assistant will be more efficient at saving time in today's lifestyle.

The key characteristic of this Personal Assistant is its simplicity of usage. The Assistant effectively completes some duties that users assign it. Additionally, this assistant can perform a wide range of tasks, including estarting or turning off our PC with a single voice command and reading out the most recent news. All generations of individuals, including those with special needs or impairments, will benefit from this project. The personal voice assistant will be simple to use and will eliminate the need for manual labour to do a variety of activities. The present voice assistant system's capability is restricted to functioning online (requires an internet connection to accomplish tasks) and on desktop computers. Because the voice assistant system is modular, new features can be added without affecting existing system functionality. The project is extremely

beneficial and has broad application possibilities across numerous sectors. Although the programme primarily focuses on how to use voice commands to operate the personal assistant on an Android phone, the idea of voice recognition can be applied in a variety of industries as it is often more convenient, saves a lot of time, and is especially useful for those who find it difficult to perform manual tasks. Consequently, the idea is limited to Android application programming. Additionally, those who prefer voice operation and those who have trouble or a disability with manual operations can benefit from

the programme that uses voice commands. The program's main goal is to provide voice services, and doing so makes it more accessible to a wider audience.

REFERENCES

[1]   *Ahmed J Abougarair,1 Mohamed KI Aburakhis,2 Mohamed O Zaroug1 1 Department of Electrical and Electronic Engineering, University of Tripoli, Libya 2 Department of Engineering Technology, Clark State College, USA*

[2]   *Abeed Sayyed1 , AshpakShaikh2 , AshishSancheti3 , Swikar Sangamnere4 , Prof. Jayant H Bhangale5 UG Scholar, Department of Mechanical Engineering1,2,3,4 Faculty, Department of Mechanical Engineering 5 Matoshri College of Engineering and Research Centre, Eklahare, Nashik*

[3]   *Damarla Kavya, Daddanala Suvarna, Javisetti Srinivas and Chintha Venkata Ramaiah from the Department Of Electronics And Communication Engineering, Chalapathi Institute Of Engineering And Technology, Guntur, Andhra Pradesh.*

[4]   *Dhiraj Pratap Singh, Deepika Sherawat, Sonia School of Computer Science & Engineering, Galgotias University Greater Noida-203201*

[5]   *Song Qunying, Bachelor in Computer Software Development 2012 Shen Hui, Bachelor in Computer Software Development 2012 School of Health and Society Department Computer Science Kristianstad University SE-291 88 Kristianstad*

[6]   *Mr. Abhijit J. Patankar Research Scholar, Computer Science and Engineering, VTU, Belgaum, Karnataka.*

[7]   *Divisha Pandey1, Afra Ali2, Shweta Dubey3, Muskan Srivastava4, Shyam Dwivedi5, Md. Saif Raza6 1, 2,3,4, Student of B. Tech fourth year, Department of Computer Science and Engineering, Rameshwaram Institute of Technology & Management, Lucknow, India*

[8]   *Jatu Naazneen Abdul Gaffar Dtss College Of Commerce*

[9]   *P Srinivas B. Tech IV year, Dept. of Electronics and Communication Engineering, Gandhi Institute of Technology and Management, Hyderabad, Telangana, India*

[10]   *Mrs.A.M.Sermakani.,M.E.,(Ph.D)1 , J.Monisha2 , G.Shrisha3 , G.Sumisha4 UG Scholar Department of Information Technology,S.A Engineering College,Chennai2-4*

[11]   *1Nivedita Singh, 2Dr. Diwakar Yagyasen, 3Mr. Surya Vikram Singh, 4Gaurav Kumar, 5Harshit Agrawal 1,2,3,4,5Department of CSE, Babu Banarasi Das National Institute of Technology and Management, Lucknow, India*

[12]   *Harkishen Singh Muskan Khedia Jayashree Panda Subham Mishra Ankit Singh college of Engineering and Technologo Bhuvneshwar.*