AI Voice Assistant

Ayush Thorat, ShivaniSawarkar, Yashshree Mendhe, Manisha Vaidya Alok Gajbhiye

[1] *Ayush Thorat, Artificial Intelligence, Priyadarshini J.L College of Engineering, Nagpur, Maharashtra, India*
[2] *ShivaniSawarkar, Artificial Intelligence, Priyadarshini J.L College of Engineering, Maharashtra, India*
[3] *Yashshree Mendhe, Artificial Intelligence, Priyadarshini J.L College of Engineering, Maharashtra, India*
[4] *,Alok GajbhiyeArtificial Intelligence, Priyadarshini J.L College of Engineering, Maharashtra, India*
[5.] *Manisha Vaidya, Artificial Intelligence, Priyadarshini J.L College of Engineering, Maharashtra, India*

ABSTRACT

This paper introduces a sophisticated Voice recognition technology has witnessed significant advancements in recent years, enabling the development of intelligent virtual assistants capable of understanding and responding to natural language commands. Among these assistants, Jarvis, inspired by the AI assistant in popular culture, represents an intriguing concept. In this paper, we aim to review existing research and advancements in the field of voice recognition, focusing on Jarvis-like models. We explore the underlying technologies, architectures, applications, challenges, and future directions in this burgeoning area of research.

Keyword : *Voice Assistant, NLP, Neural Network, Google Search*

1. INTRODUCTION

In the digital age, technology has advanced to make our lives easier and more efficient. One such technology is voice recognition, which allows us to interact with our devices through speech. Jarvis is a popular voice recognition model that works for personal laptops and computers. In this paper, we will explore what Jarvis is, how it works, its benefits and limitations, and its impact on personal computing. Voice recognition technology has become an integral part of daily life, with applications ranging from virtual assistants in smartphones to smart home devices and automotive systems. The concept of a Jarvis-like assistant, as popularized by science fiction, has sparked considerable interest in developing intelligent voice recognition models capable of performing various tasks.

In today's fast-paced world, technology has become an integral part of our daily lives. From smartphones to smart homes, we are constantly surrounded by devices that make our lives easier and more efficient. One such technology that has gained popularity in recent years is voice recognition software. Voice recognition software allows users to interact with their devices using spoken commands instead of typing or clicking. This technology has been widely used in smartphones and smart speakers, but now it has also made its way to personal laptops and computers. One such voice recognition software is Jarvis, which is a voice recognition model that works on personal laptops and computers.

2. PROBLEM DEFINITION

In an increasingly interconnected world, technology plays a pivotal role in simplifying tasks and enhancing productivity. Voice assistants have emerged as a convenient interface for interacting with digital devices and accessing information through natural language. 3.

3.LITERATURE REVIEW

Literature survey is also commonly known as literature review. This part of the paper is an explanation and understanding of all the academic material that is used to gather information about our project.
We are collect some review from the article which is mentioned below :-
Kumar, C. Raju. "Virtual assistant using artificial intelligence and Python." J. Emerg. Technol. Innov. Res 7.3 (2020): 1116-1119.
Gupta, Ujjwal, et al. "Desktop Voice Assistant." International Journal for Research in Applied Science & Engineering Technology (IJRASET) (2022).
Umapathi, N., et al. "DESKTOP'S VIRTUAL ASSISTANT USING PYTHON."
Jain, Shubham, Shreya Joshi, and Ruchi Parashar. "ALPHA: Personal Assistant in Chat Box." (2021).
Sahu, Ajay, et al. "Voice Assistant Using Artificial Intelligence." Available at SSRN 4384623 (2023).

4. PROPOSED WORK

The voice assistant initiates voice mode and prompts the user to provide input in voice/text format for best results from the voice assistant. As this program can also be controlled with your phone with help of an application 'WO-MIC', it just turns any android phone into a wireless microphone and helps in the reduction of unwanted noise in the environment.
• Using this application, which is Wikipedia's search engine, users can contact the wizard and the wizard will retrieve the data from the internet. The results are displayed in the console window in audible format, up to a limited number of lines.
• Getting Current News about his/her motherland, about world, about technologies, about sports or about entertainment of the industry and much more, the user can easily get the news just by giving voice input to assistant to open news so it will open new tab and it can also fetch the data from the websites and return it to the console and read out for user without any labor.
• Weather Forecast, through this feature users can see the weather forecast for any location. In addition, the temperature and humidity of Kelvin will return the weather.•
Open Applications like , YouTube, google search engine , launching websites , system applications with the help of web browser python library and os for opening system applications(like, code editor, notepad,chrome,etc.)
• Close Applications, the application work perfectly by providing a command 'TASKKILL/ F/im file.exe'. The assistant close that application asked to close.
• Automation, the application performs automation for YouTube and any search engine with the help of keyboard python library. The user just need to give input and the assistant will perform the automation ask.
• Voice Assistant caneven repeat the user's words by takeCommand function and speak function.
• WhatsApp Messages, the application work by taking mobile number of the receiver or the name of the receiver, message to send , time when to send as a query. As the result , voice assistant will send the message and inform you. This is done with help of pywhatkit python library. And the history of messages will be saved in pywhatkit database file .
• Sending Mails, this feature allows users to send an email to someone whose contacts include an email address. It then sends the successful execution of the task back to the user via the Hearing Assistant.
• Listening to music, the voice assistant plays the music requested by the user, either from the user's system or through an online search, without the user having to do it.
• TimeTable Notification, the voice assistant will remind you the work accord-ing to the user's time table schedule and as a result it also give notification with the help of notify python library .
• The Voice Assistant can answer any query with the help of wikihow python library and wolframalpha algorithm.
• A setting alarm is a basic function of any device, this allows the user to set the alarm to a specific time.
• Chatbot, this feature communicates with the user on a case-by-case basis. It also works whenever the user provides voice input to the assistant and the user receives the output in the voice response of the voice assistant ChatBot.
• ScreenShot, this feature allows the user to take a screenshot of the current window or photo or current

5. OBJECTIVES

The voice assistant seeks to enhance the overall user experience by offering a more intuitive and personalized interaction model. Through natural language understanding and context awareness, the assistant aims to provide relevant and timely responses that cater to the user's needs and preferences.
One of the key objectives of a voice assistant is to make technology more accessible to a wider range of users, including those with disabilities or limited technical expertise. Voice-based interaction eliminates the need for manual input, making it easier for users to access information and perform tasks using natural language.

6. METHODOLOGY

In Virtual Voice Assistant initially, we used SAPI5 API, a Python package called pyttsx3 converts text into speech. It is compatible with Python 2 and 3 and functions offline, unlike competing libraries. To enable the use of speech synthesis and recognition in Windows programs, Microsoft created the Speech Application Programming Interface, or SAPI. Next comes the definition of the program's main function, which contains all of its capabilities. The following functionality is intended to be included in the suggested system: While continuing to listen for commands, the assistant requests feedback from the user. A user's request can be followed when determining the listening time.

(a) The assistant will repeatedly ask the user to repeat the command if it does not understand it well.

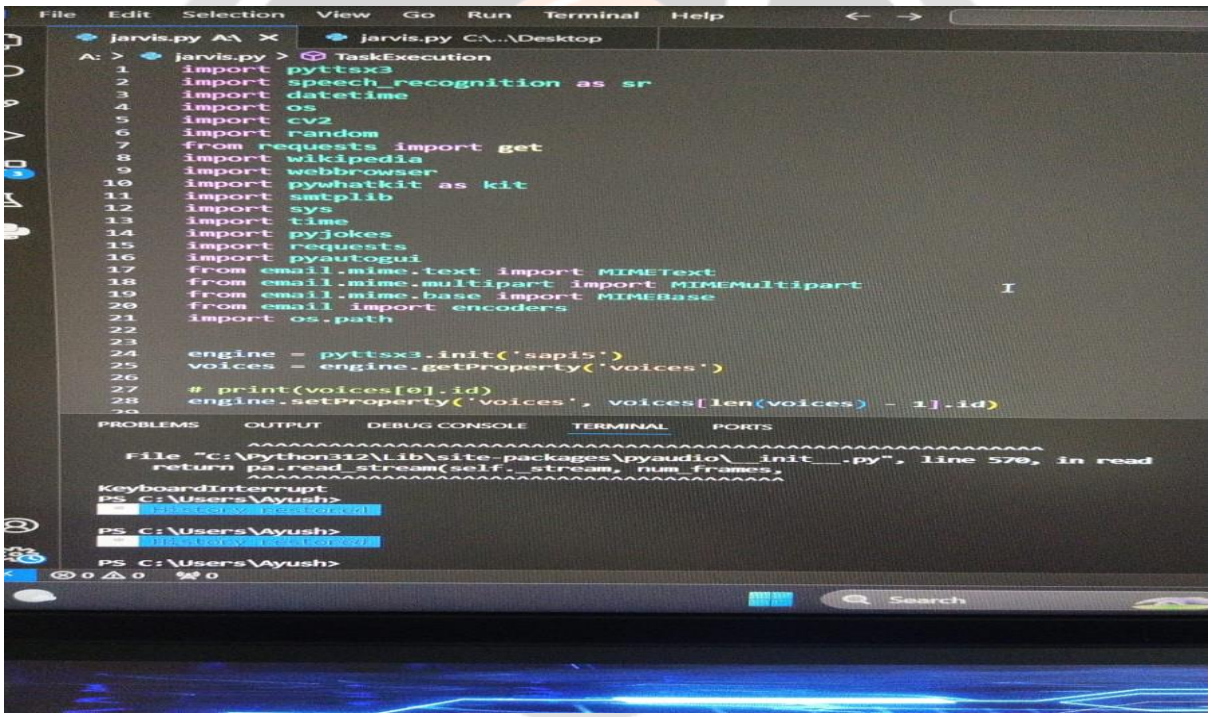
(b) The user has the option to customize the voice of this assistant to be either male or female.

(c) The assistant's current version includes functions like searching Wikipedia, opening apps, checking the clock, taking notes, showing notes, opening, and closing YouTube, Google, and opening and closing Programmes

7. WORKING

Step1.The following are the results obtained by Virtual Voice Assistant .As we give command in voice to assistant it will performs actions accordingly .These are the screenshots of outputs where the input commands are in voice and in return outputs are in tasks that are told to be followed

Figure 1: Listening to user command



Step2 : Opening YouTube on command. The assistant will also open different applications based on the voice command ,the keyword being “open YouTube” so whenever the user will mention the keyword the assistant will recognize it will do the further actions. Like shown in picture as soon the user said the command YouTube has been successfully opened on the system.

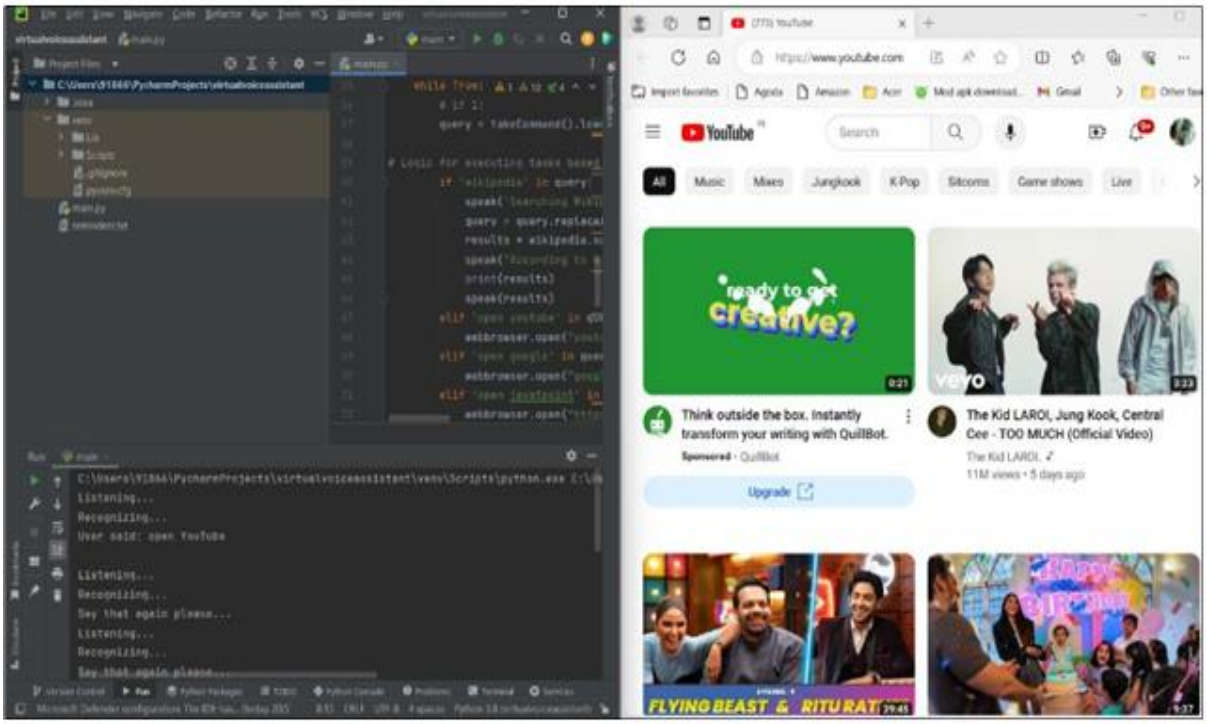
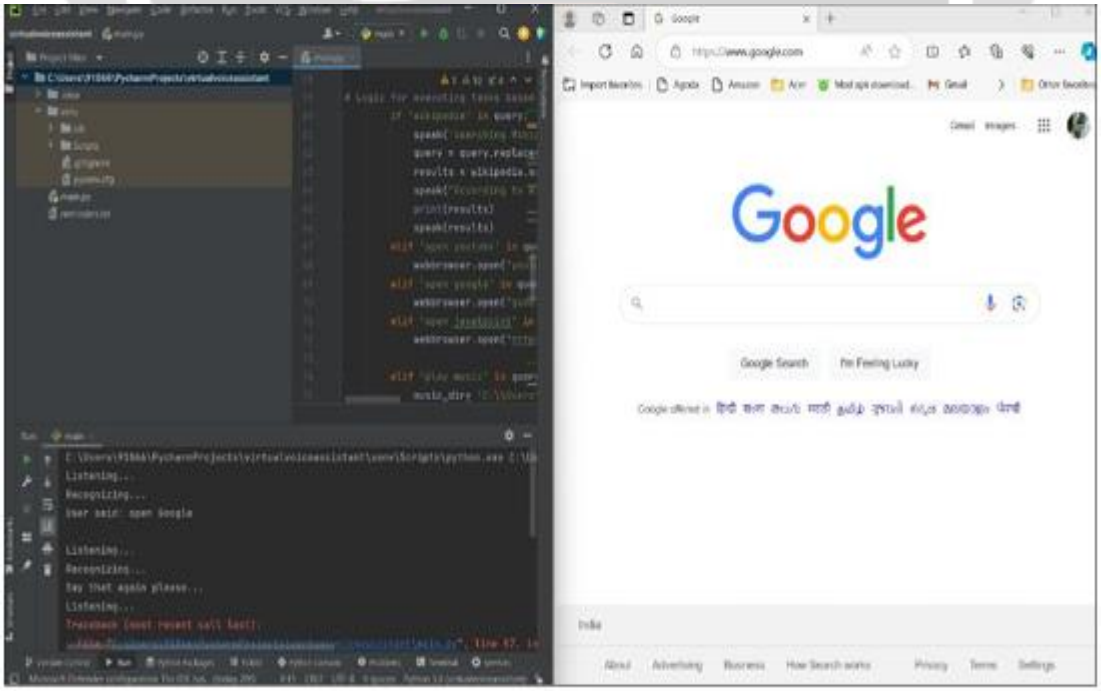
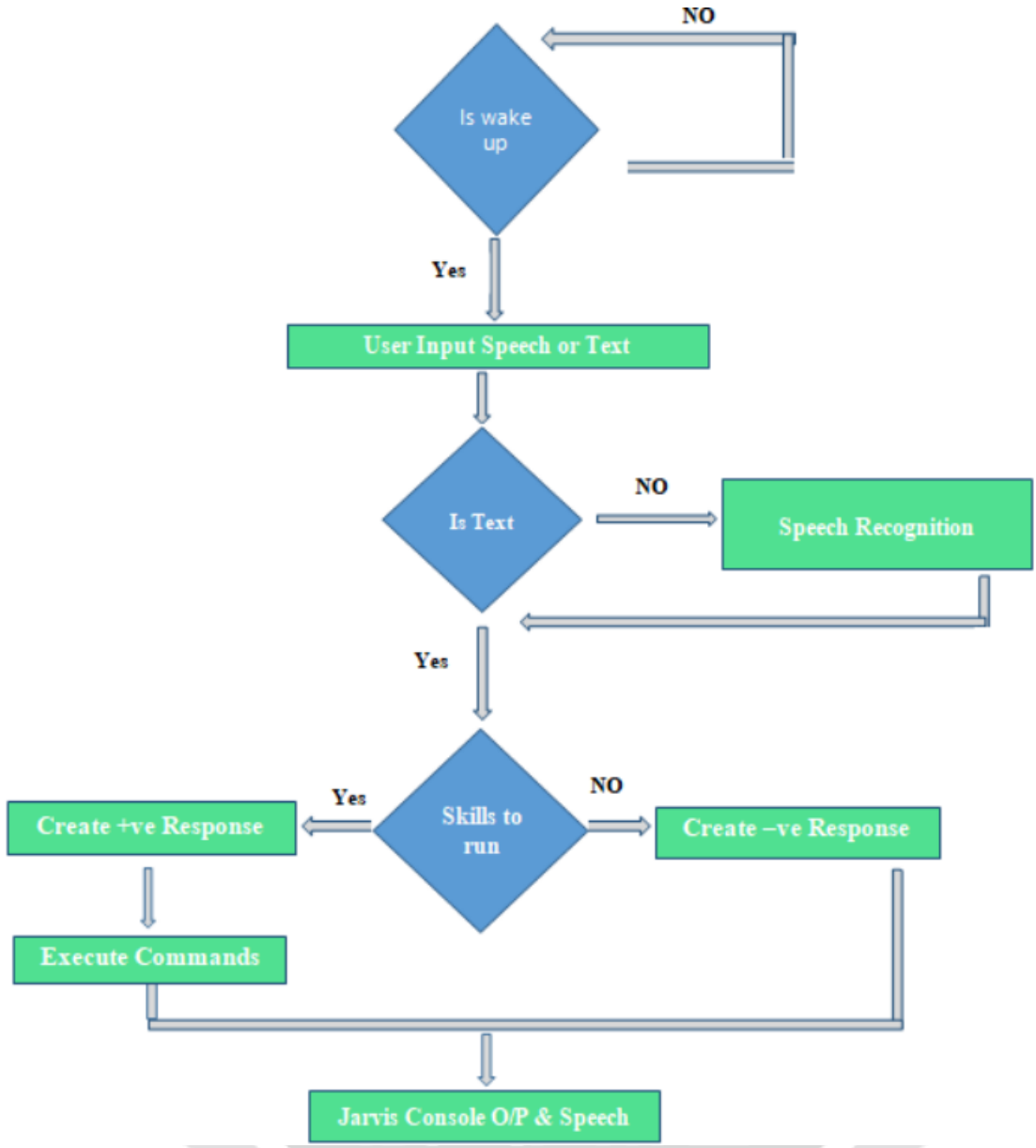


Figure 2: Opening YouTube on command.

Step3:The assistant will also open different applications based on the voice command ,the keyword being “open YouTube” so whenever the user will mention the keyword the assistant will recognize it will do the further actions. Like shown in picture as soon the user said the command YouTube has been successfully opened on the system.

Figure 3: Opening Browser on command.





Flowchart:

8.Advantages

Convenience: Users can interact with Jarvis hands-free using voice commands, eliminating the need for manual input via keyboards or touchscreens. This convenience is particularly valuable in scenarios where users have limited mobility or are engaged in other tasks.

Efficiency: Jarvis streamlines tasks by providing quick access to information, managing schedules, controlling smart devices, and performing various actions with minimal effort. Its ability to understand natural language commands enhances user productivity by reducing the time spent on manual interactions.

Personalization: Through machine learning algorithms, Jarvis can learn user preferences and tailor responses and recommendations accordingly. This personalization fosters a more customized and engaging user experience, enhancing benefit from technology.

9.Disadvantages

Privacy Concerns: Voice assistants like Jarvis often collect and store sensitive user data to improve performance and personalize interactions. However, this raises privacy concerns regarding the security and confidentiality of personal information, leading to potential breaches or misuse.

Security Risks: Hackers could exploit vulnerabilities in Jarvis's infrastructure or voice recognition system to gain unauthorized access to user data or perform malicious actions. Security breaches could result in identity theft, financial loss, or other damaging consequences.
Accuracy and Understanding: Despite advancements in natural language processing, Jarvis may still struggle to accurately interpret complex or ambiguous voice commands, leading to misunderstandings or incorrect responses

10. APPLICATION

Education: Jarvis AI can personalize learning experiences for students by adapting educational materials to their individual needs and preferences. It can also provide tutoring, answer questions, and offer feedback on assignments.
Smart Homes: In smart homes, Jarvis AI can control various connected devices, such as thermostats, lights, security cameras, and appliances. It can respond to voice commands, learn user preferences, and automate routine tasks to enhance convenience and energy efficiency.
Content Creation: Some content creators use AI tools like Jarvis to generate written content, such as articles, blog posts, and product descriptions. These tools can help speed up the writing process and generate ideas for new content.
Research and Analysis: Researchers and analysts use Jarvis AI to gather information from various sources, analyze data sets, and generate insights. This can be particularly helpful in fields such as market research, competitive intelligence, and scientific discovery.

11. FUTURE SCOPE

Enhancement in the capacity of database or the data training sets can be done in this for more situations or the acquaintances that can be faced by JARVIS. This would upgrade its effectiveness and the wide range ability of producing responses. Further addition of more voices can also be done as an additional feature. So these limitations can be broken with the increase in data training sets. The interface
of the system can be improved more or we can say can be optimized. From saying more optimized it is meant that the interface can be more user friendly, comprehensive and easy to use for more percentage of users. So the Jarvis would become more accessible and intractable
.

12. CONCLUSION

Jarvis - An AI Voice Assistant System uses speech recoginitio, gTTs and other AI techniques along with Neural Networks and Natural Language Processing for a smart responsive system to the given circumstances or conditions. It can reduce the workload of basic human activities or the daily activities and can replace some human working posts like personal secretaries employed for scheduling a person's per day time table. Critically, the system is designed to interrelate with other sub-systems smartly and comprehensively. The system will have the following phases: Input phase in which data or query given in form of text or speech, Interpretetion of voice to text, Processing and storing of data, producing output in the form of voice from the refined text to Jarvis console. The information produced at each step can then be used to retreive patterns and analyze them for later use. This could be the main basis for artificial intelligence machines to learn and recognize patterns for people. So, based on a literature study and analysis of persisting systems, the conclusion is derived that our provided system will not only facilitate interaction with systems and modules but also keep users more organized.

13. REFERENCES

[1] Alotto, F., Scidà, I., and Osello, A. (2020). "Building modeling with artificial intelligence and speech recognition for learning purpose." Proceedings of EDULEARN20 Conference, Vol. 6.