



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 4, April 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com

Desktop Voice Assistant Integrated with ChatGPT

K Raja Simha Reddy, S. Kirubakaran

U.G. Student, Department of Computer Engineering, CMR College of Engineering and Technology, Hyderabad,
Telangana, India

Professor, Department of Computer Engineering, CMR College of Engineering and Technology, Hyderabad,
Telangana, India

ABSTRACT: This project aims to develop a voice-based desktop virtual assistant (using artificial intelligence) leveraging advanced machine learning algorithms, speech recognition, and contextual understanding technologies. The primary objective is to enhance the capabilities of desktop voice assistants by integrating with ChatGPT. This project offers desktop assistant tasks such as knowing the weather reports of various countries, knowing the daily news, opening browsers and various musical apps, and even more. It includes automatic Speech Recognition (ASR), which is used for automatic caption generation for videos, etc. Automatic Speech Recognition (ASR) is a NLP task used to detect accurate words from the audio that is used as an input. It interprets the data accurately. It also stores the historical input data, so it increases the user experience. As it is integrated with ChatGPT, it leverages productivity. This project mainly highlights the advantages and capabilities of using a desktop voice assistant integrated with ChatGPT. It provides adaptive assistance at the tip of our finger.

KEYWORDS: Artificial Intelligence, Desktop Voice Assistant, Automatic Speech Recognition (ASR), NLP, ChatGPT.

I. INTRODUCTION

In today's digital age, the integration of artificial intelligence (AI) into everyday tasks has become increasingly advantageous, with virtual assistants serving as one of the most important applications. These intelligent systems, capable of understanding and responding to human commands, have evolved to become great tools for enhancing productivity and convenience. This project seeks to further advance the capabilities of desktop virtual assistants by leveraging cutting-edge technologies such as machine learning algorithms, speech recognition, and natural language processing (NLP).

The primary focus of this endeavor is to develop a voice-based desktop virtual assistant that seamlessly integrates with ChatGPT, a sophisticated language model developed by OpenAI. By harnessing the power of ChatGPT alongside advanced machine learning algorithms and speech recognition technologies, the aim is to create an assistant that not only understands user commands but also provides intelligent and personalized responses.

Key functionalities of the proposed desktop virtual assistant include retrieving weather reports for various locations, delivering real-time news updates, and facilitating access to music browsing and playback applications. Central to the functionality of this assistant is Automatic Speech Recognition (ASR), which accurately transcribes spoken words into text, enabling seamless interaction with the system, and it also produces the output in a text file. Moreover, the integration with ChatGPT allows for natural language understanding and adaptive assistance, enhancing the user experience and productivity.

This introduction sets the stage for a comprehensive exploration of the project, highlighting the significance of leveraging AI, ASR, NLP, and ChatGPT to create a powerful and intuitive desktop virtual assistant. Through this integration, users can expect a seamless and efficient means of accessing information, performing tasks, and enhancing overall workflow productivity.

II. RELATED WORK

A) Voice Control Using AI-Based Voice Assistant by S.Subhash et al

The study mentioned, conducted by S. Subhash et al. in 2020, titled "Voice Control Using AI-Based Voice Assistant," highlights the implementation of artificial intelligence (AI) and voice technology to automate and control tasks. The voice assistant developed in the study captures audio from a microphone, converts it into text, and then utilizes Google Text-to-Speech (GTTS) to translate the text into an English audio file. This audio file is then played using the Python programming language's play sound module.[3]

The key features of the voice assistant developed in the study include:

- 1. AI-based Voice Assistant:** The voice assistant is powered by artificial intelligence, enabling it to understand and respond to voice commands effectively.
- 2. Intelligent Personal Assistant:** The study introduces an Intelligent Personal Assistant feature, which utilizes Voice User Interface (VUI) to listen to and process audio commands. This allows users to control various tasks, such as turning on and off smartphone applications, using voice commands.
- 3. Voice User Interface (VUI):** VUI enables users to interact with devices and applications through voice commands. It processes audio input, interprets commands, and executes actions accordingly.
- 4. Automation of Tasks:** The voice assistant automates tasks by converting voice commands into actionable instructions, thereby enhancing user convenience and efficiency.
- 5. Seamless Operation:** The study suggests that the voice assistant operates without causing any trouble, indicating smooth functionality and reliability.

Overall, the implementation of AI-based voice assistants, as demonstrated in the study, showcases the potential of voice technology to streamline tasks and improve user experience. Features such as Intelligent Personal Assistant and VUI contribute to making interactions more intuitive and efficient. This research highlights the advancements in voice technology and its applications in various domains, including automation and control.

B) Voice-based person identification using mean clustering and medoid clustering approaches

The paper titled "Voice-based person identification using mean clustering and medoid clustering approaches" by Anil B V et al. focuses on utilizing clustering algorithms for voice-based person identification. The authors explore three classes of c-means algorithms (Hard c-means, fuzzy c-means, and rough c-means) and compare them with three classes of medoid algorithms (hard c-medoid, fuzzy c-medoids, and rough c-medoid).[4]

The findings reveal that the fuzzy c-medoid algorithm outperforms other voice recognition systems in terms of recognition accuracy. This suggests that clustering techniques, particularly fuzzy c-medoid, hold promise for improving the effectiveness of voice-based biometric recognition systems.

The paper emphasizes the importance of robust biometric recognition systems in minimizing security threats and authenticating individuals with a high level of confidence. While traditional validation methods such as identity cards or memorized entities like PINs or passwords are effective, they can be prone to vulnerabilities such as loss or forgetfulness.

Biometric recognition systems, on the other hand, rely on measuring physical or physiological features with precision to identify individuals accurately. Voice recognition, as a biometric modality, offers advantages in terms of universality, permanence, collectability, and acceptability. These attributes make voice-based recognition systems suitable for various applications, including security authentication and access control.

In summary, the paper highlights the significance of clustering algorithms, particularly fuzzy c-medoid, in improving voice-based person identification. It underscores the importance of biometric recognition systems in

enhancing security and authentication processes, while also emphasizing the advantages of voice recognition in terms of its universal applicability and user acceptance.

III. METHODOLOGY

When the desktop assistant runs, it takes the input from the user (voice), which is actually called a command. By using machine learning algorithms and Natural Language Processing(NLP) it reframes the command to identify a syntactical meaning and identify the keyword in the command to perform a specific operation. In case they are not identified, the voice assistant throws a warning and waits for another input. This process continues until a valid input is found.

The following steps are involved in implementation:

Initially, we need to load all the required modules, such as Appopener, OpenAI, Speech Recognition, Web Browser, Tkinter, PIL, etc.

1. Initialize GUI: This step involves setting up the graphical user interface (GUI) components, including the text area, microphone image, and stop button. It prepares the interface for user interaction.

2. Start Listening: Once the GUI is initialized, the program starts listening for voice commands. This is represented as a decision point because the program continuously checks whether it should keep listening or stop based on user input.

3. Process Command: When the program is actively listening, it captures audio input from the microphone, attempts to recognize speech, and executes the corresponding command based on the recognized speech. This step includes various actions, such as opening websites or applications, retrieving weather information, displaying news, or speaking the current time.

4. Execute Command: After processing the command, the program executes the corresponding action. This may involve interacting with external services or updating the GUI to provide feedback to the user.

5. Stop Listening: At any point, the user can choose to stop the listening process by clicking the "Stop" button. This decision point determines whether the program should continue listening or stop based on user input.

6. Close GUI: Finally, when the user decides to stop listening and exits the program, the GUI is closed, and the program ends.

Modules:

Speech Recognition:

Speech Recognition package basically a wrapper for several popular speech APIs, it includes the Google Web Speech API, openai whisper api and many more which comes with a default API key already integrated into the [1]SpeechRecognition package. This means you can try without signing up for a service. This level of flexibility and easy to integration make it better choice for python projects.

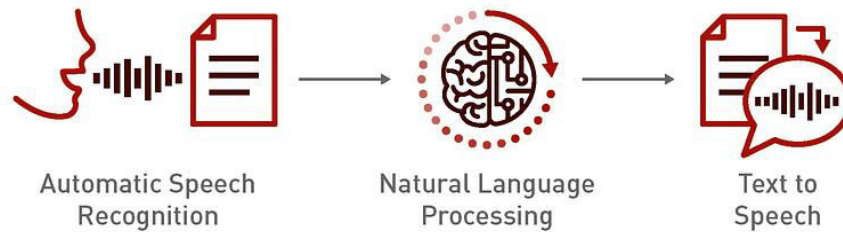


Fig-2 Speech recognition

Openai:

The OpenAI Python library offers seamless integration with the OpenAI API, allowing developers to access its functionality directly from Python applications. This library provides a predefined set of classes for API resources, which dynamically initialize themselves from API responses. This design ensures compatibility with various versions of the OpenAI API, simplifying the development process for users. With the OpenAI Python library, developers can easily leverage the power of the OpenAI API to incorporate natural language processing and AI capabilities into their Python applications with ease and efficiency.[2]

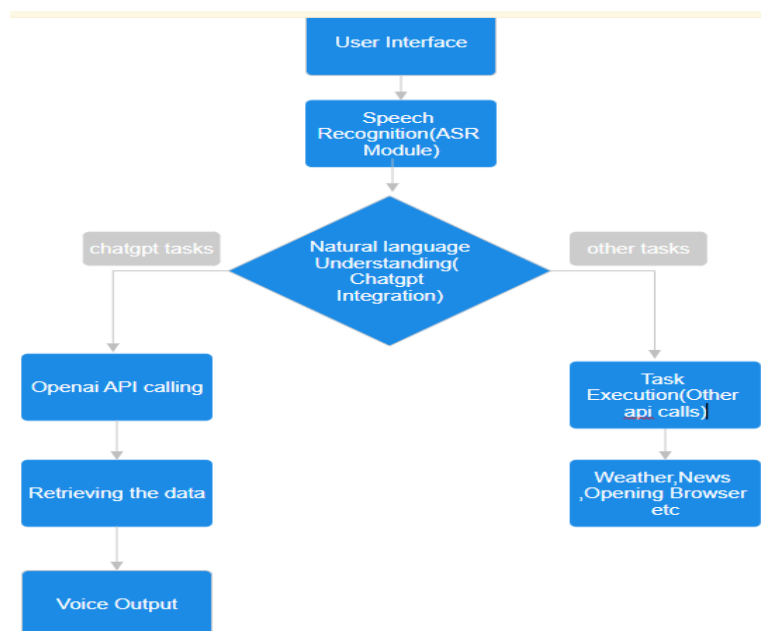


Fig.1 Architecture



IV. EXPERIMENTAL RESULTS

Integration of Advanced Technologies: The project proposes integrating cutting-edge technologies such as machine learning, speech recognition, and natural language processing (NLP) to develop a sophisticated voice-based desktop virtual assistant. By leveraging these advanced technologies, the virtual assistant can understand user commands, execute tasks, and provide personalized responses, enhancing user experience and productivity.

Utilization of ChatGPT: Integrating ChatGPT, a powerful language model developed by OpenAI, into the virtual assistant offers several advantages. ChatGPT enhances natural language understanding and response generation capabilities, allowing for more intuitive and conversational interactions with users. Additionally, ChatGPT's continuous learning capabilities enable the virtual assistant to improve its performance over time, adapting to user preferences and refining its language understanding abilities.

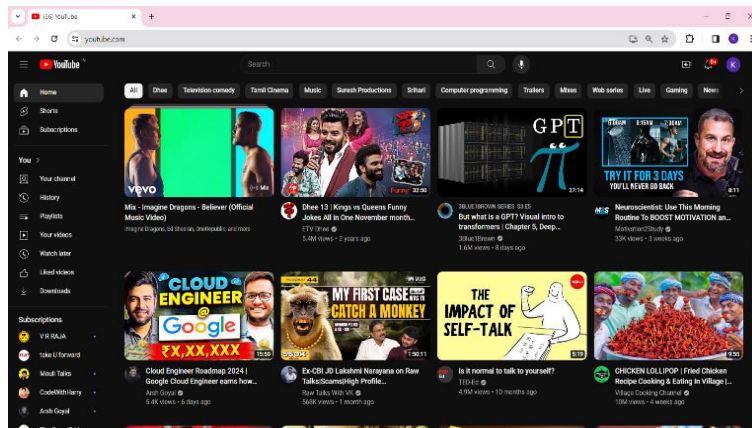


Fig-3 Opening youtube

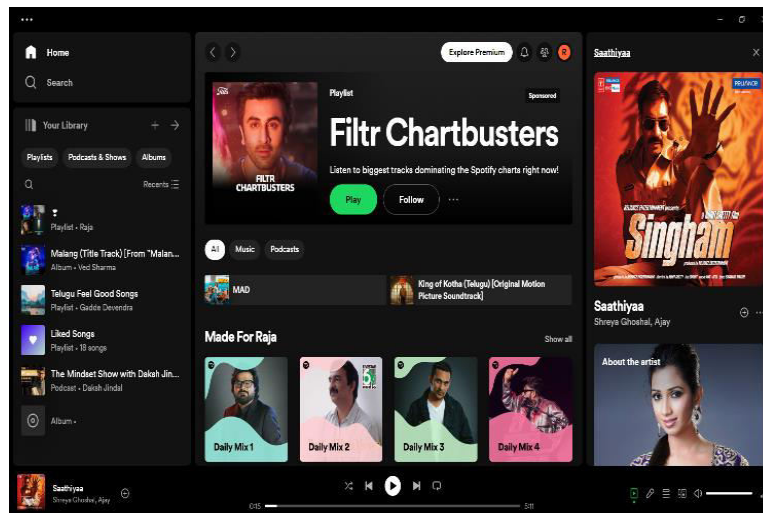


Fig-4 Opening Music App

V. CONCLUSION

The desktop voice assistant project has successfully crafted a sophisticated software solution employing artificial intelligence and natural language processing to interpret and act upon voice commands on desktop computers. This innovation offers a seamless and effective tool, meeting the diverse needs of our user base and receiving commendations from users.

Despite the many advantages of personal desktop voice assistants, there are several security and privacy issues to be aware of. To guarantee the security of their personal data, users must be informed about the data being gathered and how it be used. All things considered, personal desktop voice assistants hold great promise for improving our everyday lives and streamlining our digital connections. It will be interesting to watch how these helpers develop and advance in the coming years as technology keeps advancing.

Through meticulous evaluation and ongoing refinement based on user input, we have ensured that our voice assistant remains a versatile and indispensable asset across a wide spectrum of applications. We see this project as a pivotal stride forward in integrating voice assistant technology into desktop computing, and we eagerly anticipate further advancements in this realm.

REFERENCES

- [1] Speech recognition at: <https://pypi.org/project/SpeechRecognition/>
- [2] Open ai at: <https://platform.openai.com/docs/introduction>
- [3] S. Subhash et al., "Voice Control Using AI-Based Voice" Assistant," 2020 International Conference on Smart Electronics and Communication (ICOSEC), Bangalore, India, 2020, pp. 592-596, Doi:10.1109/ICOSEC49019.2020.9230327
- [4] A. B. V and R. M. S, "Voice Based Person Identification Using c-Mean and c-Medoid Clustering Techniques," 2020 IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER), Udupi, India, 2020, pp. 121-126, doi: 10.1109/DISCOVER50404.2020.9278042
- [5] S.S. Biswas Potential use of chat GPT in global warming Ann. Biomed. Eng. (2023), pp. 1-2
- [6] Ankit Lal Sinha; Hardik Muley; Jaydeep Ghosh; Padmavati Sarode. "AI based Desktop Voice Assistant for Visually Impaired Persons", 2023.
- [7] Jarvis-Chatgpt at <https://github.com/gia-guar/JARVIS-ChatGPT>
- [8] Muneer, S., & Khan, F. A. (2022). Artificial Intelligence and Its Applications in Healthcare. International Research Journal of Modernization in Engineering, Technology and Science, 1(4), 23-27.
- [9] Oracle. (2018). Advantages and Disadvantages of Voice Assistants for Marketers. Oracle Marketing Cloud Blog.
- [10] Mukherjee, A. (2019). Survey on Virtual Assistant (Google Assistant, Siri, Cortana, Alexa). 4th International Symposium SIRS 2018, Revised Selected Papers, 65-74.
- [11] Sweigart, A. (2021). PyAutoGUI. PyPI
- [12] Kumar, A., Kumar, A., & Kumar, P. (2021). Development of Voice Assistant using Python. International Journal of Research in Engineering and Science, 10(2), 43-50.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details