AI Based Virtual Assistance

Anjali Baliyan
Department of Electronics and Communication Engineering
Maharaja Surajmal Institute of Technology
New Delhi, India
anjalibaliyan@msit.in

Abstract

This paper explores the development and implementation of a voice assistant A common and simple method of human-computer interaction, voice assistant development and use are examined in this research. The project uses speech recognition and natural language processing techniques to enable voice commands and answers, with a focus on design and usefulness. The underlying technologies are explored, tackling issues like noise reduction, language understanding, and context awareness. These technologies include machine learning techniques and Python libraries. Data collection, model training, and integration with outside services are all included in the implementation phase. To evaluate performance and user satisfaction, evaluation metrics and user feedback are analysed. The research advances our knowledge of voice assistant technology and their uses in areas including smart homes, automotive systems, and customer service.

Keywords— natural language processing, Speech to text, artificial intelligence, virtual assistant

I. INTRODUCTION

Frameworks are being learned step by step nowadays and will assist people in their daily lives [1]. In the IoT world, intelligent virtual assistant (IVA) is a popular service to interact with users based on voice command [2]. These intelligent systems enable users to carry out a variety of tasks, such as organizing schedules, getting information, and operating smart gadgets. Although VAs are especially successful in home automation, they also show great potential as artificial intelligence-driven laboratory assistants [3].

The rapid development of artificial intelligence and machine learning has promoted the development of speech recognition technology [4]. A voice assistant is a digital assistant that helps people through gadgets and voice recognition software by using speech synthesis, natural language processing, and voice recognition. The foundation of this research is speech recognition, one of the fundamental ideas in artificial intelligence. The software functions similarly to Siri and Google Assistant. However, the primary focus of the application is the computer.

I.I. Features of Virtual Assistant

1) *Tasks:* Our project's primary concept is that users ask voice assistants to perform tasks using the device's microphone, and the commands are then converted into text. After that, the text request is handled, and any voice assistant work is delivered together with a text response to the text. We are striving to incorporate the concept of facial identification for security considerations in our voice assistant in addition to basic daily tasks to give it more adaptability and personality. Because our programme doesn't directly interface with servers, it consumes the fewest system resources, reducing the requirement for expensive systems and the risk to your system.

- **To enable an extremely captivating user experience:** Voice help retains consumers' attention better than any other interface.
- Users can communicate naturally with the programmes and ask for everything they like.

- **To stop users from becoming annoyed by the application:** The existing machine system requires us to touch, type, and utilise a mouse in order to finish our task, which can occasionally irritate users. Using a voice assistant, users can directly ask for the task they want to perform.
- Voice assistants can actually respond to every user, taking into account their location, preferred language, and preferences. This enables you to personalise each user's experience with your software.

2) ***Virtual assistant applications****:* Voice assistants can do the following standard tasks:

- Web search
- Play music or videos
- Set reminders and alarms
- Launch any programme or application
- Receive weather updates or send emails, WhatsApp, etc.
- Showing date and time
- Chat using artificial intelligence chatbot

These are only a few examples of the tasks that voice assistants are capable of performing; depending on our needs, there are many more. The capabilities and developments of voice assistants are always expanding to provide users with better performance. Our voice assistant for the desktop runs swiftly and effectively since it is developed with Python modules and libraries.

I.II. Benefits of Voice Assistant

There are many benefits to using voice assistants, including:

- **Increased convenience**: For persons who are busy or have restricted mobility, voice assistants can be used to manage gadgets without having to touch them.
- **Increased safety**: Voice assistants can be used to operate equipment in risky or hazardous settings, such factories or labs.
- **Increased productivity:** By automating processes, voice assistants can help people do jobs more rapidly.

I.III. Issues Of Virtual Assistant

The Despite the many advantages of voice assistants, there are several issues that need to be resolved. A few of these issues include:

- **Accuracy:** Because speech recognition technology is still in its infancy, there may be transcription errors in uttered words.
- **Naturalness:** Voice assistants may have an artificial or robotic quality, which can make users find them less appealing.
- Privacy: Since voice assistants can record a lot of information about what users say, there are worries regarding the privacy of users' data.

I.IV. Internet Applications

The implementation of virtual assistants, the automation of homes, smart cities, management of remote sensors, among other technologies related to the internet of things (IoT) [5]. t. It has some new highlights like posting remarks on the web-based media sites like Facebook, Twitter, and so on By only not many basic orders. You can likewise know the climate around you and can get the environment conditions in your district. It can open and dispatch web-applications and the nearby stockpiling of the client PC [6]..

All about distributed voice-enabled material known as voice sites that can be viewed from any telephone using a standard voice browser is the foundation of the upcoming Voice Web (or voice portals). Internet material may now be accessed from any phone, anywhere, using human voice thanks to the confluence of the richness of the internet with the accessibility and mobility of the phone, creating a large new network called the Voice Web.

II.TECHNOLOGY USED IN PERSONAL DESKTOP VOICE ASSISTANT

The Users of the desktop voice assistant can provide voice commands to carry out a variety of tasks. The system should be capable of accurately recognizing voice instructions, responding quickly, and carrying out the specified duties effectively.

- **Python:** A common programming language for creating personal desktop voice assistants is python. For implementing speech recognition, natural language processing, and machine learning, it provides a number of libraries and frameworks.
- **Speech recognition APIs:** Popular speech recognition APIs include Microsoft Azure Speech Services, Amazon Transcribe, and Google Cloud Speech-to-Text API. These APIs can be incorporated into voice assistant applications and offer speech-to-text conversion capabilities.
- **Natural Language Processing (NLP) libraries:** Python users can use the Natural Language Toolkit (NLTK), spacy, and Stanford Core NLP among other NLP packages. These libraries offer assistance with tasks like part-of-speech tagging, named entity recognition, and sentiment analysis. Engines for Text to Speech (TTS) Google Text-to-Speech, Amazon Polly, and Microsoft are a few well-known TTS engines.
- **Speech Services**: These engines can be used to produce speech output that sounds human from text input.
- **Machine learning frameworks:** The well-known machine learning frameworks TensorFlow, PyTorch, and Scikit-learn are all listed. To train machine learning models for tasks like speech recognition and natural language processing, use these frameworks.
- **Graphical User Interface (GUI) libraries:** The voice assistant application can have a graphic user interface made using GUI tools like PyQt and Tkinter. Information like weather updates, news articles, and reminders can be shown using the GUI.
- **Web APIs:** The voice assistant can be integrated with third-party services using web APIs like OpenWeatherMap, News API, and Spotify Web API. The voice assistant can access weather predictions, news stories, and music streaming thanks to these APIs.

II.I. FUNCTIONAL REQUIREMENTS

THE FOLLOWING FEATURES SHOULD BE AVAILABLE ON THE DESKTOP VOICE ASSISTANT:

• **Wake word detection:** A wake word like "Hey, assistant" should be able to be detected by the system to turn on the personal assistant.

• **Voice recognition:** The system should be able to recognize and understand user voice commands with accuracy.

• **Natural language processing:** The system need to be able to discern the user's purpose and react appropriately.

II.II. NON-FUNCTIONAL REQUIREMENTS

The voice assistant on your computer should have the following qualities:

- **Accuracy**: The system should recognize voice commands with a high degree of accuracy.
- **Speed of response**: The system must react quickly to user commands.
- **Security**: The system must maintain user privacy and be secure.
- **Usability**: The system ought to be simple to operate and have an intuitive user interface.
- **Limitations**: Windows, Mac, and Linux operating systems should all be compatible with the desktop voice assistant. Additionally, a variety of microphones and audio input devices should work with the system.
- Assumptions and Dependencies: For some operations, like sending emails and conducting web searches, the desktop voice assistant assumes that the user has a dependable internet connection.

II.III. SYSTEM REQUIREMENTS

1. Software Requirements:

• Windows OS.

2. Hardware Requirements:

• Minimum Requirement – 2 Gb RAM, Microphones.

• Recommended – 4 Gb RAM, Microphones.

Other Requirements:

- Internet Connection.

III.METHODOLOGY

NLP is used by virtual assistants to translate user text or voice input into actionable commands. The natural language audio signal is transformed into an executable command or digital data that the software can analyse when a user asks their personal assistant to carry out a task. To determine an acceptable response, this data is then compared with software data.You can operate machines using your own commands by using VirtualAssistant. For creating virtual assistants, we use several installer packages for Python, such as

1) **Speech recognition**

The device converts speech input to text using Google's online speech recognition algorithm. Through this, users can talk into a microphone to record their words, which are then briefly saved in the system before being transferred to Google Cloud for speech recognition. Special corpora are organized on the computer network server at the information center. The voice assistant programme then receives the same text and sends it to it.

2) **Python Backend**

Python Backend was used to write the entire programme. The speech recognition module's voice input is converted to output by the Python backend, which then determines whether the given instruction is a Context Extraction, API Call, or System Call. The response is then forwarded to produce the desired result..

3) **API calls**

Application Programming Interface is a type of software intermediary that facilitates communication between two applications. In other words, API is the message conduit that transmits users' requests to providers, who subsequently transmit users' responses.

4) **Content Extraction**

Machine-readable documents that are unstructured or semi-structured can automatically extract structured information via context extraction. NLP is being used in this endeavour to process texts written in human language. The term "content extraction" refers to processes like automatic annotation and content extraction from various photos, videos, and audio files.

5) **System Calls**

In system calls, a programmatic approach is used to ask the operating system's kernel for a service. These requests can be made for hardware services like accessing a hard drive, starting and running new processes, or interacting with process scheduling.

It also offers a crucial point of contact between the process and the OS.

6) **Google-Text-to-Speech**

The main purpose of text-to-speech technology is to turn user-provided text into speech. To put it another way, a TTS engine transforms written text into phonetic representation, which is subsequently transformed into waveforms, producing sound. The third-party publishers that use TTS have helped it evolve and add other languages.

The "component heading" is an additional heading style that is utilised for other elements not included in the main text.

IV.SYSTEM ARCHITECTURE

The following components make up a voice assistant's system architecture:

- **Speech recognition**: The user's spoken words are translated into text by the speech recognition component.
- **Natural language processing:** This element is in charge of deciphering the text's meaning.
-

- **Dialogue management:** This part controls how the user and the voice assistant communicate with one another.
- **Action execution:** This part is in charge of carrying out user commands.
- **User interface:** This part is in charge of inputting user input and showing information to the user..

We have used some of the packages as follows:-

- The os module, which gives access to operating system capabilities, is imported using the import os command. This module is used to perform system commands, get the current working directory, and communicate with the file system.
- importing speech recognition as sr imports the speech recognition module's sr alias. This module offers a straightforward speech recognition API. It can listen for speech instructions and translate them into text.
- The pyttsx3 module is imported using import pyttsx3. The text-to-speech engine offered by this module can be used to read text aloud.
- import pywhatkit imports the module for pywhatkit. Numerous functions for communicating with the web are offered by this module. It can be used to communicate via text message, conduct internet research, and manage smart home appliances.
- Importing the datetime module using import datetime. For representing dates and times, this module offers classes. You can format dates and times, retrieve the time right now, and figure out how much time has passed between two dates.
- wikipedia imports the module for wikipedia. A Python interface to the Wikipedia API is provided by this module. It can be used to conduct a Wikipedia information search and retrieve an article's summary.
- import imports the module for Pyjokes. There are several jokes in this module that can be spoken aloud.
- import json imports the module for json data. This module offers a method for presenting data in a way that is both machine- and human-readable. JSON files can be read and written with it, and Python objects can be converted to and from JSON strings..
- Importing the cv2 module using import cv2. A Python interface to the OpenCV library is provided by this module. A computer vision library called OpenCV can be used to carry out a number of tasks, including object identification, facial recognition, and picture processing.
- import webbrowser imports the module for the web browser. Opening web pages in a web browser is made possible by this module.

V.FUTURE SCOPE

The field of voice assistance is expanding quickly, and new innovations are being produced all the time. Here are a few trends that are anticipated to influence voice help in the future:

- **Improved accuracy and dependability:** Voice assistants will become more accurate and dependable as speech recognition technology advances. As a result, they will be more advantageous for carrying out a variety of functions, including placing calls, sending texts, and managing smart home gadgets.
- **More natural interactions**: Speaking with voice assistants is becoming more natural. They are becoming better at comprehending more intricate instructions and giving human-like responses. Users will find them more desirable as a result, which will motivate them to utilise voice assistants for additional activities.
- **Greater device integration:** Voice assistants are becoming more integrated with other devices, like smartphones, smart speakers, and automobiles. As a

result, voice assistants will be more advantageous in a wider variety of situations and users will find it simpler to control their gadgets with their voice**.**

- **New uses:** Voice assistants are being used in a wide range of new situations, including customer service, healthcare, and education. We may anticipate seeing even more cutting-edge uses for voice assistants as technology progresses**.**

Voice help appears to have a promising future overall. New applications are always being created, and technology is evolving quickly. The accuracy, dependability, and ease of use of voice assistants are all improving. Because of this, voice assistants are likely to grow in popularity over the next years.

Here are some concrete instances of novel and creative ways that voice aid is being applied:

- **Customer service**: Voice assistants are utilised in a range of businesses to provide customer support. Customers can make reservations, schedule appointments, and troubleshoot issues with voice assistants, for instance.
- **Healthcare:** Voice assistants are utilised in a variety of ways to deliver healthcare services. Voice assistants, for instance, can be used to connect patients with healthcare professionals, remind patients to take their medications, and give educational resources.
- **Education:** Voice assistants are utilised in a variety of ways to deliver educational services. Voice assistants, for instance, can be used to access educational resources, offer tutoring, and assist students in learning new ideas.These are just a few examples of how voice assistance is being used in new and innovative ways. As the technology continues to develop, we can expect to see even more innovative applications for voice assistants in the years to come.

VI.CONCLUSION

Voice Assistant helps the users with hand free voice control of their system. Speech recognition is the technology which provides a new way of human interaction with machines. It is very much helpful to the physically challenged people. ARA helps the visually impaired to have access to the most important features of the system enhancing the quality of the system by making use of the different custom layouts and using text to speech. It not only responds to human orders, but also to questions that are posed or phrases that are stated by the user, such as opening tasks and operations. It speaks to the user in a way that makes speaking with the voice assistant more fun and liberating for the user. Instead of text input, the entire system relies on speech input.

REFERENCES

[1] Lakshmi Chandana, C., Ashita, V., Neha, G., Sravan Kumar, K., Suresh Babu, D., Krishna Kishore, G., & Vijaya Bharathi, Y. (2022). Voice-Enabled Virtual Assistant. In *Sustainable Communication Networks and Application: Proceedings of ICSCN 021* (pp. 335-346). Singapore: Springer Nature Singapore.

[2] Chung, H., & Lee, S. (2018). Intelligent virtual assistant knows your life. *arXiv preprint arXiv:1803.00466*.

[3] Austerjost, J., Porr, M., Riedel, N., Geier, D., Becker, T., Scheper, T., ... & Beutel, S. (2018). Introducing a virtual assistant to the lab: A voice user interface for the intuitive control of laboratory instruments. *SLAS TECHNOLOGY: Translating Life Sciences Innovation*, *23*(5), 476-482.

[4] Patil, M. A. A., Gavali, R. A., & Shetty, S. S. (2021). Voice Assistant-A Review. *Science*, *26*, 275-284.

[5] Chilcañán, D., Navas, P., & Escobar, M. (2018, July). Virtual Assistant for IoT process management, using a middleware. In *Proceedings of the 2nd International Conference on Algorithms, Computing and Systems* (pp. 209-213).

[6] Jakhar, V. Virtual Assistant with Applications of IOT.

[7] Khobragade S. (2013)'Jarvis, Digital Life Assistant', International Journal Of Engineering Research & Technology Vol-02,Issue-01,ISSN 2278-0181

[8] Iyer V.,Shah T.K. ,Sheth S.,Kailas D.,(2020)'Virtual assistant for the visually impaired',5th International Conference on Communication and Electronics Systems (ICCES)

[9] Shende D.,Umahiya R.,Raghorte M.,Bhisikar A.,Bhange A.,(2019),'AI Based Voice AssistaUsing Python',JETIR February 2019, Volume 6, Issue 2 www.jetir.org (ISSN-2349-5162)

[10] Dhanraj K.V.,Kriplani L. and Mahajan S., 'Research Paper on Desktop Voice Assistant'(2022),International Journal of Research in Engineering and Science (IJRES),Vol 10,Issue 2.

[11] Damarla K.,'Virtual Assistant Using Python', Retrieved from https://extrudesign.com/virtual-assistant-using-python/