DESKTOP VOICE ASSISTANT USING ARTIFICIAL INTELLIGENT

Prof. C. D. Sawarkar*1, Miss. Akanksha S. Tripathi*2

*1Professor, Computer Science And Engineering, Shri Shankarprasad Agnihotri College Of Engineering Ramnagar, Wardha, Maharashtra, India.
*2Students, Computer Science And Engineering, Shri Shankarprasad Agnihotri College Of Engineering Ramnagar, Wardha, Maharashtra, India.

DOI: https://www.doi.org/10.56726/IRJMETS48988

ABSTRACT

Integrating artificial intelligence (AI) in desktop voice assistants has revolutionized how we interact with computers. These advanced AI-powered voice assistants, such as Jarvis AI, can understand natural language commands, process information, and provide personalized assistance to users seamless and intuitive manner. The voice assistant uses speech recognition modules which is useful for recognizing and understanding human input voice and on the basis of user input command it gives the required input queries or performs the given task like opening and closing different applications, can search and send messages on whatsapp without using keyboard or mouse. In this paper artificial intelligence technology is used to create an desktop voice assistant which will be helpful for visually impaired and the people with disabilities. This desktop voice assistant can also be useful for normal people as it saves time and provide efficiency in doing our day to day tasks. The assistant incorporates features such as voice recognition, natural language processing, and integration with external APIs to enhance its functionality and user experience. The assistant differentiates itself from existing solutions by offering a highly customizable and extensible platform. Users can tailor the assistant’s behavior and functionality to their specific needs, while also benefiting from integration with popular tools and services. The user interface is designed to be intuitive and user-friendly, providing a seamless experience for both novice and experienced users. By creating a personal desktop assistant that combines convenience, automation, and personalized features, this project aims to enhance users' productivity and efficiency in their day-to-day computer tasks. As a personal assistant, Jarvis assists the end-user with day-to-day activities like general human conversation, searching queries in google, searching for videos, retrieving images, live weather conditions, word meanings, searching for medicine details, health recommendations based on symptoms and reminding the user about the scheduled events and tasks. The user statements/commands are analysed with the help of machine learning to give an optimal solution. A speech synthesizer takes as input and produces an audio stream as output. A speech recognizer on the other hand does opposite. It takes an audio stream as input and thus turns it into text transcription. The voice is a signal of infinite information. A direct analysis and synthesizing the complex voice signal is due to too much information contained in the signal. Therefore the digital signal processes such as Feature Extraction and Feature Matching are introduced to represent the voice signal.

I. INTRODUCTION

What is voice assistant and how it works. Many of us might have already known about this voice assistant and we use this in our day-to-day life. A voice assistant is a digital assistant that uses voice recognition, language processing algorithms, and voice synthesis to listen to specific voice commands and return relevant information or perform specific functions as requested by the user. A brief description is given about them in this chapter. Speech is an effective and natural way for people to interact with applications, complementing or even replacing the use of mice, keyboards, controllers, and gestures. A handsfree, yet accurate way to communicate with applications, speech lets people be productive and stay informed in a variety of situations where other interfaces will not. Speech recognition is a topic that is very useful in many applications and environments in our daily life. Generally speech recognizer is a machine which understands humans and their spoken word in some way and can act thereafter. A different aspect of speech recognition is to facilitate for people with functional disability or other kinds of handicap. To make their daily chores easier, voice control could be helpful. With their voice they could operate the light switch turn off/on or operate some other domestic appliances. This leads to the discussion about intelligent homes where these operations can be made
available for the common man as well as for handicapped. This application acts as your personal assistant such as google or Siri, which features in making notes, searching information on Wikipedia, opening youtube, google, stackoverflow, it can play music from your local directory, it can tell you the current date or time, it also has the feature to send emails. It uses the google calendar api by which you will stay updated by your current events also. A brief description is given about them in this chapter.

These personal assistants can be easily configured to perform many of your regular tasks by simply giving voice commands. The most famous application of iPhone is “SIRI” which helps the end user to communicate end user mobile with voice and it also responds to the voice commands of the user. Same kind of application is also developed by the Google that is “Google Voice Search” which is used for in Android Phones. But this Application mostly works with Internet Connections. But our proposed System has capability to work with and without Internet Connectivity. It is named as Personal Assistant with Voice Recognition Intelligence, which takes the user input in form of voice or text and process it and returns the output in various forms like action to be performed or the search result is dictated to the end user. In addition, this proposed system can change the way of interactions between end user and the mobile devices. The system is being designed in such a way that all the services provided by the mobile devices are accessible by the end user on the user’s voice commands.

II. LITERATURE SURVEY

Chen, Liu, and Guo (2020) [1] conducted a comprehensive survey on voice assistant systems. They explored the evolution, technologies, and challenges in the field, covering topics such as speech recognition, natural language understanding, dialogue management, and user experience. The survey provides valuable insights into the current state and future directions of voice assistants. Kundu and Ekbal (2021) [2] presented their work on the design and development of a voice-based intelligent personal assistant. The paper focuses on leveraging natural language processing and machine learning techniques to build an efficient and user-friendly voice assistant system. The study provides insights into the implementation and performance evaluation of the developed assistant, demonstrating its potential in various real-world applications. Sarma, Sengupta, and Gupta (2019) [3] conducted a review of voice assistant systems. They explored various aspects including the architecture, features, applications, and challenges associated with voice assistant systems. The study provides a comprehensive overview of the existing research and advancements in the field, serving as a valuable resource for researchers and practitioners working in this domain. Huang and Sharma (2021) [4] conducted an overview of technologies and applications related to personal voice assistants. They explored the underlying technologies such as automatic speech recognition, natural language processing, and dialogue management, and discussed the wide range of applications where voice assistants are being utilized. The survey provides valuable insights into the current state and potential future developments in the field of personal voice assistants. Wang, Li, Yu, and Zhang (2021) [5] presented their work on the design and implementation of a voice assistant based on deep learning. They focused on leveraging deep learning techniques for speech recognition and natural language understanding, and discussed the system’s architecture and functionalities. The paper provides insights into the development and deployment of a voice assistant system powered by deep learning, showcasing its potential in improving user interaction and convenience. Bullich and Grivolla (2020) [6] addressed the issues and opportunities surrounding voice assistants and language technologies. They discussed challenges such as speech recognition accuracy, multilingual support, privacy concerns, and bias in voice assistant systems. The paper also highlighted opportunities for improving user experience, expanding language coverage, and leveraging voice assistants for language-related tasks. The study offers valuable insights into the current landscape and future directions of voice assistants and language technologies.
III. METHODOLOGY

In this work, we are implementing a desktop virtual assistant, where the speech recognition library has many build functions that will help the virtual assistant to understand the command given by the user and it will respond to the user invoices.

The NLP algorithm will convert the user voice into text and according to the keywords present in the text respective action will be performed by the desktop virtual assistant. Also, our assistant will be able to perform some other functions like searching history, mathematics, and science questions also extracting news with the help of API Wolfram Alpha. We are using libraries like random, OS, Turtle, My alarm etc. each corresponding to these techniques. We will use the libraries random for gaming, OS to implement operating system function like battery percentage, sleeping system, volume up, volume down. By using pyautogui will be able to take the screenshot. Distutil library is useful for checking battery status. Some other modules are:

1. Speech recognition — Speech recognition is an important feature used in various forms like house automation and artificial intelligence devices. The important function of this library is to try to understand whatever the humans speak and converts the speech to text form.
2. pyttsx3 — pyttsx3 is a conversion library OF text to speech in python. This module supports text to speech engines on mac, Linux, and windows.
3. Wikipedia — Wikipedia is an online encyclopedia used by many people from the academic community ranging, freshmen to professors to students who want to gain knowledge over a particular topic. This module in python extracts data required from Wikipedia.
4. CV2 — This module is used to capture images from your camera
5. DateTime — This is a default module in python and it works on time and date.
6. OS — It provides the function to interact with the operating system and this module is a standard library in python
7. Time — This module helps us to showtime.
8. Web browser — It extracts data from the internet and this is an in-built package in python.
9. Subprocess — This is a standard library used to process various system commands like to sleep or to restart and shut down your PC.
10. Json- The json module is used for storing and exchanging data from API (Application Programming Interface).
11. Request- The request module is used to send all types of an HTTP requests. Its accepts URL as a limitation and gives access to the given URL’S.
12. WolframAlpha - WolframAlpha is an API that can compute expert-level answers using Wolfram's algorithms, information base, and AI technology.

IV. MODELING AND ANALYSIS

A voice assistant system typically consists of several modules working together to provide a seamless user experience. Here are the key modules commonly found in voice assistant systems. A voice assistant is a software program that uses speech recognition and natural language processing to understand and respond to user commands or queries. Fig.1 Block Diagram It leverages artificial intelligence and machine learning techniques to provide information, perform tasks, and interact with users through voice-based communication. Voice assistants are commonly found on smartphones, smart speakers, and other devices, acting as virtual personal assistants. They can answer questions, provide recommendations, control smart home devices, set reminders, play music, and perform a wide range of tasks based on user requests, enhancing convenience and efficiency in daily activities
Speech Recognition Module:

Speech Recognition Module: This module converts spoken words into text through speech recognition algorithms. It analyzes and processes audio input, identifying words and phrases spoken by the user. This technology enables applications and devices to understand and respond to voice commands, making it possible to interact with them through spoken language. Speech Recognition modules are widely used in virtual assistants, voice-controlled systems, transcription services, and various other applications.

Natural Language Processing (NLP) Module:

Tokenization: Tokenization is essentially splitting a phrase, paragraph, or an entire text document into tokens. Each of these smaller units is known as a token. For example, Sara eats cake. So the sentence is broken down as 'Sara' 'eats' 'cake'. Stemming: Stemming is a technique that is used to extract the base form of the words by removing affixes. For example, the root of the words eating, eats, eaten is eat. Lemmatization: In Lemmatization root word is called Lemma. For example, eats, eating, eat are all forms of the word eat, therefore eat is the
lemma of all these words. POS: PoS tagger is to resolve the ambiguity accurately based on the context of use. For example, the word "eat" can be a noun or a verb. Name entity recognition: Named entity recognition (NER) helps you easily identify the key elements in a text like names of people, places, brands, etc. For example, Sara is a name. Chunking: Chunking is a process to take small pieces of information and group them into large units.

V. RESULTS AND DISCUSSION

We have used the VS Code IDE in this video. Feel free to use any other IDE we are comfortable with. I have started a new project and make a file called jarvis.py. Visual Studio Code is a freeware source-code editor made by Microsoft for Windows, Linux and macOS. Features include support for debugging, syntax highlighting, intelligent code completion, snippets, code refactoring, and embedded Git. How to install Visual Studio Code on Windows? We can download Visual Studio code from URL "https://code.visualstudio.com/download" by selecting the right platform: Firstly, download the Visual Studio Code installer for Windows. Once it is downloaded, run the installer (VSCodeUserSetup-{version}.exe). It will only take a minute.
VI. CONCLUSION

The development of a voice assistant project brings forth a range of benefits and opportunities for users in various domains. Voice assistants have witnessed significant advancements in speech recognition, natural language processing, and personalized assistance, leading to improved user experiences. Through integration with devices, applications, and IoT systems, voice assistants have become valuable tools for controlling smart homes, accessing information, and performing tasks hands-free. However, there are still challenges to overcome in the realm of voice assistants. Issues such as accuracy, contextual understanding, personalization, privacy, and bias need to be addressed to enhance the overall user experience. Ongoing research and development efforts are essential to ensure voice assistants continue to evolve and meet user expectations.

In this paper, we have discussed a VA developed using python. This assistant currently works online and performs basic tasks like weather updates, streaming music, searching Wikipedia, playing music, opening desktop applications, etc. The system requires an internet connection. This Personal Assistant has been designed with ease of use as the main feature. The Assistant works properly to perform tasks given by the user. It will overcome the drawbacks of the existing solutions. A similar application is Cortana which comes with Windows operating system does not work in other operating systems but our system will work on any of the operating systems. There are many benefits of using this assistant such as it reduces labor costs, improves work quality, increases productivity, increases flexibility, etc.

VII. REFERENCES


