



INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Desktop Task Assistant

Mr. Aayush Jain
Student, M.Sc. (IT)
K.C College, Churchgate

Dr. Rakhi O Gupta
Coordinator Department
K.C College, Churchgate

Mrs. Nashrah Gowalker
Assistant Professor
K.C College, Churchgate

Abstract: In today's digital landscape, personal or virtual assistants have become essential tools, significantly enhancing the productivity and efficiency of both individuals and businesses. This project focuses on the development of an advanced desktop-based virtual assistant designed to automate a variety of tasks by accepting, comprehending, analyzing, and executing user commands. By streamlining routine activities, the assistant saves time and increases overall user efficiency. The virtual assistant employs cutting-edge machine learning techniques and natural language processing technologies to interpret voice commands accurately. It draws data from multiple online sources, such as search engines, social media, and other information platforms, to continuously enhance its capabilities. This allows the assistant to adapt to a wide range of user requests, providing dynamic and precise responses. For instance, the assistant can retrieve real-time data, such as weather forecasts or news updates, ensuring that users always have access to the latest information.

The core technologies that power this system include voice activation, automatic speech recognition, text-to-speech synthesis, natural language understanding. These technologies work together to create a smooth and intuitive user experience, where the assistant can execute desktop tasks, provide information, and respond to queries in real-time. Additionally, advanced web scraping techniques enable the assistant to gather up-to-date information from the internet, enhancing its ability to handle complex requests.

As technology evolves, the need for more accessible and efficient ways to interact with computers becomes increasingly important. While traditional methods of input, such as keyboards and mice, are still widely used, voice-based interaction offers a more convenient alternative. This is especially beneficial for individuals with visual impairments who may struggle with traditional input methods. A voice assistant allows them to interact with their computer easily, executing commands and receiving feedback through voice or other accessible means.

The primary goal of this project is to create a versatile and inclusive virtual assistant capable of automating tasks, retrieving real-time information, and improving user productivity. By combining voice recognition, task automation, and dynamic information retrieval, the assistant provides a seamless and efficient experience for all users. In doing so, it not only enhances daily productivity but also ensures greater accessibility for those who rely on voice commands to interact with their devices.

Keywords: Virtual Assistant, Voice Command, Task Automation, Machine Learning, Natural Language Processing, Speech Recognition, Text-to-Speech, Dynamic Information Retrieval, Web Scraping.

I. INTRODUCTION

In our project, we mainly use voice as communication means, so the basically the Speech recognition application. 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. The concept of speech technology really encompasses two technologies: Synthesizer and recognizer. A speech synthesizer takes as input and produces an audio stream as output. A speech recognizer, on the other hand, does the direct opposite. It takes an audio stream as input and thus turns it into text transcription. 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. In this project, we directly use the speech engine which uses Feature extraction technique.

An voice assistant is a software agent that can perform tasks or services for an individual based on commands or questions. Desktop voice assistants are able to interpret human speech and respond via synthesized voices. 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 computer devices are accessible by the end user on the user's voice commands.

II. LITERATURE SURVEY

Personal or virtual assistants have become an integral part of our lives today. All businesses or individuals are switching to these types of technologies to make their jobs easier. This system is based on desktop applications. The system includes a virtual assistant that can accept, understand, analyse, and perform the appropriate tasks from users. This saves a lot of time for the user. Each developer at an intelligent assistant company applies their own methods and approaches to development, which impacts the final product. One assistant can synthesize speech more qualitatively, another assistant can perform tasks more accurately without additional explanation or modification, and the other assistant can perform a narrower range of tasks, but most accurately. It can be done according to the user's wishes. Obviously, there is no one-size-fits-all assistant that does all the tasks equally well. The properties of the assistant are entirely dependent on which areas the developer paid more attention to. All systems are based on machine learning techniques, using large amounts of data from different sources to create and train them, so the sources of this data are search systems, different sources of information, and social. It plays an important role such as network. The amount of information from different sources determines the types of assistants that can emerge from it. Despite different learning approaches, different algorithms and methods, the principles for building such a system are about the same. The most important technologies are voice activation, automatic voice recognition, Text to Speech, voice biometrics, dialogue manager, natural language understanding, and recognition of named entities.

III. OBJECTIVES

- 1) Main objective of building personal assistant is to answer questions that users may have.
- 2) This may be done in personal computer where the virtual assistant is available as a service where a voice asks the user "What can I do for you?" and then responds to verbal input.
- 3) Virtual assistants can tremendously save you time. When we need to search something while working we can just give a voice command and HARRY can do that for you.
- 4) Provide a topic for search and continue with your tasks while HARRY does the searching. Also before you go out you can ask Harry for the weather. While you are busy in some work Harry can send Mail, WhatsApp messages on your voice command.
- 5) One of the main advantages of voice searches is their rapidity. In fact, voice is reputed to be four times faster than a written search: whereas we can write about 40 words per minute, we are capable of speaking around 150 during the same period of time.

IV. METHODOLOGY

1. System Architecture:

The overall system design consists of following phases:

- (a) Data collection in the form of speech.
- (b) Voice analysis and conversion to text
- (c) Execute Python script
- (d) Generating speech from the processed text output

In first phase, the data is collected in the form of speech and stored as an input for the next phase for processing. In second phase, the input voice is continuously processed and converted to text using STT[1]. In next phase the converted text is analyzed and processed using Python Script to identify the response to be taken against the command.

Finally once the response is identified, output is generated from simple text to speech conversion using TTS[2].

2.Data Flow Sequence

- Initialize device: Initialize the device by calling its name.
- Task Manager: Conversion of Speech-to-Text and Text- to Speech is performed by task manager.
- Service Manager: Analysis of commands and matching them with web service and applications.
- Execute Command: After finding the match for the given command, run the respective python script and give the output..

V. PROPOSED SYSTEM

In this proposed concept effective way of implementing a Personal voice assistant, Speech Recognition library has many inbuilt functions, that will let the assistant understand the command given by user and the response will be sent back to user in voice, with Text to Speech functions. When assistant captures the voice command given by user, the under lying algorithms will convert the voice into text. And according to the keywords present in the text (command given by user), respective action will be performed by the assistant. This is made possible with the functions present in different libraries. Also, the assistant was able to achieve all the functionalities with help of some API's. These APIs were used for several functions, such as performing calculations and extracting news from web sources. Send a request and get the appropriate output via the API. APIs like WOLFRAMALPHA are very useful for performing calculations and small web searches. In this way, you can extract messages from a web source and send them as input to a function to further achieve your goals. There are also libraries like Random and many others, each compliant with a different technology. I used the library operating system to implement operating system-related features such as system shutdown or reboot the system.

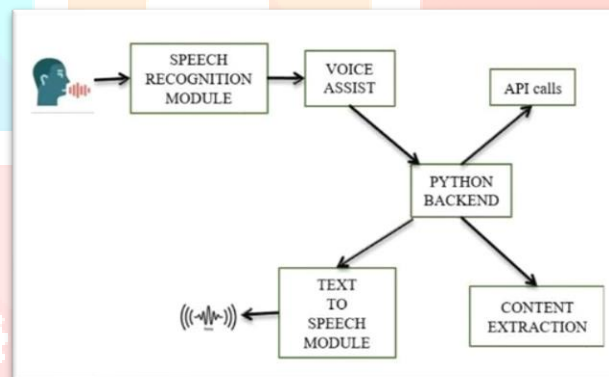


Figure-1:Detailed Workflow

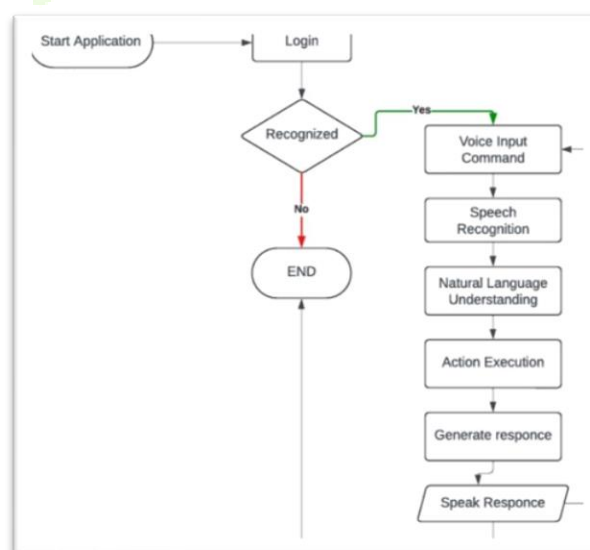


Figure 2:Flow Chart

Figure 2 shows a flowchart where a user starts the voice assistant, it will greet the user and wait for the user to provide an input command. If the user provides an input command through voice, the assistant will capture it and search for any related keywords. If the assistant finds a keyword related to the input command, it will perform the task and return the output to the user both in voice and in text form on the terminal window. If the assistant cannot find any relevant keywords, it will continue waiting for the user to input a valid command. Each of these functionalities plays an important role in the overall functioning of the system.

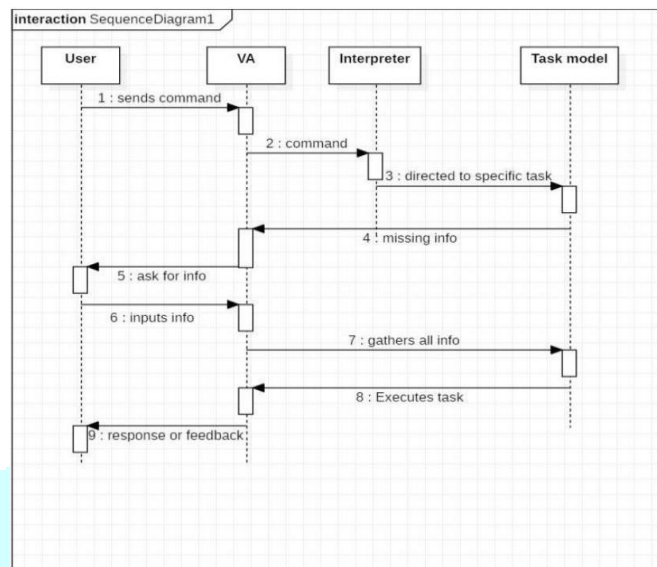


Figure 3: Sequence Diagram

The user sends command to virtual assistant in audio form. The command is passed to the interpreter. It identifies what the user has asked and directs it to task executer. If the task is missing some info, the virtual assistant asks user back about it. The received information is sent back to task, and it is accomplished. After execution feedback is sent back to user.

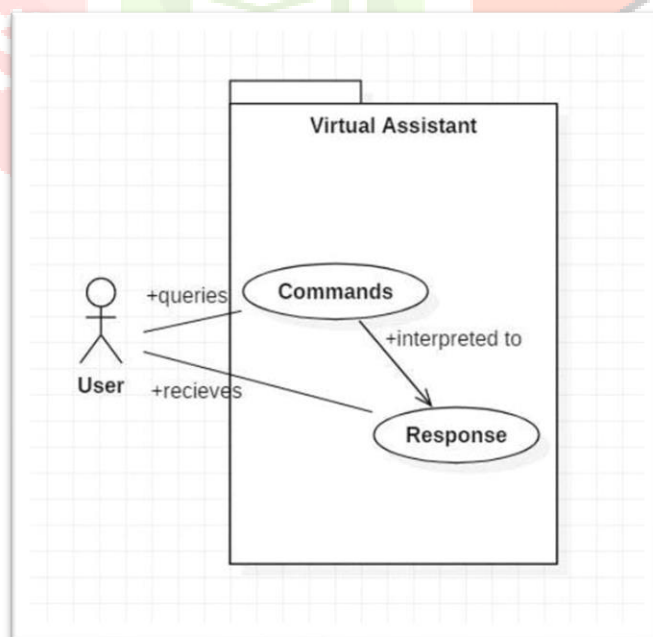


Figure 4: Use Case Diagram

In this project there is only one user. The user queries command to the system. System then interprets it and fetches answer. The response is sent back to the user.

VI. BENEFITS TO THE SOCIETY

1) MINIMAL EFFORT

IT'S EASIER TO SAY A FEW WORDS THAN TYPE THEM ON A SMALL SMARTPHONE SCREEN.

2) EYES FREE

ARE YOU WEARING GLASSES OR LENSES? ME TOO. YOU CAN BE AS BLIND AS A BAT, BUT A VOICE ASSISTANT WILL ALWAYS HELP YOU. YOUR EARS ARE ENOUGH. YOU CAN ALSO ASK THE BOT ABOUT SOMETHING WHILE COOKING AT THE SAME TIME.

3) HANDS-FREE

THE COMMUNICATION WITH A VOICE ASSISTANT CAN TAKE PLACE EVEN WHEN YOUR HANDS ARE OCCUPIED. ARE YOU PAINTING YOUR NAILS NOW? ARE YOU INSTALLING A CABINET NOW? DON'T WORRY, THE VOICE ASSISTANT WILL BE HAPPY TO HEAR WHAT YOU HAVE TO SAY.

4) MINIMALLY DISRUPTIVE

HAVE YOU EVER WONDERED HOW MANY TIMES YOU TURN ON AND OFF YOUR MOBILE WHILE TEXTING VIA MESSENGER? IF YOUR SMARTPHONE IS NOT MUTED, YOU HEAR A NOTIFICATION SOUND EACH TIME. THAT'S IT. COMFORTABLY READ YOUR BOOK AND CHECK WHAT IS HAPPENING IN THE WORLD USING YOUR VOICE ONLY.

5) FAST RESPONSE

IMAGINE HOW MUCH TIME YOU HAVE TO SPEND TO FIND SOME INFORMATION ON A WEBSITE? OR HOW MANY CLICKS DO YOU NEED TO MAKE BEFORE YOU FIND THE THING YOU NEED IN A MOBILE APPLICATION? VOICE ASSISTANTS DON'T GENERATE SUCH DIFFICULTIES. YOU ASK A QUESTION AND YOU HAVE THE ANSWER.

6) EASY FOR CHILDREN

TWO-YEAR-OLD CHILDREN TYPING VIA MESSENGER? IMPOSSIBLE? TWO-YEAR-OLD CHILDREN READING WEBSITES? UNLIKELY. 2-YEAR-OLD CHILDREN SPEAKING WITH A VOICE ASSISTANT? NOTHING HARD.

VII. FUTURE SCOPE

The virtual assistants which are currently available are fast and responsive but we still have to go a long way. The understanding and reliability of the current systems need to be improved a lot. The assistants available nowadays are still not reliable in critical scenarios. The future of these assistants will have the virtual assistants incorporated with Artificial Intelligence which includes Machine Learning, Neural Networks, etc. and IoT. With the incorporation of these technologies, we will be able to achieve new heights. What the virtual assistants can achieve is much beyond what we have achieved till now. Most of us have seen Jarvis, that is a virtual assistant developed by iron man which is although fictional but this has set new standards of what we can achieve using voice- activated virtual assistants

REFERENCES

[1] Dr.C K Gomathy, Article: A Semantic Quality Of Web Service Information Retrieval Techniques Using Bin Rank, International Journal Of Scientific Research In Computer Science Engineering And Information Technology (IJSRCSEIT) Volume 3 | Issue 1 | ISSN : 2456-3307, P.No:1563-1578, February-2018

[2] Dr.C K Gomathy, Article: A Scheme Of ADHOC Communication Using Mobile Device Networks, International Journal Of Emerging Technologies And Innovative Research (JETIR) Volume 5 | Issue 11 | ISSN : 2349-5162, P.No:320- 326, Nov-2018.

[3] Dr.C K Gomathy, Article: Supply Chain-Impact Of Importance And Technology In Software Release Management, International Journal Of Scientific Research In Computer Science Engineering And Information Technology (IJSRCSEIT) Volume 3 | Issue 6 | ISSN : 2456-3307, P.No:1- 4, July-2018

[4] Dr.C K Gomathy, Article: The Efficient Automatic Water Control Level Management Using Ultrasonic Sensor, International Journal Of Computer Applications (0975 – 8887) Volume 176 – No. 39, July 2020.

[5] C K Gomathy And V Geetha. Article: A Real Time Analysis Of Service Based Using Mobile Phone Controlled Vehicle Using DTMF For Accident Prevention. International Journal Of Computer Applications 138(2):11-13, March 2016. Published By Foundation Of Computer Science (FCS), NY, USA,ISSN No: 0975-8887

[6] C K Gomathy And V Geetha. Article: A Real Time Analysis Of Service Based Using Mobile Phone Controlled Vehicle Using DTMF For Accident Prevention. International Journal Of Computer Applications 138(2):11-13, March 2016. Published By Foundation Of Computer Science (FCS), NY, USA,ISSN No: 0975-8887

[7] G. Bohouta, V. Z. Kėpuska, "Comparing Speech Recognition Systems (Microsoft API Google API And CMU Sphinx)", Int. Journal of Engineering Research and Application 2017, 2017.

[8] Subhas mani Kaushal,megha Mishra(IRJET)volume:07 Issue:09|sep 2020

[9] Dhiraj Pratap singh,Deepika Sherawat,Sonia school of computer science and engineering,Galgotias University volume 26,Issue 6,2020.

[10] Kei Hashimoto, Junichi Yamagishi, William Byrne, Simon King, Keiichi Tokuda, "An analysis of machine translation and speech synthesis in speech-to-speech translation system" proceedings of 5108978-1-4577-0539- 7/11/\$26.00 ©2011 IEEE.

