

NLP based AI Voice Assistant

Prof. Sonali Pakhmode¹, Vishranth Poojary², Pratik Bhore³, Kushal Thakur⁴, Vishal Dethe⁵

Department of Information Technology

Vasantdada Patil Pratishthan's College of Engineering & Visual Arts, Sion, Mumbai, India

ABSTRACT - All electronic devices will operate in the very near future using a remote helper, which is easily accessible but requires a certain level of weakness. This setup allowed users to access the system via voice commands. The user can ask the assistant to do anything that the framework is capable of doing, such as play music, open a certain application, open tabs, open websites, and so forth. Programming experts known as voice associates are able to understand human speech and respond with carefully crafted voices. Our voice assistant's primary goal is to educate individuals and provide quick, calculated results. The voice help receives voice input from our microphones (wired and Bluetooth) and turns it into computer-understandable English before providing the user with the necessary solutions and replies. This assistant makes a connection to the Internet to deliver answers to the user's queries. A natural language processing algorithm enables computer systems to communicate using many forms of natural human language.

Keywords: NLP (Natural language processing), TTS (Text-to-Speech), AI (Artificial Intelligence), Voice Assistant, NLTK(Natural Language Toolkit), Python.

1. INTRODUCTION

The most well-known voice assistants include Siri (Apple Inc.), Alexa (Amazon), Cortana (Microsoft), and so forth. Voice assistants are defined as software agents that can comprehend or modify spoken language and respond with synchronized voices.

Although we can see that the futuristic idea of talking to the computer is now a reality, technology is advancing day by day. Since people have sought to communicate with computers since they were first created, this is now possible thanks to the technology developed by determined computer scientists. Several products deployed in the last few years which bring out the inexpensive use of the voice assistants to our day-to-day life, with regard to the time more features and a platform are being added and gets updated from time to time.

In this fast-developing modern world, everyone likes to

get their work done in easier and more fast manner. For that, we will be creating an AI Voice Assistant that will assist you to get your job done. From opening your files to search things on Google, it would assist you for every single thing. Moreover, you would have your personal assistant which would make you feel like a boss. It would use voice commands from the user to receive and interpret directives. All your tasks would be completed easily at user's voice command and with this user can interact and respond with the machine in natural language. The purpose of this project is to create and develop an AI Voice Assistant using NLP which is helpful for the people especially visually impaired and handicapped providing a hands-free tool that eases your daily tasks.

1.1 Project Scope

How rapidly the time changed? If we look back around twenty year ago, Voice Recognition was in its infant stage. As when the computer system came into the existence, it was the dream to fully fledged interaction with the computer machine. Now, we can eventually talk, ask and as well order to do the assigned task. The world is moving forward as a result of this technological improvement. The already accessible virtual assistants are quick and helpful, but there is still a long way to go. The current systems' comprehension and dependability require significant improvement. The modern aides are nevertheless unreliable in urgent situations. The virtual assistants of the future will be equipped with artificial intelligence, such as machine learning, neural networks, and the internet of things. We shall be able to scale new heights with the help of these technologies. What the virtual assistants are capable of is far beyond what we have so far accomplished.

1.2 Project Goals and Objectives

1.2.1 Project Goals

The primary goal of building an AI Voice Assistant Project for desktop is to provide users with a hand-free, convenient tool that can help them perform various tasks on their computer without having to use their hands. The AI Voice Assistant can use Natural Language Processing (NLP)

algorithms to understand user voice commands, interpret them, and execute actions on the user's behalf.

The AI Voice Assistant can be integrated with various software and applications installed on the user's computer, allowing users to perform tasks such as opening applications, searching the web, opening YouTube, text reader using voice commands. This feature can be particularly beneficial for individuals with mobility impairments, who may find it challenging to use a keyboard or mouse to interact with their computer. Overall, the purpose of building an AI Voice Assistant Project for desktop is to provide users with a tool that can simplify their daily tasks, increase their productivity, and enhance their overall computing experience.

1.1.1 Objectives

- i. To create a system designed to be used efficiently on desktops.
- ii. To develop a software tool which is simple & easily accessible and operable.
- iii. To create an interactive software which is cost effective.
- iv. To ensure the AI Voice Assistant can recognize and respond accurately to a wide range of voice commands, including variations in accents and speech patterns, to enhance user experience and accessibility.
- v. To integrate the AI Voice Assistant with commonly used desktop applications, such as email clients, web browsers, and productivity software, to provide a seamless user experience.
- vi. To conduct user testing and gather feedback from visually impaired and handicapped users to ensure the AI Voice Assistant is accessible, user-friendly, and meets their specific needs. This will help ensure that the AI Voice Assistant is truly effective and helpful for this target group.

2. PROBLEM STATEMENT

Implementation of AI based Voice assistant using NLP techniques. Our aim is to help the visually

impaired or handicapped people by providing them hands free tool and to use technology without using hands.

We all are aware of the Google voice assistant, Siri, Cortana but these Voice assistants can be used only on their respective devices. There is no such voice assistant which can be used across all platforms.

3. LITERATURE SURVEY

Artificial intelligence (AI) systems that can orchestrate natural human-machine interaction (via speech, gestures, facial expressions, etc.), are currently gaining popularity. The direction of interaction based on the machine's interpretation of real human language was one of the most researched and well-liked topics. Now, a machine learns to converse with a human by observing his activities, routines, and behavior in an effort to become his personal assistant rather than the other way around. Personal assistants are increasingly an essential part of our life, whether we realize it or not. It is because of all the features and ease of use they provide. A user can concentrate on what is most important to them by using a personal assistant to automate some of their daily duties. Features like, making calls, writing messages, taking photographs, storing to-dos on the go, browsing internet etc., are offered by personal assistants. So, utilization of these features of a virtual assistant will save an individual a lot of time, and effort. It's critical to concentrate more on what matters most to a person, whether it be personal or professional job. These kinds of personal assistants can automate mundane tasks that people frequently spend more time on. When someone works in an environment with which he/she is not familiar with, they often find it difficult to locate applications that they need, like browser, any IDE or any other software. The majority of the time, people will waste hours looking for the application. This results in unnecessary time wastage. A personal assistant with voice recognition will therefore aid in automating this process. User is expected just to give a voice command, and the assistant will take care of the rest. The use of a voice-activated personal assistant is discussed in this paper. This device can help people do tasks more quickly and easily by responding to voice instructions.

The final product is impacted by the unique methods and approaches that each organization using the intelligent assistant uses to produce it. One assistant can execute duties more accurately and in accordance with the user's

preferences, while another can accomplish jobs more accurately and without the need for extra explanations or corrections. There isn't a single assistant that could do everything perfectly, of course. The set of traits an assistant possesses entirely depends on the region to which the developer has given the most consideration. The source of this data—whether it be search engines, different information sources, or social networks—plays a significant influence. The type of the assistant is determined by the quantity of information from many sources, which may lead to a result. Although there are various learning strategies, algorithms, and procedures, the fundamentals of creating such systems are essentially the same.

The first approach that we looked at involves acoustic analysis, which often completes a partially pronounced feature by recording and analyzing it using learned data. The user's command is broken up into words, checked against the command list, and then a response is sent in accordance with that.[1]

The second evaluation is of a smartphone application with a few features that combines natural language processing to convert audio to text and decision tree algorithm to classify photographs and detect objects in them. It utilizes keywords supplied by the user, which are compared to training data, and the reply is returned.[2]

The models under review included each task or service corresponding to a class of service modules, each of which provided the same service and exposed the same interface but differed in terms of performance, computational complexity, memory footprints, degree of portability (some modules may rely on proprietary services), and potential reliance on external (cloud) services.[3]

The following assistant under examination is efficient with resources, interactive, and adaptable, and the realized prototype works on a reasonably priced, diminutive Raspberry PI 3 device. The system was connected into an open source home automation environment for testing purposes, and it ran for many days with users being urged to engage with it. It turned out to be precise, dependable, and appealing.[4]

These are systems that process two or more combined user input methods, such as speech, image, video, touch, and manual gestures, are known as multi-modal dialogue systems. The original framework of generic dialogue systems included elements such the ASR Model, Gesture Model, Graph Model, Interaction Model, User Model, Input Model, Output Model,

Inference Engine, Cloud Servers, and Knowledge Base for generating responses.[5]

The reviewed DIVA assistant design functions as a plan that considers accent recognition, speech registration, and allows consumers more manual control while acting as a mobile application. In order to achieve the set user goals, the virtual assistant uses techniques including natural language processing, speech pattern recognition, and machine learning.[6]

The next system uses a carefully curated data set. The dataset has then been completely cleaned by removing punctuation, superfluous symbols, and other extraneous items. The redundant Logos were then found and removed from the dataset using the Bag of Words technique after we had fully prepared our data. Despite our best efforts to clean up the incoming data and turn it into beautiful phrases and clean terminology.[7]

NLP techniques are used by the chatbots we studied to recognize user commands and respond. The chatbot's main architecture consists of deep neural networks and APIs for text-to-speech and speech-to-text conversion. The robot receives the voice sign commands that have been converted to text over a Bluetooth network and uses open cv. The chatbot uses stemming, lemmatization, and word filtering in conjunction with the NLTK module to perform natural processing in order to respond to user requests.[8]

The further NLP- based chatbot under study is built in such a way that users can provide audio notes rather than texts, as with other chatbots. This architecture initially converts the user-sent text to text. The chatbot's main architecture consists of deep neural networks and APIs for text-to-speech and speech-to- text conversion.[9]

The most recent system under examination is a mobile application that uses user-provided voice commands that are translated by the mobile device utilizing natural language processing. The mobile device serves as a central console, deciding which appliance must perform which action to satisfy the user's request. The cloud has the ability to process the sent data. The Internet of Things is established when the appliances are linked to the mobile device using an Arduino Board.[10]

4.PROPOSED SYSTEM

Our aim is to build a software which is simple but equally efficient. The software that we are proposing mainly focuses on helping the user go hands free and ease on the daily routine task of the user. It also intends to save time and money of the user. It is purely a desktop-based voice assistant which takes the voice commands from the user

and on the interpretation of that gives a desired output. This Assistant consists of three modules. First step of this is assistant accepting voice input from user or client. Next, determine the user's input's intent and function by analyzing and mapping it. The third step is the assistant speaking the user's outcome throughout. The assistant will initially begin taking user input. The assistant will translate the analogue voice input into digital text after receiving it. The assistant will translate the analogue voice input into digital text after receiving it. The assistant will start asking the user for input again if it was unable to translate the voice into text. If converted, it will begin to analyze the input and map it to a certain function. The assistant would use NLP algorithms to convert user input into machine understandable language and map it with appropriate responses and later, the output will be given to user via the voice command. It uses TTS modules to convert voice commands to text and text responses again to voice commands. The user can change the name of the assistant as per their choice. The proposed assistant comes very handy for the user when it comes to simplifying the day- to-day desktop task of the user. The assistant can search and open any folder the user wishes for within seconds with just a voice instruction and its very helpful especially when a folder is misplaced. The voice assistant opens Google, Wikipedia, YouTube, plays music, play videos in the system just by giving a voice modulated command. If the user wishes to know the temperature or wishes to know/change the date and time then just ask the assistant and it is done. The software can also read out aloud the text of a particular document. When a user requests that the assistant read out some news, the assistant will both read out the news and show it line by line. Also, the text reading speed can be controlled by the user. The voice assistant is very much capable and has a special feature of Speech to Text conversion for documenting any text without the chaos of typing each and every word physically. This feature is very beneficial when it comes to texting and typing emails and send them without using the keyboard at all. Additionally, the user can change the voice of the assistant to male/female whichever they are comfortable with. Whenever the user calls for their assistant, the assistant greets the user with its

username and as per the time then.

4.1 Design

Fig.1 Use-Case Diagram

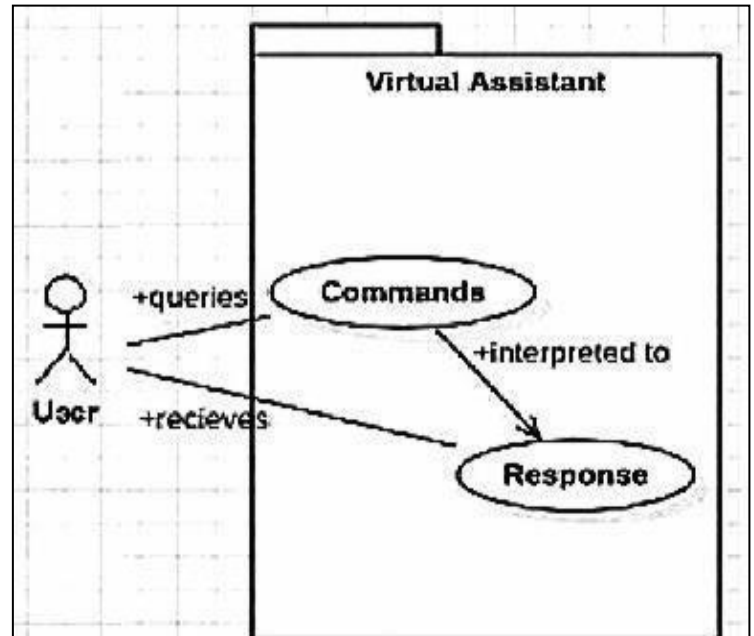


Fig.2 Data Flow Diagram

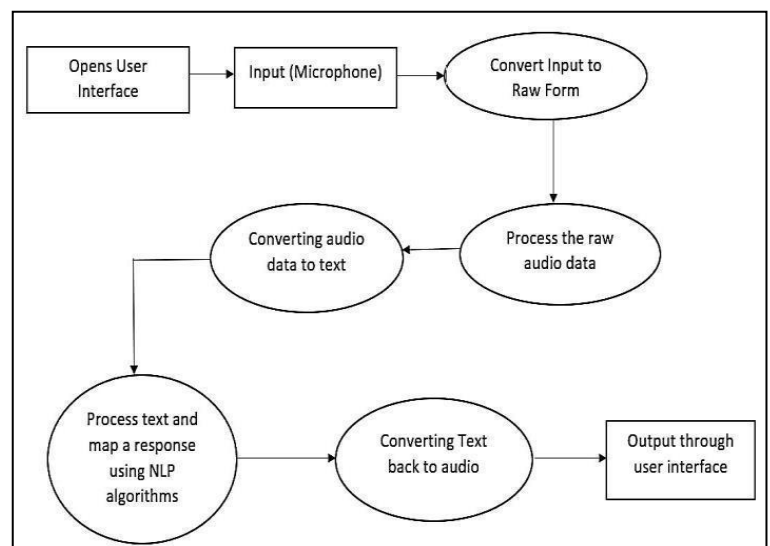


Fig.3 Block Diagram

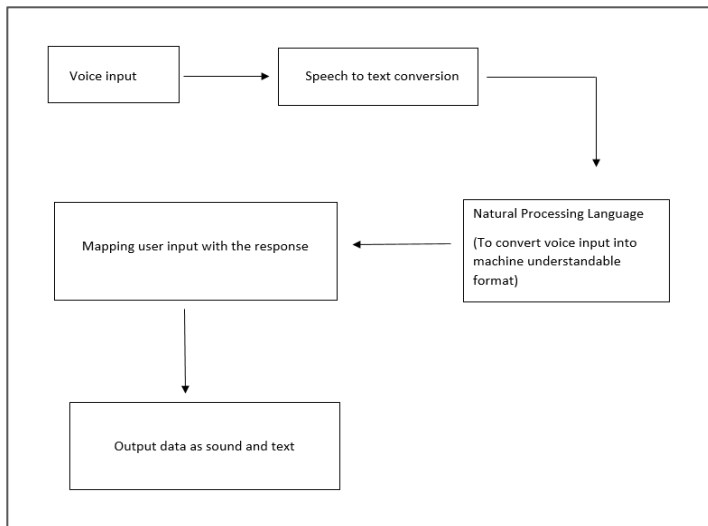
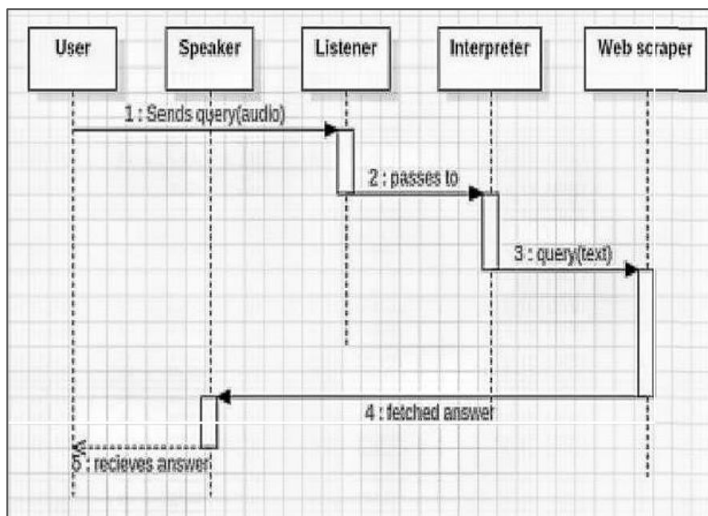


Fig.4 Sequence Diagram



5.RESULTS

The necessary Python programming language packages have been installed, and the code was written using the Visual Studio Code Integrated development environment (IDE). The Python code we created is compatible with Python 3, and the outputs that we have seen so far from our NLP-based AI voice assistant are shown below.

5.1 Open Google

As shown in Fig.5 and Fig.6. When we command the

voice assistant to open Google, it accepts the user request and executes the task by launching the Google search engine browser.

Fig.5 UI screen of Open Google command

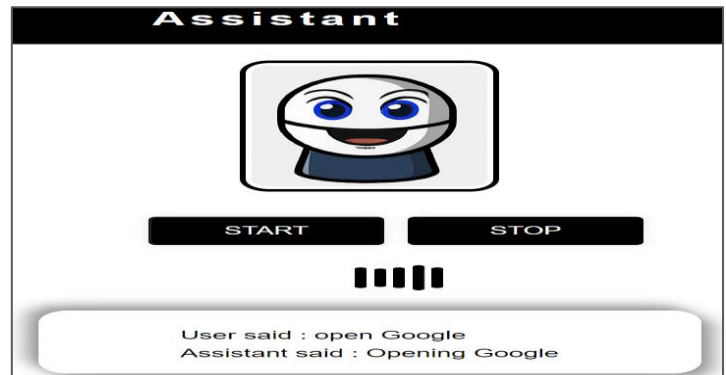
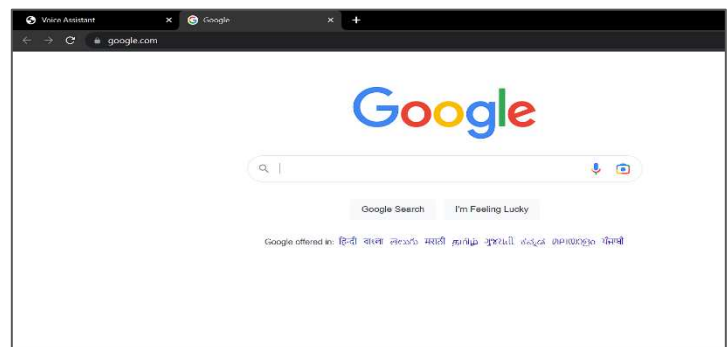


Fig.6 Output screen of Google browser



5.2 Playing videos on YouTube

As shown in Fig.7 and Fig.8. When we command the voice assistant to play a certain video/music, it performs the task by playing the video/music on YouTube.

Fig.7 UI screen of playing video on YouTube



Fig.8 Output screen of playing video on YouTube



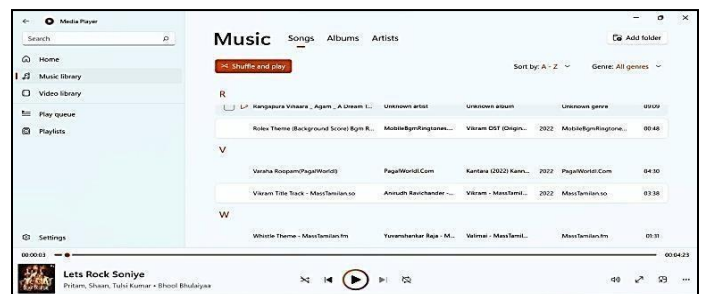
5.4 Playing music on the PC system

As shown in Fig.11 and Fig.12. When we command the voice assistant to play music which is stored on the PC system, it takes the user input and executes the task by playing music using system music player.

Fig.11 UI screen of play music command



Fig.12 Output screen of playing music



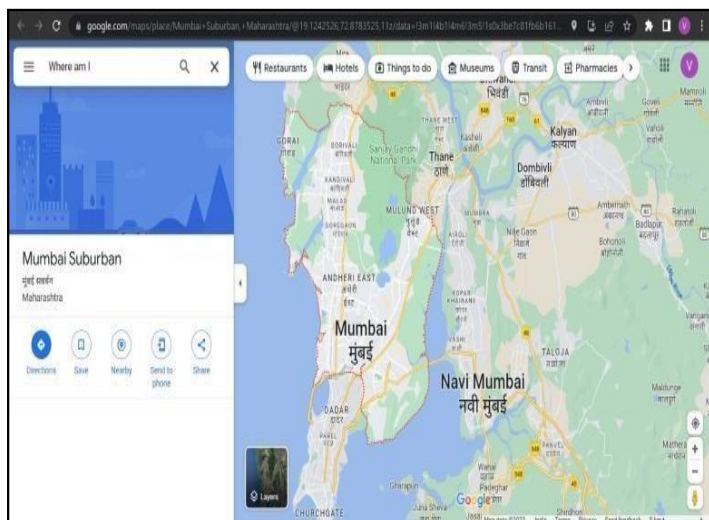
5.3 Search location

As shown in Fig.9 and Fig.10. When we a user requests look for a particular location, the assistant receives the request and searches for the location using Google Maps.

Fig.9 UI screen of searching location



Fig.10 Output screen of searching location



5.5 Searching top news

As shown in Fig.13 and Fig.14. When we ask the voice assistant to search top news, it receives the command and fetches the latest news using Google News.

Fig.13 UI screen for searching news



Fig.14 Output screen for searching news

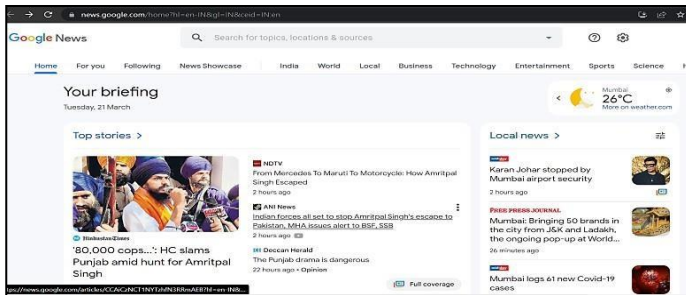
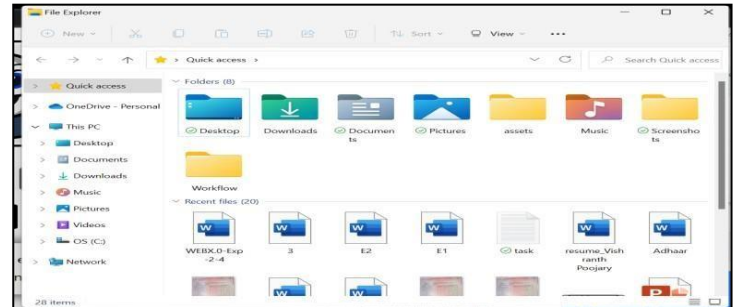


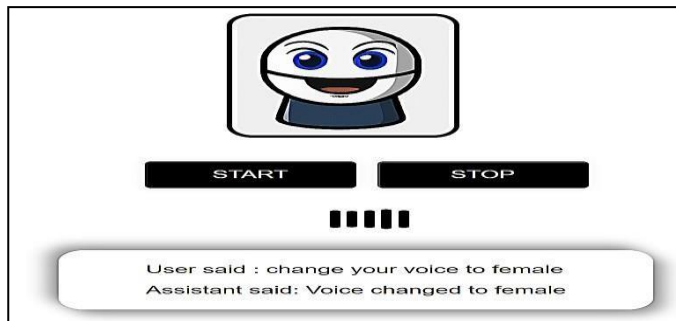
Fig.17 Output screen of opening specific folder



5.6 Changing Assistant's voice from male to female

As shown in Fig.15. User can easily change the voice of the voice assistant from male to female and vice versa by asking the assistant to change its voice from male to female.

Fig.15 Output screen of assistant's voice change



5.8 Open an application/software

As shown in Fig.18 and Fig.19. When we ask the voice assistant to open a particular application/software, it accepts the user command and opens the specific application/software.

Fig.18 UI screen of opening a software



5.7 Open a folder

As shown in Fig.16 and Fig.17. When we ask the voice assistant to open a particular folder, it takes the user input and opens that specific folder.

Fig.16 UI screen of opening specific folder

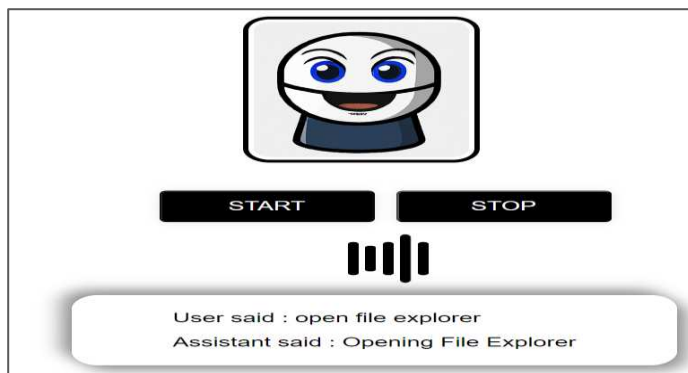
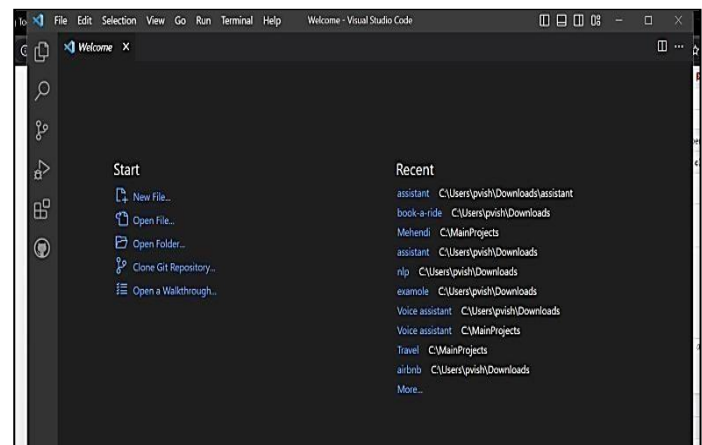


Fig.19 Output screen of opening a software



5.9 Open CMD

As shown in Fig.20 and Fig.21. When we command the voice assistant to open the command prompt (CMD), it receives the user request and performs the task by opening the command prompt (CMD) on the desktop screen.



Fig.20 UI screen of opening CMD

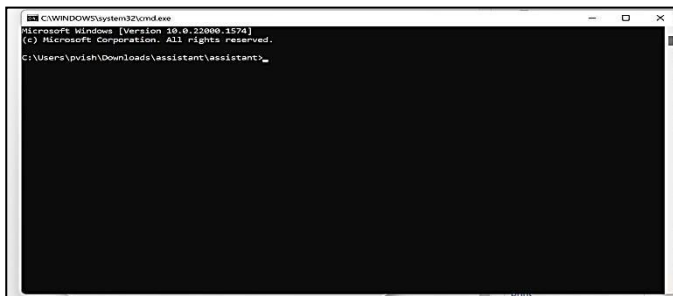


Fig.21 Output screen of opening CMD

5.10 Displaying Date and Time

As shown in Fig.22 and Fig.23. When we ask the voice assistant to show date and time, it takes user input and performs the tasks by displaying the date and time.



Fig.22 Output screen of displaying date



Fig.23 Output screen of displaying time

5.11 Increase/Decrease assistant's voice speed

As shown in Fig.24. When we ask the voice assistant to change the speed of its voice to slow/fast, it takes user request and changes the speed of the assistant's voice to slow/fast as commanded.

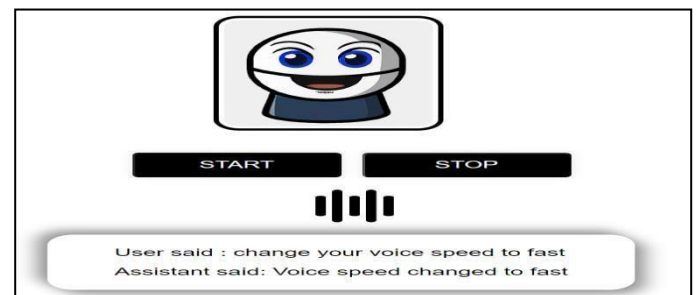


Fig.24 Output screen of changing voice speed

5.12 Assistant's introduction

As shown in Fig.25. When we ask the voice assistant to introduce itself, it takes user request and gives its introduction to the user.

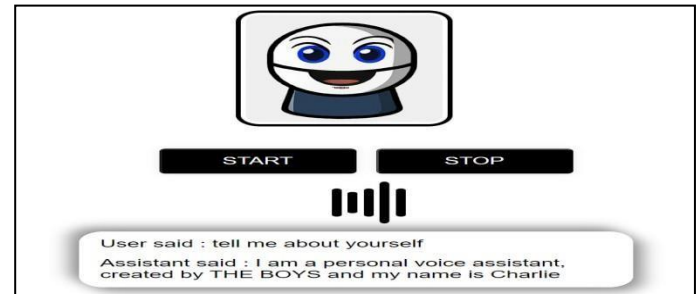


Fig.25 Output screen of assistant's introduction

6. FUTURE WORK

As AI is a developing field and people like to use it for their daily work, there is a lot of scope for it in the future. An additional feature of language would be added. With the help of this feature, user can give command in Hindi and assistant would respond in Hindi as well as English. More features would be added to make user experience better and work easier.

7. CONCLUSION

This paper “NLP based AI Voice Assistant” presents a comprehensive overview of the design and development of a Voice enabled personal assistant for desktop. The modular nature of this project makes it more flexible and easier to add additional features without disturbing current system functionalities.

This voice enabled desktop assistant, in today's life style will be more effective in case of saving time, compared to that of previous days. This Voice Assistant has been designed with ease of use as the main feature. The Assistant effectively completes some duties that users assign it. It not only responds to human orders but also to questions posed by the user or words said by the user, such as opening operations and tasks. The voice assistant is introducing itself to the user in a way that makes them feel more at ease and allows them to engage with it freely. The programme should also get rid of any extra manual labor that is required for users to complete each job. Instead of the more conventional physical input, the entire system operates on verbal input. This assistant can also perform a wide range of tasks, including restarting or turning off our PC with a single voice command.

REFERENCES

- [1] S. Subhash, P. N. Srivatsa, S. Siddesh, A. Ullas and B. Santhosh, "Artificial Intelligence-based Voice Assistant," 2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4), 2020, pp. 593-596, doi: 10.1109/WorldS450073.2020.9210344.
- [2] Felix, S. M., Kumar, S., & Veeramuthu, A. (2018). A Smart Personal AI Assistant for Visually Impaired People. 2018 2nd International Conference on Trends in Electronics and Informatics(ICOEI). doi:10.1109/icoei.2018.8553750
- [3] Ee. P. M., Santra, S., Bhowmick, S., Paul, A., Chatterjee, P., & Deyasi, A. (2018). Development of GUI for Text-to-Speech Recognition using Natural Language Processing. 2018 2nd International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech). doi:10.1109/iementech.2018.8465238
- [4] Iannizzotto, G., Bello, L. L, Nucita, A, & Grasso, G. M. (2018). A Vision and Speech Enabled, Customizable, Virtual Assistant for Smart Environments. 2018 11th International Conference on Human System Interaction (HSI).doi:10.1109/hsi.2018.8431232
- [5] Kepuska, V, & Bohouta, G. (2018). Next-generation of virtual personal assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home). 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC). doi:10.1109/ccwc.2018.8301638
- [6] J. Sayed ,M. W. Ashour, 2022 International Conference on Edge Computing and Applications (ICECAA), DOI: 10.1049/icp.2022.0334
- [7] Vindhya Vasini K; Tejasvi K; Pravallika K , 2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS) ,DOI: 10.1109/ICSCDS53736.2022.9760998
- [8] Satyendra Praneel Reddy Karri; B. Santhosh Kumar, Deep Learning Techniques for Implementation of Chatbots, 2020 International Conference on Computer Communication and Informatics (ICCCI) ,DOI: 10.1109/ICCCI48352.2020.9104143
- [9] Aashdeep Singh, P. Murugeswari, S D Prabu Ragavendiran, Amanpreet Kaur, AI-based Chatbot for Physically Challenged People, 2022 International Conference on Edge Computing and Applications (ICECAA),DOI:10.1109/ICECAA55415.2022.9936585
- [10] Rani P. J, Bakthakumar, J, Kumaar, B. P, Kumaar, U. P, & Kumar, S. (2017). Voice controlled home automation system using Natural Language Processing (NLP) and Internet of Things (IoT). 2017 Third International Conference on Science Technology Engineering & Management (ICONSTEM). doi:10.1109/iconstem.2017.8261311