AI-based Voice Assistant

Omkar Singh
Chandigarh University
Punjab, India
oksinghsengar@gmail.com

Kamya Brata Debnath
Chandigarh University
Punjab, India
erkbdunk@gmail.com

Harsh Raj
Chandigarh University
Punjab, India
harshraj3922@gmail.com

***Abstract*—Voice control is a major developmental factor that changes the way people live. The voice assistant is commonly used on smartphones and laptops. AI-based voice assistants are primarily operating systems that can capture a person's voice and respond with embedded words. This voice assistant will collect audio from the microphone and then convert that into text, later miles GTTS (Google text to speech). The GTTS engine will convert text content into audio report in English, and the audio will be made using a playback audio package of the Python programming language.**

Keywords— AI Assistant, GTTS Engine, Python Programming, Voice Assistant.

I. Introduction

Only in recent times have we seen significant changes in the way users use and find visible assistants. We already use them for a variety of purposes, including turning on / off the lights and listening to music by streaming apps such as Wynk Music and Spotify. This is a new way of communicating with technology gadgets that introduce lexical communication as a new companion to this technology.

Virtual Assistants were initially used to appoint professionals who provide additional online services. [1] The voice function is divided into three categories: Improving the current range, text to speech, text to purpose, and the purpose of the action will be fully developed. [6] Voice assistants should not be confused with visible assistants, who are people who work part-time and can do many tasks. Thanks to AI-based Voice Assistants, Voice Assistants anticipate all of our needs and take action.

Virtual Assistants were initially used to appoint professionals who provide additional online services. [1] The voice function is divided into three categories: Improving the current range, text to speech, text to purpose, and the purpose of the action will be fully developed. [6] Voice assistants should not be confused with visible assistants, who are people who work part-time and can do many tasks. Thanks to AI-based Voice Assistants, Voice Assistants anticipate all of our needs and take action.

II. Background

A. History of Voice Assistants

When it comes to the history of voice assistants, the first voice-activated toy was produced in 1911 with the 'Radio rex.'. IBM laid the groundwork for smart virtual assistants. Simon in the year 1994 as we know today. Digital speech recognition technology has become an aspect of personal computers in the 90s with Microsoft, Apple, Philips, etc., After various researches ‘Siri’ was introduced as the first modern digital Voice Assistant as the feature in iPhone 4S in 2011.

Many firms have used verbal communication to create system devices such as Amazon Alexa, Microsoft Cortana, Apple's Siri, Google Assistant, and more. Voice recognition, voice language nervousness, chat manager, native language production, text-to-speech converter, and knowledge base six components of a normal chat system.

B. Future Applications

Voice assistants may be utilised for two purposes in the future: Because broadband allows for more complex data processing in powerful data centres, the first quality of dialogue recognition will improve. Second, from the standpoint of the user, VAs facilitate engagement. Voice assistants can be used in the workplace to automate repetitive operations, such as opening video conferencing and booking meeting spaces.

C. Aim of this Study

Voice assistants may be utilised for two purposes in the future: Because broadband allows for more complex data processing in powerful data centres, the first quality of dialogue recognition will improve. Second, from the standpoint of the user, VAs facilitate engagement. Voice assistants can be used in the workplace to automate repetitive operations, such as opening video conferencing and booking meeting spaces.

III. Proposed Design

The project will provide a good grasp of an intelligent assistant that can recognize the commands issued by the user. Our assistant can easily understand the commands given by the user through vocal media and responds as required. Our assistant performs the most frequently asked requests from the user and makes their task easier.Through the mic, the voice assistant listens to the user's command. After listening it will say “done listening” and displays what the user said and acts accordingly.

In our project we have installed gTTS engine package to make the voice assistant speak like a normal human being. We have defined a function called ‘voice assistant speak’, as

explained in (1) The gTTS will analyze the command given by the user through microphone and searches in the browser the required response and convert that response into text.

tts = gTTS(text=audio_string, lang='en') (1)

gTTS is a programme that converts an audio string into text. This audio string is the response that the voice assistant should provide to the user. The text will be written in English, and the code for English is 'en'. This complete function is saved as 'tts'. This text is being saved as an audio file with the extension '.mp3'. A random number between 1 and 20000000 is assigned to each audio file. The command 'random.randint()' can be used to produce a random number. The moniker 'audio file' is used to save the entire '.mp3' extension file. Finally, we used the command provided in to save this audio file (2).

tts.save(audio file) (2)

This command (2) saves the audio file in the system. (Ex - 'audio24854.mp3').

IV. Tasks Performed by the Voice Assistant

- Can remember any person name till the usage session.
- Voice assistant name can also be changed unlike in other voice assistants.
- Play/download a YouTube music or video. When a user requests that the assistant "play/download me a song" or "play a movie," the assistant opens YouTube and plays or downloads the requested video/song.
- Google searches anything and returns the required content. When prompted to 'google search,' the assistant searches Google for the content requested and opens it in a browser.
- Opens the maps and displays the exact location requested by the user. When asked to 'find location' or 'google maps,' the assistant asks for the user's desired location and then opens Google Maps, highlighting the desired area.
- Provides reliable weather information for the location requested by the user. When asked for the 'current weather in,' the assistant provides the actual weather of the selected area in degrees Celsius, both maximum and minimum.
- The display is captured in a screenshot. When the assistant is asked to "capture," "capture my screen," "my screen," "screenshot," or "take screenshot," the assistant grabs the user's display and saves it in the provided path.
- Gives the live news around the world. When asked 'news for today', 'tell me the news', 'what's the news', 'what is the news', 'news', the assistant reads the first 5 updated news headlines from the website.
- Can able to tell whether the password has been hacked or not.
- If a person is in danger, our voice assistant can notify the user's location to the authorities or family members by saying "I'm in danger."
- It sends a mail to the username specified by the user. When told "send mail", the assistant asks to whom the mail has to be sent and it will send a mail according to that.

It can translate the user's words into any language and display the words in that language that the user specifies. User commands can be used to shut down or restart the system. These are some of the capabilities that we have introduced to our AI-based voice assistant so far, and we are working on adding many more.

V. Methodology

All voice assistants are created in programming languages, and they listen to the user's commands and respond accordingly. To create the AI-based Voice assistant in this project, we employed the Python programming language. When a user says, "Play me a Song" or "Open facebook.com," the voice assistant will answer by playing that song or opening the Facebook website. The voice assistant waits for a pause to indicate that the user has completed their request, then transmits the request to its database to be searched for.

- The user's request is divided into different commands so that our voice assistant can interpret it.
- Once within the commands list, our request is searched and compared with the other requests.
- The Voice assistant receives these commands from the commands list.
- Once the voice assistant receives those commands, then it knows what to do next.
- If the request is not clear enough for the voice assistant to handle, it will ask a question to ensure that it knows what we want. The voice assistant will accomplish the task that the user has requested if it believes it understands it well enough to process it.

A. Working of ASR

As shown in Fig 1 :The primary premise behind the operation of AI-based Voice Assistant is Automatic Speech Recognition, or ASR. [4] ASR systems first record the speech, after which the device creates a wavefile containing the words it hears, after which the wavefile is cleaned to remove background noise and normalise the volume, after which it is broken down into elements and analysed in sequences, after which the ASR software examines these sequences and uses statistical probability to find the entire words, and finally it is processed. [5] Element Recognition is a superior way for recognising elements than word decoding since it produces better outcomes.

Fig. 1. Process of ASR

It doesn't matter what kind of speech recognition software we employ because everything is handled by its ASR. In a nutshell, the process starts with the device collecting audio from the source, which is a microphone, and then passing the recorded speech waveforms to acoustic analysis, which is performed on three levels, as illustrated in Fig 2.

B. *Acoustic Analysis*

- **Acoustic Modelling:** It reflects whether the elements were pronounced or not, as well as what words can complete these elements in this procedure.
- **Pronunciation Modelling:** This examines the manner in which these elements are pronounced, as well as any accents or other anomalies.
- **Language Modelling:This is frequently used to determine contextual probabilities based on the elements recorded.**

Without any human intervention, all of the data that was captured is processed by A.i., and the speech waveforms data is then passed to the decoder, where it is finally transformed into text for further use as a command.

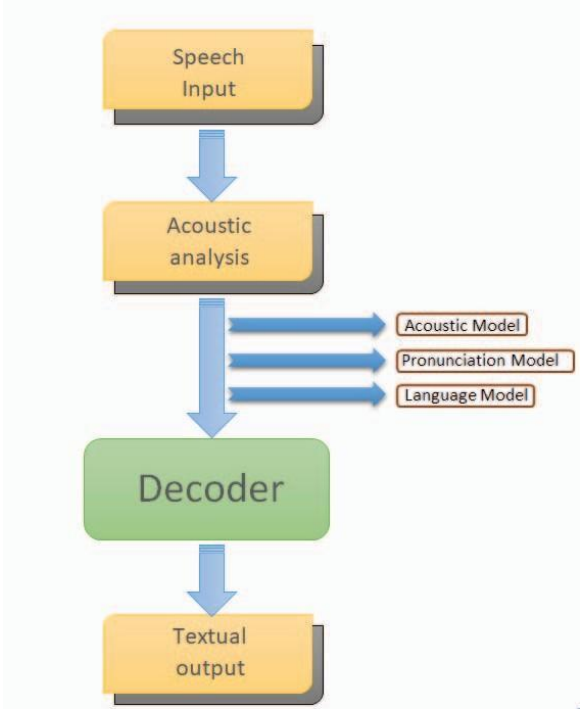


Fig. 2. Acoustic Analysis

VI. How our Voice Assistant is Useful

In this section, we'll go over a series of scenario events and show how our voice assistant can help. For example, if we want to go to a specific area, rather than opening Google Maps and inputting the destination, we can just tell our VA to identify the location and it will open the map and mark it. Playing a music on YouTube is a feature for modern VAs, but they can't download it. For example, if you want to download the song "give me some sunshine," we won't be able to do so with other voice assistants, but our VA will search the YouTube database and note the Video ID , with this the video is downloaded..

We must open Google and search for whatever information we require. However, in VA, it can be done simply by issuing a command, such as "what are super computers?"It compiles the most useful information from Wikipedia and provides it to us, and if we want to know what's going on in the world, we can ask it, "What are the top five news headlines of the day?" and it will retrieve the information and present it to us. We can learn a lot from this. When a girl states "I'm in danger," the Voice Assistant collects the girl's present location and transmits it to their parents/relatives as well as the nearest police station.

VII. Results

The required Python programming language packages were installed, and the code was implemented using the PyCharm Integrated development environment (IDE). The python code we developed runs in both Python 2.7 and Python 3.x, and the outputs we received in our AI-based voice assistant are listed below.

A. *Google Search Output*

As shown in below Fig 3. When we ask the voice assistant to search ‘MS DHONI’, it receives the request and performs the action by searching google.

Fig. 3. Output screen of performing Google Search

B. *Playing song/video on YouTube*

As shown in below Fig 4. When we ask the VA ‘Play song’, it responds by saying ‘which song you want to play’ and we ask the VA a particular song/video then the VA will perform the task by playing song/video on YouTube.

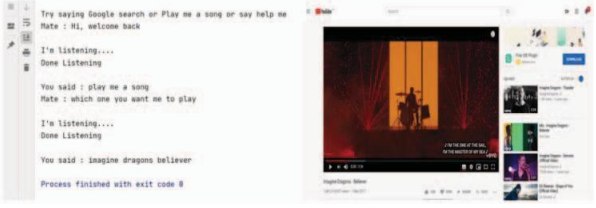


Fig. 4. Output screen of Playing song on YouTube

C. *Searches any location*

Figure 5 illustrates this. When we ask the voice assistant to search for a specific location, it receives the request and uses Google maps to find it.

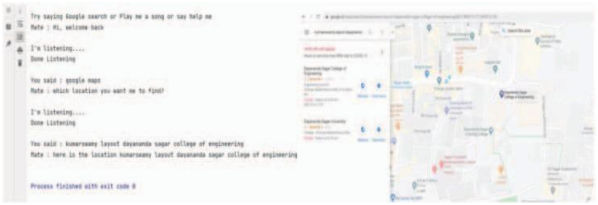


Fig. 5. Output screen of Searching Location

D. *Current News*

As demonstrated in Figure 6, When we ask the voice assistant for current news, it receives the request and answers with the top 5 stories for that day, as well as the website name.

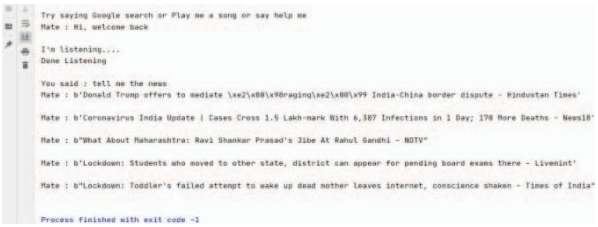


Fig. 6. Output screen of Showing Current News

VIII. CONCLUSION

In comparison to other helpers, we have implemented several things in our project. Because it is a hands-free application, it is quite useful in human life nowadays. It's a very straightforward application. It is also utilised in the corporate world, for example, in laboratories where people wear gloves and body suits for safety reasons, making it difficult to type. With a voice assistant, they can receive any information they need, making their work easier.

Voice assistants are useful in a variety of fields, including education, daily life applications, home appliances, and more. Voice assistants are also useful for illiterate people, as they can obtain any information simply by speaking to the assistant. Thanks to AI-based voice assistants, people can enjoy luxury.

Voice assistants are becoming more prevalent in everyday life. Many voice assistant firms are working to improve interaction and provide additional capabilities, and many young people have begun to use voice assistant in their daily lives, with positive results reported by many sources.In the previous few years, voice assistants have come a long way.

REFERENCES

[1] Saadman Shahid Chowdury, Atiar Talukdar, Ashik Mahmud, Tanzilur Rahman, “Domain specific Intelligent personal assistant with bilingual voice command processing,” IEEE 2018.

[2] Polyakov EV, Mazhanov MS, AY Voskov, LS Kachalova MV, Polyakov SV “Investigation and development of the intelligent voice assistant for the IOT using machine learning,” Moscow workshop on electronic technologies, 2018.

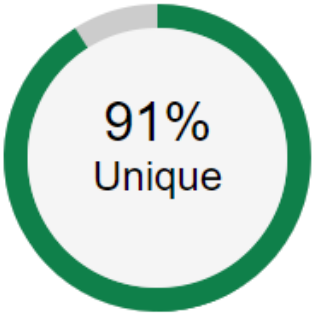
[3] Khawir Mahmood, Tausfer Rana, Abdur Rehman Raza, “Singular adaptive multi role intelligent personal assistant (SAM-IPA) for human computer interaction,” International conference on open source system and technologies, 2018.

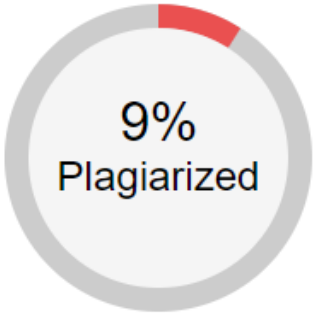
[4] Veton Kepuska and Gamal Bohota, “Next generation of virtual assistant (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home),” IEEE conference, 2018.

[5] Piyush Vashishta, Juginder Pal Singh, Pranav Jain, Jitendra Kumar, “Raspberry PI based voice-operated personal assistant,” International Conference on Electronics And Communication and Aerospace Technology, ICECA, 2019.

[6] Laura BURbach, Patrick Halbach, Nils Plettenberg, Johannes Nakyama, Matrina Ziefle, Andre Calero Valdez, “Ok google, Hey Siri, Alexa. Acceptance relevant of virtual voice assistants,” International communication conference, IEEE 2019.

Plagiarism Report





100%

View Plagiarized Sources

9% Plagiarized +