Voice Assistant using Artificial Intelligence

Jothika R
Department of Data Science
Vellore Institute of Technology
Vellore, India
Jothika.r2031@gmail.com

Abstract - The Voice assistance is a recent and major advance feature and technology used as a human assistance. User can easily interact with computer by speaking. In recently the assistant are user in many places like mobile, laptop, electronic devices etc. Artificial Intelligence based voice assistance work as a operating system and that recognize the voice of humans and accordingly reply for it. This process works by the Speech Recognition and using GTTS (Google Text To Speech). Using speech recognition audio will be received from the humans and converting into text of computer language for that they use GTTS. It will convert the audio file into a speech of English language. This are the module and packages were used in python programming language.

Keywords : Speech recognition, voice assistance using AI, Desktop voice assistance, GTTX, Pyaudio,playsound, python.

I. INTRODUCTION

In recent time, The virtual assistance plays a major role in daily of human beings. It will interact with user as human beings works as a assistance. They will schedule the task of do list. What ever we ask the question it will replay for the questions. Example like weather report, current news, meeting schedule, reminder and also it will easily connect with other devices like home automation, car device . With that feature we using this like Alexa, Siri, Google Assistant etc.

Artificial Intelligence based voice assistance can be used in many task like switching on/off the lights, play music on spotify or wynk music etc. It is easily interact and communicate with human behaviour. The main methodology of voice assistance is easy accessible for any user. The process of voice assistance is getting the speech from the user and converting the voice into text using speech recognition. It will recording the speech of voice by using microphone as input device and convert into machine language or system understandable format of text or word phrases and reply according to the questions.

AI based Voice Assistance are used in many sector such as It field, Medical, Personal assistance, Home automation etc. Voice based search and device are going to play a major role in future generation as well as now also. In this proposal we are going to build the AI based voice assistance for system desktop usage Its going to do the all task in the system by the voice without any inconvenience.

II. LITERATURE REVIEW

The iPhone's most common speech recognition is "SIRI" which allowas the end client to offer voice flexibility to end client and additionally responds to the client's voice charges. With the support of the client with voice or content tools to be processed, it has features of intelligent speech recognition and returns the yie ld in varios systems such as the operation to be done or the item to the end client [2]. This virtual assistant, Cortana, is named after the Microsoft Artificial Intelligence (AI) device in the Halo video game. In this game, this fictional AI character provides resources such as back stories to aid the gamer. According to Wikipedia, This design was based on the Egyptian queen Nefertiti [3]. Not only for Google Android, but also for Apple iOS platforms, Google Now is an intelligent personal assistant feature. Same as Apple's Siri, Google Now can

answer questions and act accordingly. Also, Google Now uses user's preferences and search history to deliver a better result. Because of this feature, it is capable of providing personalized information before the user even requests it. [4]Amazon also created Alexa, a virtual digital assistant for the Amazon Echo and Echo Dot series of computing devices. Alexa responds to the user with the Amazon products such as music, weather, sports, and more. Amazon's Alexa Back-Engine is running on Amazon's Web Servers in the cloud. Besides Echo products, Alexa also supports Amazon's Fire HD tablet and Fire TV set-top box products. A variety of chosen third-party devices have already begun to support Alexa, including Nucleus Intercom, Ford SYNC vehicle infotainment systems, and Invoxia's Triby speaker and messaging system [5].

III. PROPOSED DESIGN

The project will give a proper knowledge about AI assistant which have a capable to understand the command and voice from the humans or user. Our voice assistance can easily understood the voice note given by the users though the microphone and responds according to required. Our voice assistant performs the frequently asked question session like that they reply the answers from the user and the task makes more easier. Initially, The AI based voice assistant listens the audio voice from the user by the command of recording or listening through the microphone. After recording the voice it shows done listening command then try to answer the question accordingly.

In our project we use the speech recognition and gTTS engine package from the python programming to make a normal conversation with human beings by voice assistant. The gTTS will analyze the command of voice by from the user through microphone device and search the thing through the browser.

gtts = gTTS(Text=Audio_string,lang ='en')

gTTS is used to convert the audio string to text or word. The audio is nothing but the response of voice assistance from the user. For the conversation the language of text should be used in 'English Language' as 'en'. From the equation(1) save the text data in gtts variable. The file extension should be in 'mp3' file. Saved the file by using the gtts variable as 'audio file'.

gtts.save(audio file)

The command save the file using the random number to identify the file name (example: 'audio1267864.mp3').

IV. PERFORMANCE OF VOICE ASSISTANT

- AI based voice assistant remember the name of the user or person name till the session complete.
- Voice assistant name also changed while in the conversion by the user or other voice assistants.
- *Greeting :* Voice assistance can also perform the normal conversation like greeting (example : "How are you ?" , "good Morning" etc.) reply for the user questions.
- *Time report* : AI voice assistance response, for current time and date for the particular day from conversation by the user.
- *Weather Reporting :* It will reporting the weather and climate for the particular day by using the browser it will reply accordingly.
- *Google search :* search anything from the google search engine or browser asked by the users and tells the required information by reply for the user. If user asked 'Google search' it will the default open browser.
- *Video play / YouTube search :* When user asked to search something on YouTube it browse the page show the relevant video according to the question from the user.

- *Mail Sending :* It will send a mail to the username or email id specified by the user. When user tells send the email to other email id. The assistance send the mail to the other person id.

- *Map :* when user ask find the location for particular place it will open 'Google map' search for the location what user send display the direction for the place.
.
- *Screenshot capture:* whether the user ask to take the screen shot for current display page , AI voice take a screen hot of the page and save the page file for the user.

- *Current news* : Give the current news and affairs around the world. When user asked the 'news of today ' ,It will tells topic trending news from the browser search as a headlines.

- *Wikipedia :* If the user need to know any information. AI voice assistance search the information from Wikipedia and reply for the user.

- *Social Media / Instagram :* User ask to open ask to open any social media like Instagram, facebook, twitter etc. It browse the page and display the page to the user.

- *Math Calculation:* If user ask any arithmetic calculation. AI voice assistance calculate the math and reply the answer for the question.

- *Translation :* It can translate the words and audio text into any language using 'Google translator' which is specified by the user.

- It can able to shutdown or restart the system by the user voice command etc.

These are the features of our AI voice assistant as of now. We can add so many features for future development and we are working on many more features to embed into this assistant.

V. METHODOLOGY

The methodology of AI voice assistant has the speech recognition to know the speech of user and use the gTTS to convert the audio into text after that browse, what user ask to do the task. (example : user ask to open social media like Instagram it will open the page or if user ask to 'play a music' it will play on Spotify. The process of voice assistance will given below.

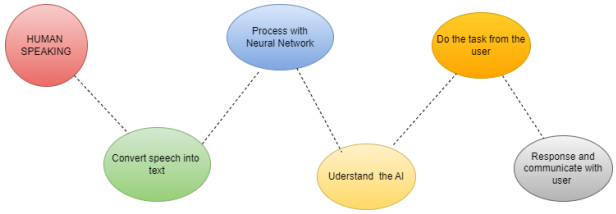

Fig. 1. Process of voice assistant

From the above Fig. 1, User or human start a conversation with AI based voice assistant. It will listening the audio and recording the voice note they converting the speech into text then do the neural network process for understand the text then understanding the AI by using GTTS (google text to speech) it will search according to that and done the task for the user then reply for the question and task do the process again after the voice from user and repeat the process until the conversation end from the user.

VI. SPEECH RECOGNITION:

Speech recognition is a package or AI based package by using in python language. It has the ability to listen the voice spoken words
And identify them. It will convert the speech into text of words make conversation with user. It allow the computer to understand what the user spoke and reply for it accordingly. It takes the input of voice from the microphone and does the process.

In the above Fig. 2. The flow chart show the process of voice assistant how it works . It will get the voice note from user and it use to speech recognition. If its recognize the voice

then convert into text or ask the user to speak again.

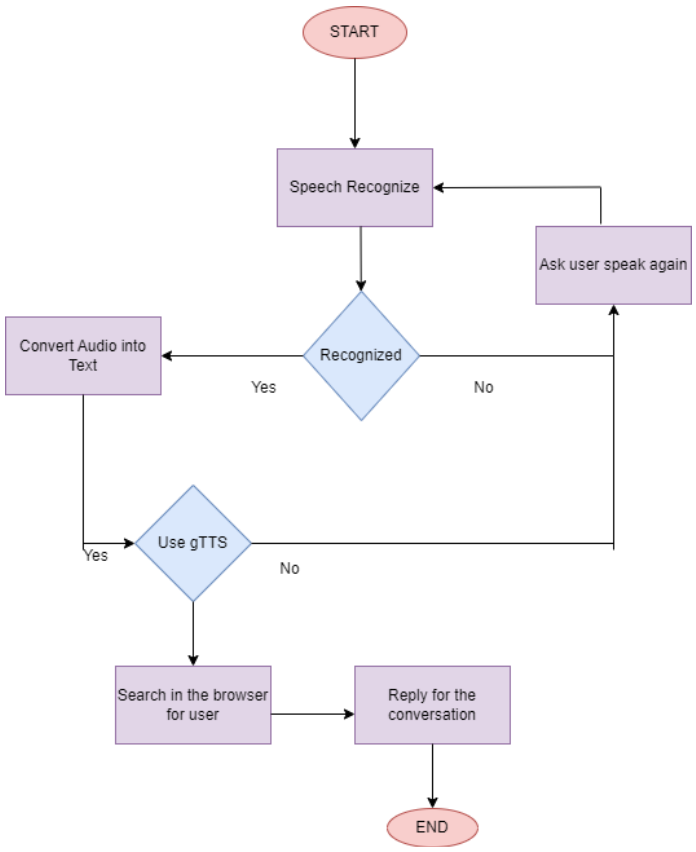

Fig. 2. Flow chart and process of voice assistant

After converting into text then, use the gTTS (Google text to speech) and then do the searching process in the browser and reply for the conversation this process do again and again until the conversation end from the user. All the data are recorded and processed by the artificial Intelligence without any interaction from user or Human beings.

VII. RESULTS

We show the result and output for AI based voice assistant. In this project , we use Python language and install required packages for it. For developing implementation are used by PyCharm IDE (Integrated development environment) and the python version is .11 and the few output are given below .

A. SELF INTRODUCE AND GREETING

As shown in below Fig.3. When user ask about self introduction and usage of AI assistant. It will introduce it self and make a formal conversation user like 'Hello', 'What your name' 'Good Morning' etc.

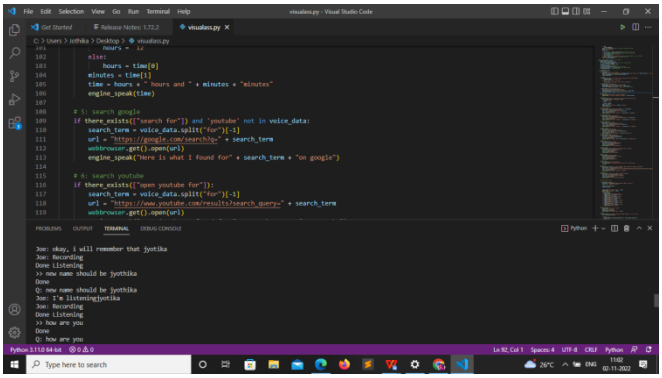

Fig. 3. Output screen for self introduction and greeting

B. GOOGLE SEARCH

As shown in Fig.2. when user ask the voice assistant to search for "something". It recording the voice and do the task by the action. It will the content on google search.

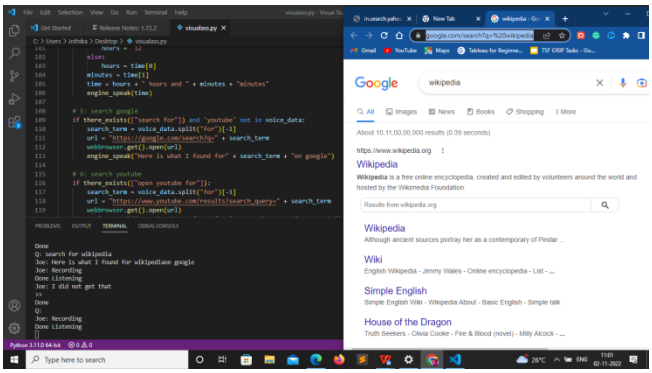

Fig.4. Output screen for Google search

C. YOUTUBE SEARCH / PLAY MUSIC

As shown in Fig The user ask the voice assistance to "open YouTube" and the AI voice assistant reply "What content need to search or play". In the fig user ask to play the video about data science so the VA search according the reply.

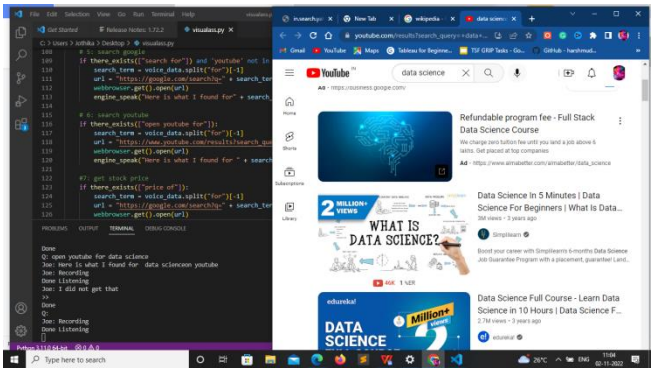

Fig. 5. Output screen for YouTube search

D. WEATHER REPORT

As shown in Fig.6. When User ask the voice assistant “Tell me the weather report” It will search the weather on google an it shown the weather result by browser search.

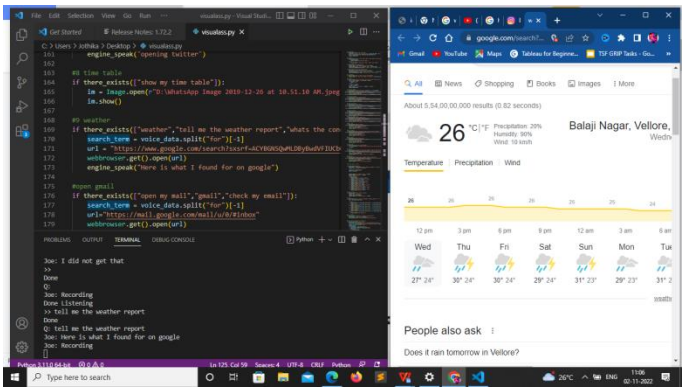

Fig. 6. Output screen for Weather report

E. MAIL INBOX

As shown in Fig.7. User asks to open the mail inbox to the voice assistant. It will open the mail inbox and show it to the user.

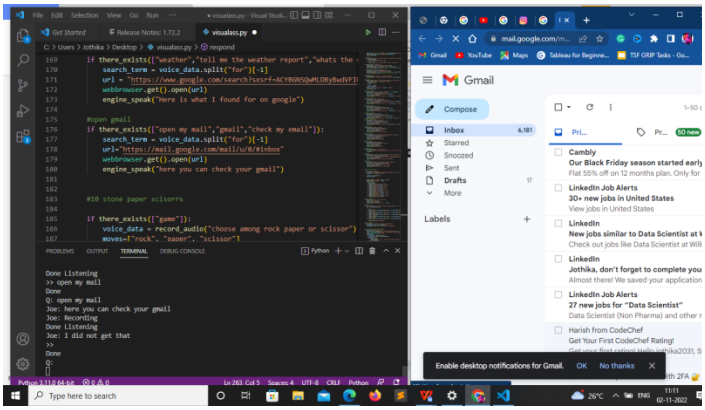

Fig. 7. Output screen for Mail Inbox

F. MATH CALCULATION

As shown in Fig.8 . AI based voice assistant also do the Arithmetic calculation when user ask to calculate. “9 plus 4”. It will reply 14 after do the math of calculation.

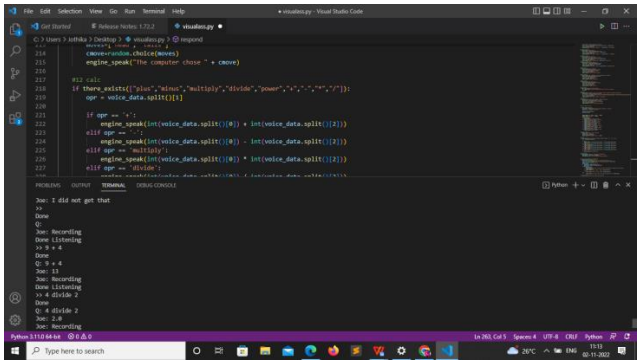

Fig. 8. Output screen for Math calculation

VII. CONCLUSION AND FUTURE WORK

In our project we have implemented many feature of Voice based assistant compared to other. Now days, This kind of assistant are used in wide place and it very useful in human daily life with smart technology and it is user friendly. Easy to access and simple handling application. As well as it is used I many field such as daily life application, education, health, marketing ,finance, home appliances etc. Its works as a personal assistant and guide , planner as like so many role its for a user to handling a smart work for daily life. Many companies and software , mobile devices in that AI based voice assistant plays a major role.

We develop a desktop based Voice AI assistant. It will work in system only and need develop good looking GUI design for the future development and also need add more upcoming feature like device access technology voice of user and mobile based feature as a application.

VIII. REFERENCE

[1]Subhash, S., Srivatsa, P. N., Siddesh, S., Ullas, A., & Santhosh, B. (2020, July). Artificial intelligence-based voice assistant. In 2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4) (pp. 593-596). IEEE.

[2] Nasirian, F., Ahmadian, M., & Lee, O. K. D. (2017). AI-based voice assistant systems: Evaluating from the interaction and trust perspectives.

[3] Lee, K. Y., Sheehan, L., Lee, K., & Chang, Y. (2021). The continuation and recommendation intention of artificial intelligence-based voice assistant systems (AIVAS): the influence of personal traits. Internet Research, 31(5), 1899-1939.

[4] Maedche, A., Legner, C., Benlian, A., Berger, B., Gimpel, H., Hess, T., ... & Söllner, M. (2019). AI-based digital assistants. Business & Information Systems Engineering, 61(4), 535-544.

[5] Desjardins, A., Psarra, A., & A. Whiting, B. (2021, June). Voices and Voids: Subverting Voice Assistant Systems through Performative Experiments. In Creativity and Cognition (pp. 1-10).

[6] Lee, O. K. D., Ayyagari, R., Nasirian, F., & Ahmadian, M. (2021). Role of interaction quality and trust in use of AI-based voice-assistant systems. Journal of Systems and Information Technology.

[7] Desjardins, A., Psarra, A., & A. Whiting, B. (2021, June). Voices and Voids: Subverting Voice Assistant Systems through Performative Experiments. In Creativity and Cognition (pp. 1-10).

[8] Rzepka, C., Berger, B., & Hess, T. (2022). Voice assistant vs. Chatbot–examining the fit between conversational agents' interaction modalities and information search tasks. Information Systems Frontiers, 24(3), 839-856.

[9] Zwakman, D. S., Pal, D., & Arpnikanondt, C. (2021). Usability evaluation of artificial intelligence-based voice assistants: the case of Amazon Alexa. SN Computer Science, 2(1), 1-16.

[10] Saibaba, C. M., Waris, S. F., Raju, S. H., Sarma, V. S. R. K., Jadala, V. C., & Prasad, C. (2021, August). Intelligent voice assistant by using OpenCV approach. In 2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC) (pp. 1586-1593). IEEE.