

Jarvis: Artificial Intelligence-based Voice Assistant

Brindha R¹, Jasmine Hepziba S², Nehashree B³

¹ UG – Information Technology, Bannari Amman Institute of Technology, Erode, Tamil Nadu

² UG - Information Technology, Bannari Amman Institute of Technology, Erode, Tamil Nadu

³ UG – Computer Science and Engineering, Bannari Amman Institute of Technology, Erode, Tamil Nadu

ABSTRACT

In this project, we'll create a voice assistant named Jarvis that can open files, conduct web searches, retrieve Wikipedia results, read PDF files, launch applications, and even make jokes and provide you advice. Artificial intelligence based Voice assistants are the working frameworks that can perceive human voice and answer through coordinated voices. The voice assistant will collect audio from the microphone, convert it to text, and send it over GTTS (Google text to speech). The English-language audio file created by the GTTS engine will be played using the play sound module of the Python programming language. We are importing libraries such as "os" for operating system tasks, "speech recognition" for voice command analysis, "pyttsx3" for voice command into text strings, "Time" for fetching the current time and date, "web browser" for web-based tasks on default web browsers, and "SMTP" (Simple Mail Transfer Protocol) for email transmission between accounts over the internet.

Keywords—GTTS, pyttsx3, SMTP, speech recognition

1. Introduction

Voice assistants are devices/applications that utilize voice acknowledgment innovation, natural language processing and AI to answer people. The idea of virtual assistants in prior days is to depict the experts who offer subordinate types of assistance on the web. [1] The occupation of a voice is characterized in three phases: Speech to text; Text to Intention; Intention to action; This is the principal framework proposed to intelligently address clients in natural language.[2] It is enlivened by the character Jarvis, played by Edward Jarvis in the Marvel Studio Creations Iron Man, who served Tony Stark [Robert Downey Jr.] as his most clever speculative collaborator, loaning him help into each conceivable angle like technology, medical, social, and other cognizant educational planning viewpoints. Voice assistant will be completely evolved to work on the current range. Thanks to AI-based voice assistants, we expect our every need to be met, and it makes a move.

2. Related Works

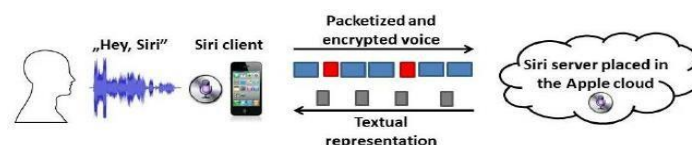


Fig 1. Apple-Siri voice service architect

A voice recognition system offers a range of services through the use of artificial intelligence systems integrated with natural language processing. [5] The system is divided into two types-Platform management type and open API artificial intelligences. Although artificial intelligence is capable of complex operations, utilizing services on other platforms is currently not possible reliant on a specific platform. Open API type is feasible to use services on different platforms. "Radio Rex" was the first voice-activated toy to be released in 1911. The 1990s saw the emergence of digital speech recognition technology in personal computers because of companies like Microsoft, Apple, Philips, etc. After

extensive study, [3] "Siri"—the first contemporary digital voice assistant—was launched in 2011 as a feature of the iPhone 4S. Oral dialogue systems have been employed by several businesses to build systems like Microsoft Cortana and [4] Amazon Alexa. Six elements make up general dialogue systems: knowledge base, text to speech converter, dialogue manager, natural language creation, voice recognition, and vocal language apprehension.

2.1 Future Applications

Voice assistants could be utilized for two improvements in the future: Firstly, the quality of dialogue recognition will improve because broadband enables more complex data processing in powerful data centers. Second, the usage of virtual assistants as an interaction tool. Voice assistants can be used in businesses to automate routine operations. For instance, Amazon's Alexa can start video conferences and reserve meeting spaces, etc.

2.2 Aim of the Study

Our project's primary goal is to develop an Intelligent Personal Assistant function that can activate and deactivate smart phone applications with the use of a Voice User Interface (VUI), which is used to hear and process audio commands.

3. Proposed System

The project will give fair information about the intelligent assistant, which is equipped to understand the orders given by the client. With the utilization of voice media, the client may effectively impart orders to our assistant, who will then complete them on a case-by-case basis. Our assistant plays out the most common demands from the client and makes their task easier. Our voice assistant pays attention to the order given by the client through the microphone. In the wake of listening, it will be able to recognize what the client said and act appropriately.

The gTTS engine brought into the system produces the total speech synthesis [1]. Deep Learning Neural Networks are used by Google's Text-to-Speech to recognise and interpret the speech provided to the module. gTTS's built-in mixer mechanism is used to interpret the speech for the engine. All of the collected data is processed to the gTTS using the Microphone as the source for the JARVIS, and the output is then either converted to ".txt" or ".mp3" format. To make the voice assistant sound more human, we incorporated the gTTS engine package in our project. The gTTS will analyse the instruction issued by the user through the microphone, search the browser for the necessary response, and translate that response into text. We have developed a function called "speak" as mentioned in (1).

$$tts = gTTS(\text{text}=\text{audio_string}, \text{lang}='en') \quad (1)$$

gTTS is essentially used to convert over the sound string into text. This sound string is only the reaction which the voice assistant should give the client. The language of the text is decided to be English, the code for English is 'en'. We save this whole function into 'tts'.

4. Task accomplished by the voice assistant

- Play a song or open YouTube. When client asked 'play music' it plays the music for the client. When the user says 'open YouTube' the assistant open YouTube for user.
- Can remember anything. The assistant queries "what should I remember" when the user says "Remember that." Further, it can reveal its memories.
- Search anything from Wikipedia. If the user says 'Wikipedia' the assistant searches the content asked by the user.
- Takes a screen capture of the display. When the user requests "Screenshot", the assistant captures the display and stores it in the path given by the user.
- It provides accurate weather information. When a user requests 'weather', the assistant gives the accurate weather of the location requested by the user.

- It broadcasts live news around the world. When a user asks for 'headlines', the assistant reads the latest news headlines from the website.
- It sends an e-mail [6] to the username indicated by the client. At the point when the client asks 'send email', the associate asks to whom the mail has to be send and it will send a mail as indicated by the client.
- Can able to up/down or quiet the volume by client order and so forth.
- Send WhatsApp messages to the phone number ordered by the user.
- Set Alarm. If the client asked for an "alarm", the assistant asked, "What time do you want to set the alarm?" and the alarm will ring at the time set by the client.

5. Methodology

Programming languages are used to create voice assistants, which listen to the verbal orders and respond in accordance with the user's requirements. In this project, we developed an AI-based voice assistant using the Python programming language. The voice assistant will answer a user's request to "Play me a song" or "Open youtube.com" by playing that specific song or launching YouTube. The voice assistant waits for a pause to verify that the user has finished their request before sending it to its database to be looked up.

- The user's request is split into individual commands so that our voice assistant can comprehend it.
- Our request is looked up and contrasted with the other requests once it has been added to the commands list.
- Following that, the commands list relays these instructions to the voice assistant.
- The voice assistant knows what to do next after receiving those commands.
- If the request is not clear enough for the voice assistant to grasp it, it asks us to say that again.
- The voice assistant will carry out the request made by the user if it thinks it has understood it well enough to process it.

6. Results

The required packages of Anaconda have been installed for Python programming and the code was implemented using Spyder-Scientific PYTHON Development EnviRonment (IDE) and the Python code we have developed runs in Python 3.9.12. Below are the few outputs which we have received in our Jarvis AI-based voice assistant.

A. Send email

It sends an email to the user name that the client has specified. When a client requests to "send email," the assistant asks to whom the mail has to be send and asks "What should I say?" then sends the email in accordance with the client's instructions.

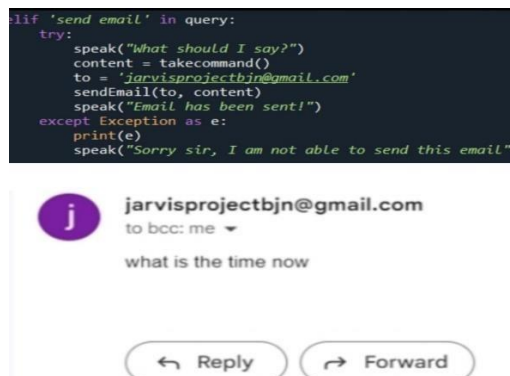
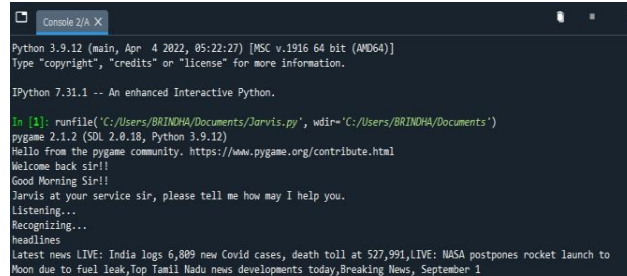


Fig. 2. Output of send email

B. Headlines

It broadcasts real-time news worldwide. The assistant reads the most recent news headlines from the website when a user requests "headlines."



```
Python 3.9.12 (main, Apr  4 2022, 05:22:27) [MSC v.1916 64 bit (AMD64)]
Type "copyright", "credits" or "license()" for more information.

IPython 7.31.1 -- An enhanced Interactive Python.

In [1]: runfile('C:/Users/BRINDHA/Documents/Jarvis.py', wdir='C:/Users/BRINDHA/Documents')
pygame 2.1.2 (SDL 2.0.18, Python 3.9.12)
Hello from the pygame community. https://www.pygame.org/contribute.html
Welcome back sir!!
Good Morning Sir!!
Jarvis at your service sir, please tell me how may I help you.
Listening...
Recognizing...
headlines
Latest news LIVE: India logs 6,889 new Covid cases, death toll at 527,991,LIVE: NASA postpones rocket launch to Moon due to fuel leak,Top Tamil Nadu news developments today,Breaking News, September 1
```

Fig. 3. Output of headlines

C. Open YouTube

It opens YouTube. When the user commands the assistant to "open YouTube," the assistance does so.

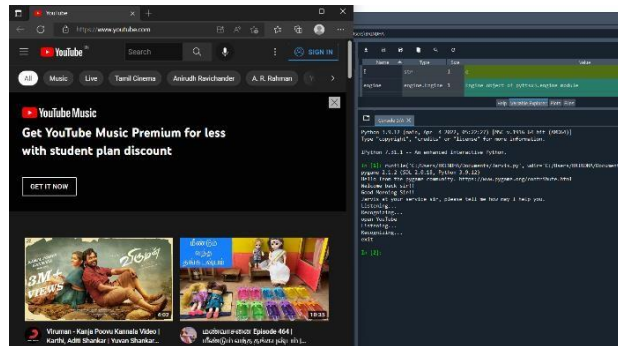
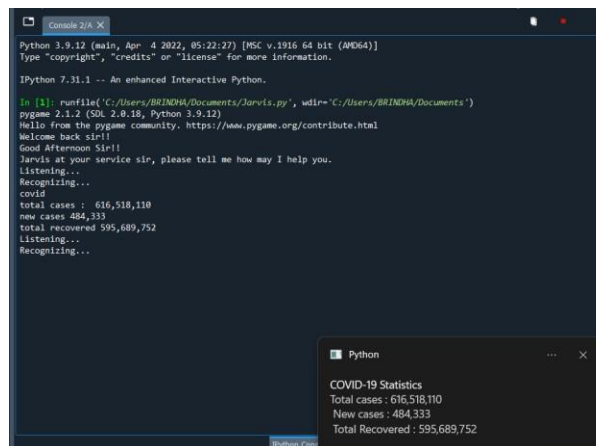


Fig. 4. Output of Open YouTube

D. Covid-19 statistics

It reports covid-19 statistics as total cases, new cases and total recovered.



```
Python 3.9.12 (main, Apr  4 2022, 05:22:27) [MSC v.1916 64 bit (AMD64)]
Type "copyright", "credits" or "license()" for more information.

IPython 7.31.1 -- An enhanced Interactive Python.

In [1]: runfile('C:/Users/BRINDHA/Documents/Jarvis.py', wdir='C:/Users/BRINDHA/Documents')
pygame 2.1.2 (SDL 2.0.18, Python 3.9.12)
Hello from the pygame community. https://www.pygame.org/contribute.html
Welcome back sir!!
Good Afternoon Sir!!
Jarvis at your service sir, please tell me how may I help you.
Listening...
Recognizing...
covid
total cases : 616,518,110
new cases 484,333
total recovered 595,689,752
Listening...
Recognizing...
```

Python
COVID-19 Statistics
Total cases : 616,518,110
New cases : 484,333
Total Recovered : 595,689,752

Fig. 5. Output of Covid-19 report

7. Conclusion

Compared to other assistants, we have implemented a lot of stuff in our project. Voice assistants are helpful in a variety of areas, including education, day-to-day use, household appliances, etc. They are particularly beneficial for the illiterate because they may obtain any information by simply speaking to the assistant. The use of voice assistants in daily life is growing. The majority of the user's tasks

are automated using Chrome, GitHub, and YouTube, including sending WhatsApp messages, playing music, performing Google and Wikipedia searches, setting alarms, sending emails, and more. The development of voice assistants has increased significantly over the past two years.

References

1. Subhash S, Prajwal N Srivatsa, Siddesh S, Ullas A, Santhosh B “Artificial Intelligence-based Voice Assistant”, IEEE 2020.
2. Ravivanshikumar Sangpal, Tanvee Gawand, Sahil Vaykar, Neha Madhavi “JARVIS: An interpretation of AIML with integration of gTTS and Python” , IEEE 2019.
3. 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.
4. Veton Këpuska, Gamal Bohouta , “ Next-Generation of Virtual Personal Assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home) “, IEEE 2018.
5. Tae-Kook Kim, “Short Research on Voice Control System Based on Artificial Intelligence Assistant” , IEEE 2020.
6. Shreyas R , Supreetha R Kashyap , Vishal M S , Vybhavi J, “Email Voice Assistant”, 2021.
7. 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
8. D. Aimless and S. Umatani., A tutorial for adding knowledge to your robot, 2006 Available at:<http://www.pandorabots.com/botmaster/en/tutorial?ch=1>. Accessed on January 05, 2019.
9. M. L. Mauldin. Chatterbots, tinymuds, and the turing test - entering the loebner prize competition In Proceedings of Twelfth National Conference on Artificial Intelligence (AAAI-94), Seattle, Washington, 1994.
10. H. P. Grice. Logic and converstion. In P. Cole and J. L. Morgan, editors, SpeechActs. Syntax and Semantics volume 3, pages 41-58. Academic Press, New York, 1975.
11. J. Weizenbaum. Eliza - a computer program for the study for the study of natural language communication between man and machine. Commun. ACM, 9(1):36-45, 1966
12. R. S. Wallace. Aiml - artificial intelligence markup language, 2006. Available at:<http://www.alicebot.org/aiml.html>. Accessed on January 05, 2019.
13. J. R. Searle. Natural Language Understanding. The Benjamin Cummings Publishing Company, Inc, New York, NY, 1995
14. E. A. Schegloff. Conversation analysis and communication disorders. In C. Goodwin, editor, Conversation and Brain Damage. Oxford University Press, New York, 2002
15. 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.