Desktop Voice Assistant

K Vishnu Vardhan
VIT University Chennai, Tamil Nadu ,India
koyyada.vishnuvardhan2018@vitstudent.ac.in

P Guru Pavan Kumar
VIT University Chennai, Tamil Nadu ,India
puretiguru.pavankumar2018@vitstudent.ac.in

S Ruthvik
VIT University Chennai, Tamil Nadu ,India
Setty.ruthvik2018@vitstudent.ac.in

N Bhanu Prakash
VIT University Chennai, Tamil Nadu ,India
bhanu.prakash2018@vitstudent.ac.in

Divya Sai Ajay Jasti
VIT University Chennai, Tamil Nadu ,India
jastidivya.saiajay2018@vitstudent.ac.in

Abstract-**As we know Python is an emerging language so it becomes easy to write a script for Voice Assistant in Python. The instructions for the assistant can be handled as per the requirement of user. Speech recognition is the process of converting speech into text. This is commonly used in voice assistants like Alexa, Siri, etc. In Python there is an API called Speech Recognition which allows us to convert speech into text. It was an interesting task to make my own assistant. It became easier to send emails without typing any word, Searching on Google without opening the browser, and performing many other daily tasks like playing music, opening your favorite IDE with the help of a single voice command. In the current scenario, advancement in technologies are such that they can perform any task with same effectiveness or can say more effectively than us. By making this project, I realized that the concept of AI in every field is decreasing human effort and saving time.**
Functionalities of this project include: 1.It can send emails.
2.It can read PDF.
3.It can send text on WhatsApp.
4. It can open command prompt, your favorite IDE, notepad etc.
5.It can play music.
6.It can do Wikipedia searches for you.
7. It can open websites like Google, YouTube, etc., in a web browser.
8.It can give weather forecast.
9.It can give desktop reminders of your choice.

I. Introduction

When Artificial Intelligence is combined with robots, it demonstrates our ability to think like humans. In this case, a computer system is built in such a way that it typically requires human intervention. Given that Python is a new language, writing a script for Voice Assistant in Python is simple. The assistant's instructions can be customised to meet the needs of the user.
Alexa, Siri, and other voice assistants use speech recognition. There is a Speech Recognition API in Python that allows us to convert speech to text. Making my own helper was a fascinating task. It got less difficult to convey messages .emails

without typing any word, Searching on Google without opening the browser, and performing many other daily tasks like playing music, opening your favorite IDE with the help of a single voice command. In the current scenario, advancement in technologies are such that they can perform any task with same effectiveness or can say more effectively than us. By making this project, I realized that the concept of AI in every field is decreasing human effort and saving time.

Because the voice assistant use Artificial Intelligence, the results it produces are extremely accurate and efficient. The assistant can assist in reducing human effort and time spent on any activity; they do away with the concept of typing entirely and act as another person with whom we are conversing and requesting to accomplish a task. Although the assistant is similar to a human assistant, we can state that it is more effective and efficient at performing any duty. The libraries and packages used to create this helper are focused on reducing time complexities.

II. Objective

The goal of the project is to create a desktop personal assistant. Jarvis is inspired by virtual assistants such as Google Assistant and Siri for Android and iOS, respectively. It was created with the goal of providing a user-friendly interface for completing a range of operations using well-defined instructions. Users can communicate with the assistant using either voice instructions or keyboard input. Jarvis assists the end-user with day-to-day activities such as general human conversation, searching queries in google, bing, or yahoo, searching for videos, sending massages, retrieving images, live weather conditions, word meanings, searching for medicine details, healthrecommendations based on symptoms, and reminding the user about scheduled events and tasks as a personal assistant.

III. Problem Statement

We are familiar with many existing voice assistants like Alexa, Siri, Google Assistant, Cortana which uses concept of language processing, and voice recognition. They listens the command given by the user as per their requirements and performs that specific function in a very efficient and effective manner.

As these voice assistants are using Artificial Intelligence hence the result that they are providing are highly accurate and efficient. These assistants can help to reduce human effort and consumes time while performing any task, they removed the concept of typing completely and behave as another individual to whom we are talking and asking to perform task. These assistants are no less than a human assistant but we can say that they are more effective and efficient to perform any task. The algorithm used to make these assistant focuses on the time complexities and reduces time.

But for using these assistants one should have an account (like Google account for Google assistant, Microsoft account for Cortana) and can use it with internet connection only because these assistants are going to work with internet connectivity. They are integrated with many devices like, phones, laptops, and speakers etc.

IV. Present System

We are familiar with many existing voice assistants like Alexa, Siri, Google Assistant, Cortana which uses concept of language processing, and voice recognition.
They listens the command given by the user as per their requirements and performs that specific function in a very efficient and effective manner.

As these voice assistants are using Artificial Intelligence hence the result that
they are providing are highly accurate and efficient. These assistants can help to reduce human effort and consumes time while performing any task, they removed the concept of typing completely and behave as another individual to whom we are talking and

asking to perform task. These assistants are no less than a human assistant but we can say that they are more effective and efficient to perform any task. The algorithm used to make these assistant focuses on the time complexities and reduces time.

But for using these assistants one should have an account (like Google account for Google assistant, Microsoft account for Cortana) and can use it with internet connection only because these assistants are going to work with internet connectivity.
They are integrated with many devices like, phones, laptops, and speakers etc

V. Proposed System

Creating my own personal helper was a fun project. It became easy to send emails without typing a single word, search on Google without opening the browser, and execute many more daily chores with a single voice command, such as playing music or launching your favourite IDE. Jarvis differs from other standard voice assistants in that it is only for desktop usage, and the user does not need to create an account to use it. It also does not require an internet connection to receive instructions to complete any activity.

The IDE used in this project is PyCharm. PyCharm was used to create all of the python files as well as install all of the necessary packages. This project made use of pyttsx3, SpeechRecognition, Datetime, Wikipedia, Smtplib, pywhatkit, pyjokes, pyPDF2, pyautogui, pyQt, and other modules and libraries. I created a live GUI for communicating with JARVIS since it adds more design and intrigue to the dialogue.

JARVIS, because to technological advancements, can now execute any task as well as, if not better than, us. I noticed while working on this project that the concept of AI in every industry reduces human effort and saves time. This project's features include the ability to send emails, view PDF files, and transmit text via WhatsApp, as well as the ability to open command prompt, your choice IDE, and notepad. It can play music, perform Wikipedia searches for you, and more. It can open Google, YouTube, and other websites on a web browser. It may display the weather forecast and desktop reminders of your choosing. It's possible to have a basic discussion with it.

VI. Literature Survey

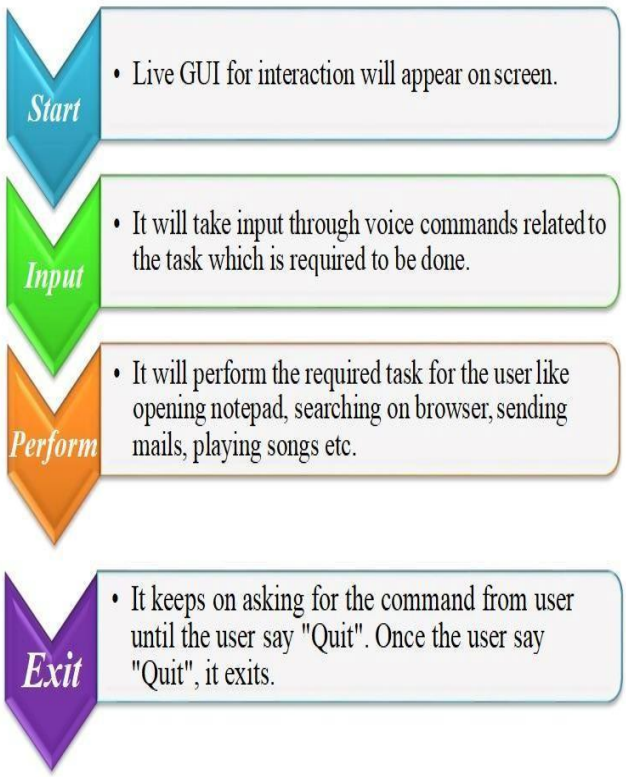
[1] The Most well known utilization of iPhone is “SIRI” which causes the end client to impart end client versatile with voice and it additionally reacts to the voice charges of the client. It is named as Personal Assistant with Voice Recognition Intelligence, which takes the client contribution to type of voice or content and process it and returns the yield in different structures like activity to be performed or the item is directed to theend client. Furthermore, this proposed framework can change the method for com-munications between end client and the cell phones.

[2] Siri: Siri has been an integral part of iOS since the dispatch of iOS5 of every 2011. It began with the nuts and bolts, for example, climate and informing ,yet has extended significantly from that point forward to help all the more outsider mix with MacOS. While Siri’s jokes are unbelievable, the virtual aide is getting more able consistently. Presently, you can request that it call individuals, send messages, plangatherings, dispatch applications and recreations, and play music, answer questions, set updates, and give climate conjectures.

[3] Alexa: While sharing different features similarlyas various VAs, Alexa is in its own one of a kind class. Amazon’s voice partner isn’tcentred on portable or PC purposes, but instead for the independent Amazon Echo speaker and a set number of Amazon Fire gadgets, with a more prominent focus on entire house administration and administrations as opposed to PC situated errands.

[4] Similarly as with any developing innovation, in any case, it can be hard to isolatethe build up from the certainties. There are four noteworthy players seeking consid-eration: Amazon (Alexa), Apple (Siri), (Google Assistant) and Microsoft (Cortana).I invested hours testing each of the four assistants by making inquiries and giving charges that numerous business clients would utilize.

[5] While it may seem like Siri, Cortana, and the mysterious Google Assistant are in general just assortments of the same virtual partners, they each have their own specific uncon-ventionality's, imperfections, and characteristics. So which one's best for clients? All things considered, that isn't a basic request to answer, as they're like the point that it'shard to take a gander at them without plunging significant into their capacities. Along these lines, we should start on this virtual right hand connection



Data Flow Diagram

VII. Requirements Specification

- **Hardware Requirements**

- Laptop

- **Software Requirements**

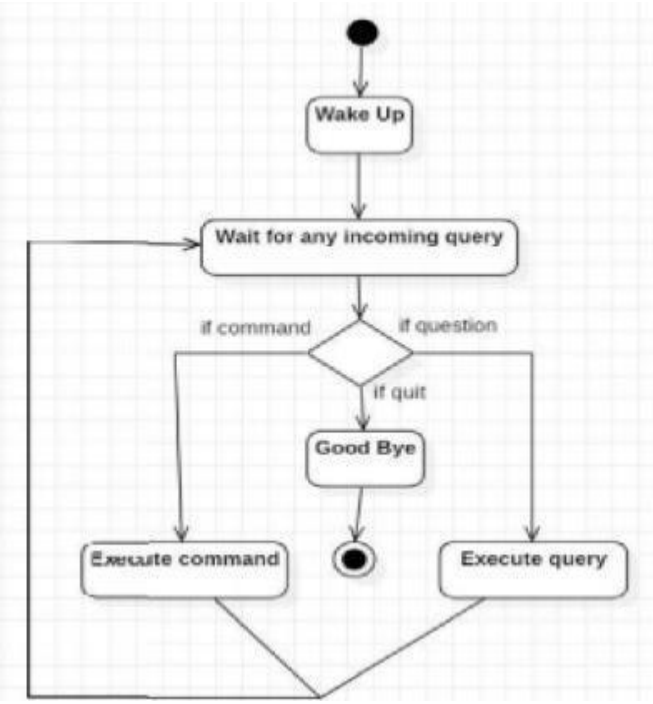
- Window 7 or above
- Python
- Visual Studio
- QT Designer

VIII. System Design

Data Flow Diagram

The system is built utilising the Artificial Intelligence concept and Python libraries. To do the tasks, Python includes a number of libraries and packages, such as pyPDF2, which may be used to read PDF files. These packages are described in full in Chapter 3 of this study.
The data in this project consists entirely of human input; whatever the user says, the assistant will carry out the task. The user input is simply a list of tasks that the user wants completed in human language, i.e. English.

Activity Diagram



Activity Diagram

The system is initially in standby mode. It begins execution as soon as it receives a wake-up call. It is determined whether the received command is a questionnaire or a job to be completed. As a result, specific actions are taken. The system waits for another command after the Question is answered or the operation is completed. Unless it receives a quit instruction, this loop will continue. It goes back to sleep at that point.

IX. Implementation of System

JARVIS, a desktop assistant, is a voice assistant that can execute numerous common desktop operations with a single voice command, such as playing music and opening your favourite IDE. Jarvis differs from other standard voice assistants in that it is only for the desktop, and the user does not need to create an account to use it. It also does not require an online connection to receive instructions to complete any activity.

Real Life Application

Saves time: JARVIS is a desktop voice assistant that responds to speech commands. It can conduct voice searches, control voice-activated devices, and assist us in completing a list of activities.

Conversational interaction: It makes it easy to perform any activity because it does so automatically by utilising Python's core modules and libraries in a conversational manner. As a result, when a user gives it a task to do, they feel as if they are giving the assignment to a human assistant due to the conversational engagement for providing input and receiving the desired output in the form of a completed task.

Reactive nature: The desktop assistant is reactive, meaning it understands human language and the context provided by the user, and responds in the same manner, i.e. in human comprehensible language, English. As a result, the user makes an informed and intelligent decision**.**

Multitasking: Its key use case could be its multitasking ability. It can keep asking for instructions until the user tells it to "QUIT."

No Trigger phase: It requests instructions and listens to the user's response without requiring a trigger phase, and then only conducts the task.

Data Implementation And Program Execution

As the first step, install all the necessary packages and libraries. The command used to install the libraries is “pip install” and then
import it. The necessary packages included are as follow

LIBRARIES AND PACKAGES

pyttsx3: It is a python library which converts text to speech.

SpeechRecognition: It is a python module which converts speech to text.

pywhatkit: It is python library to send WhatsApp message at a particular time with some additional features.
Datetime: This library provides us the actual date and time.

Wikipedia: It is a python module for searching anything on Wikipedia.

Smtplib: Simple mail transfer protocol that allows us to send mails and to route mails between mail servers.

pyPDF2: It is a python module which can read, split, merge any PDF.

Pyjokes: It is a python libraries which contains lots of interesting jokes in it.

Webbrowser: It provides interface for displaying web-based documents to users.

Pyautogui: It is a python librariy for graphical user interface.

os: It represents Operating System related functionality.

sys: It allows operating on the interpreter as it provides access to the variables and functions that usually interact strongly with the interpreter.

Functions

takeCommand(): The function is used to take the command as input through microphone of user and returns the output as string.

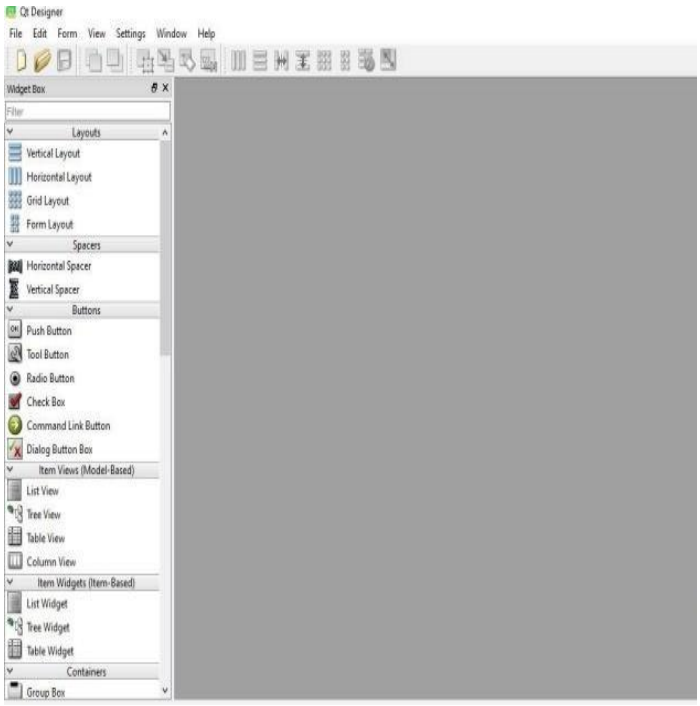
wishMe(): This function greets the user according to the time like Good Morning, Good Afternoon and Good Evening.

taskExecution(): This is the function which contains all the necessary task execution definition like sendEmail(), pdf_reader(), news() and many conditions in if condition like “open google”, “open notepad”, “search on Wikipedia” ,”play music” and “open command prompt”
etc.

VS Code

PYQT5 For Live GUI

PyQt5 is the most important python binding. It contains set of GUI widgets. PyQt5 has some important python modules like QTWidgets, QtCore, QtGui, and QtDesigner etc.



Functionality

This is where we test the system's functionality to see if it fulfils the purpose it was designed to do. Each function was checked and ran to ensure functionality; if it was able to complete the needed task correctly, the system passed that functionality test. To see if JARVIS can search on Google, for example, the user stated "Open Google," then Jarvis asked, "What should I search on Google?" Finally, the user answered, "What is Python," and Jarvis opened Google and searched for the appropriate input.

Usability

The usability of a system is determined by assessing the software's ease of use, as well as how it responds to each query posed by the user. It makes it easy to complete any activity because it does so automatically using Python's essential modules and libraries in a conversational manner. As a result, when a user gives it a task, they feel as if they are handing it to a human assistant because of the conversational engagement for providing input and receiving the desired output in the form of a completed task.

The desktop assistant is reactive, meaning it understands human language and the context provided by the user, and responds in the same manner, i.e. in human comprehensible language, English. As a result, the user makes an informed and intelligent decision. Its key use case could be its multitasking ability. It can keep asking for instructions until the user tells it to "QUIT." It requests instructions and listens to the user's response without requiring a trigger phase, and then only conducts the task.

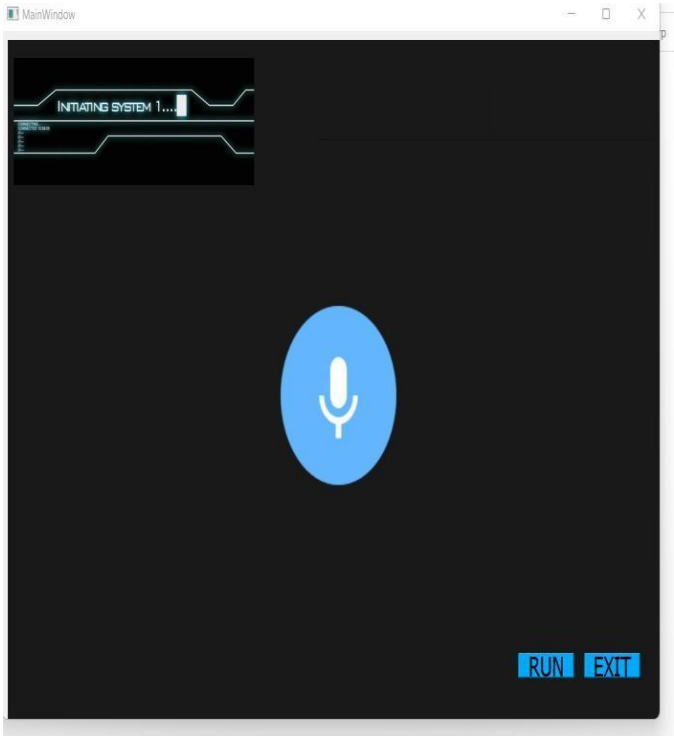
Security

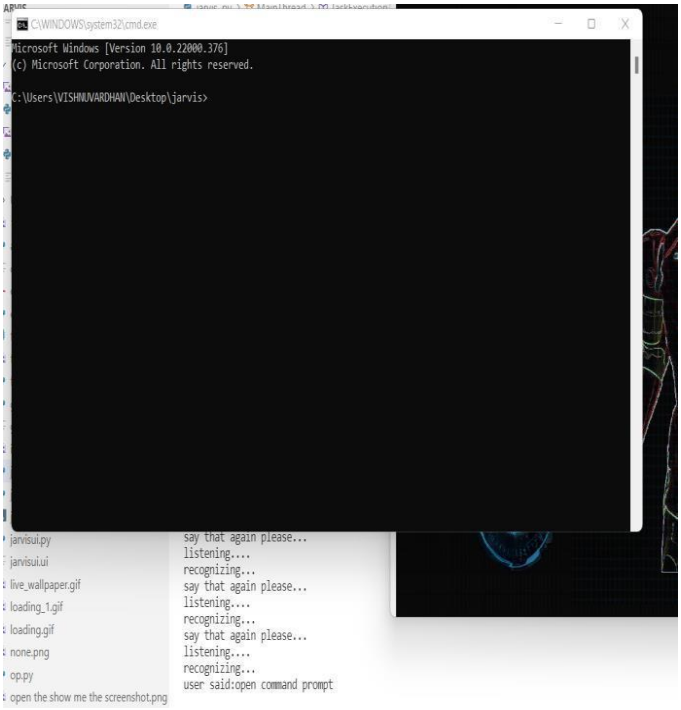
Vulnerabilities and dangers are the major emphasis of security testing. There is no risk of data breach through remote access because JARVIS is a local desktop application. Because the software is tied to a certain system, it will be launched when the user checks in.

X. Results and Discussion

User Interface

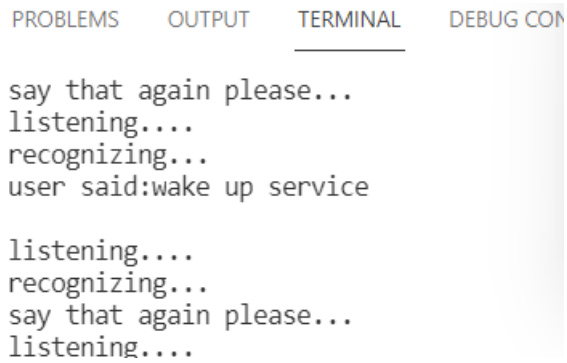
Main UI:





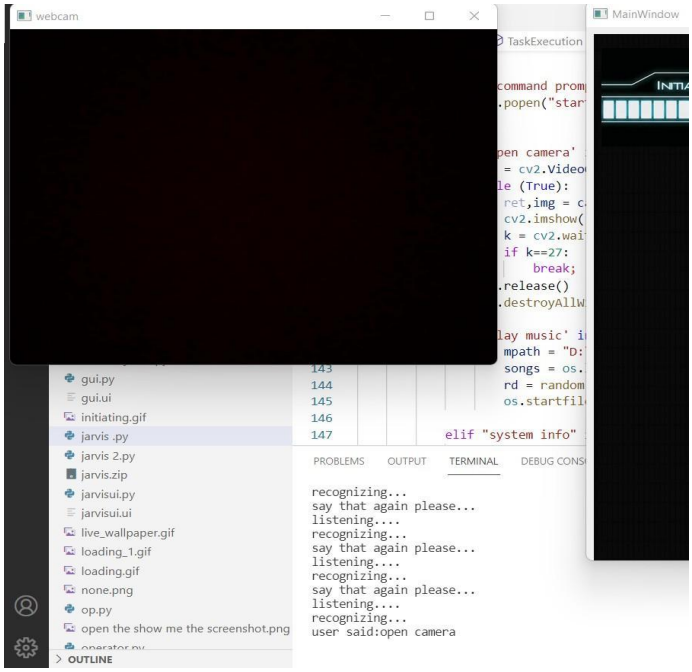
Opening notepad:

Recognizing user commands:



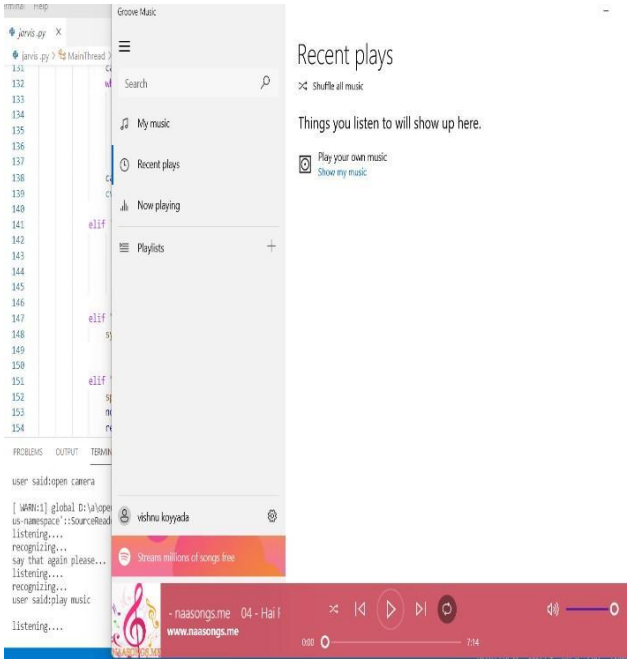
Opening commandprompt:

Opening webcam:

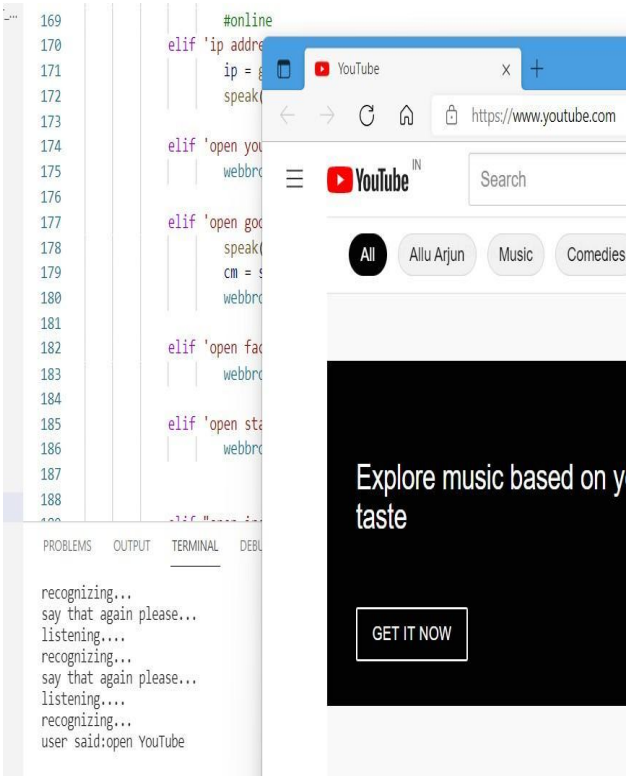



```
recognizing...
user said:play music

listening....
recognizing...
user said:system info
listening....
recognizing...
user said:today is holiday

listening....
recognizing...
user said:do you have anything

listening....
recognizing...
say that again please...
```

Playing music:

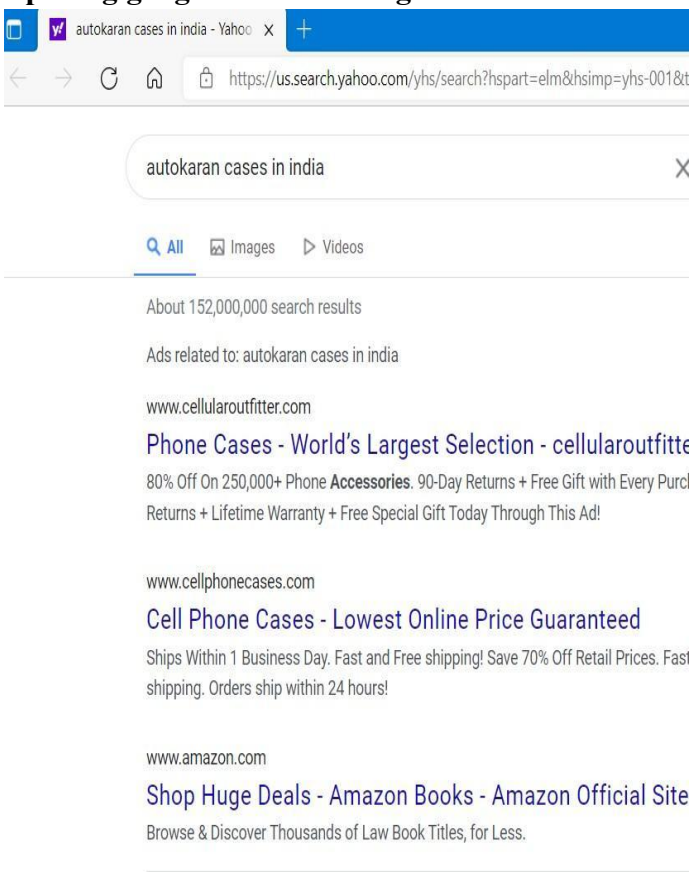


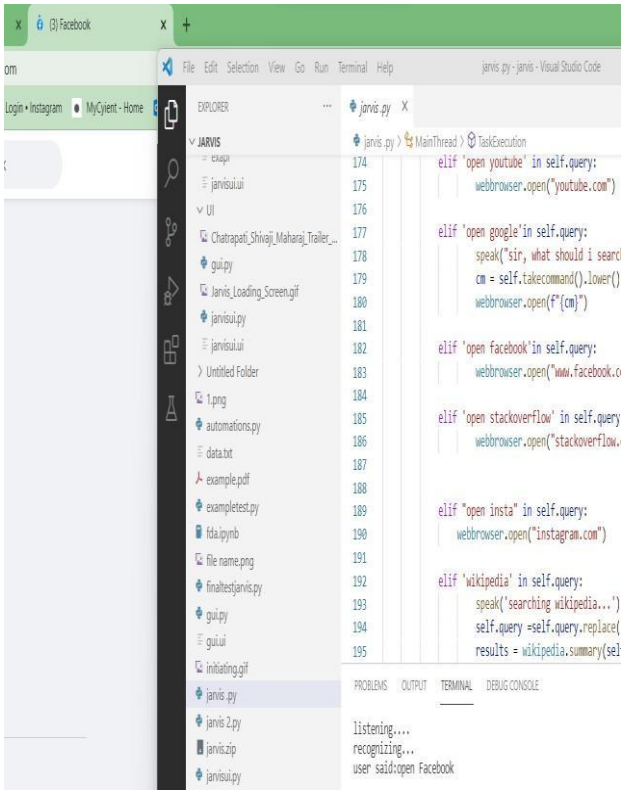
Opening youtube:



Saying to remember the text:

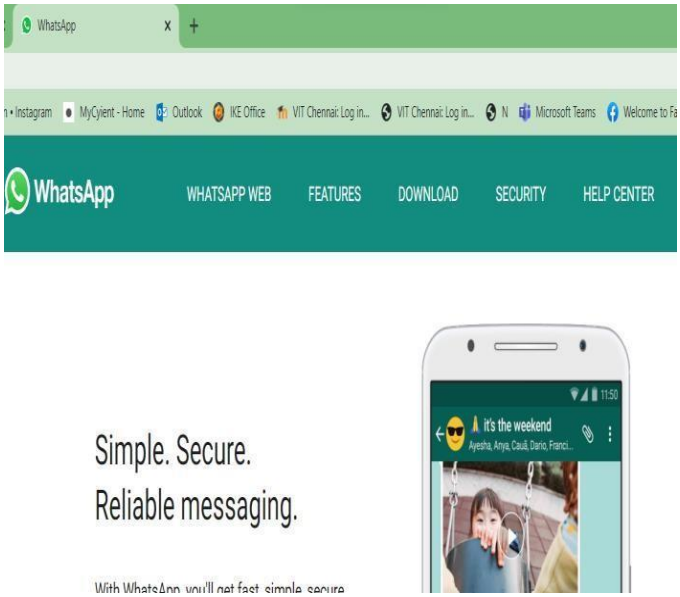
Opening google and searching:



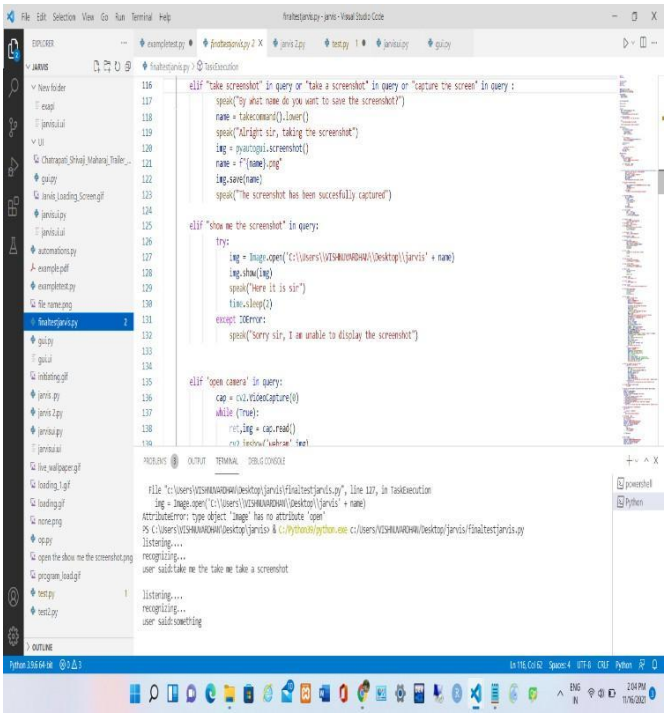


Internet speed:

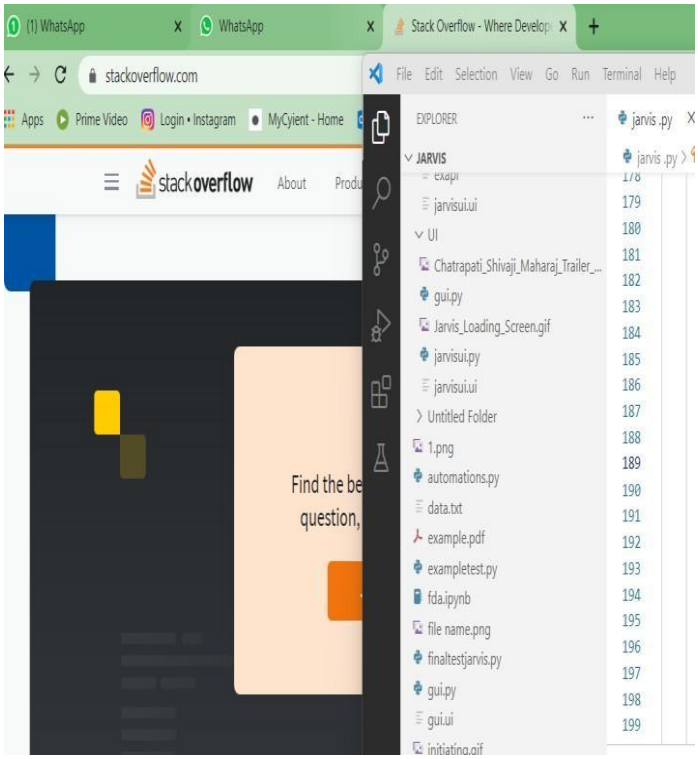
Opening webwhatsapp:



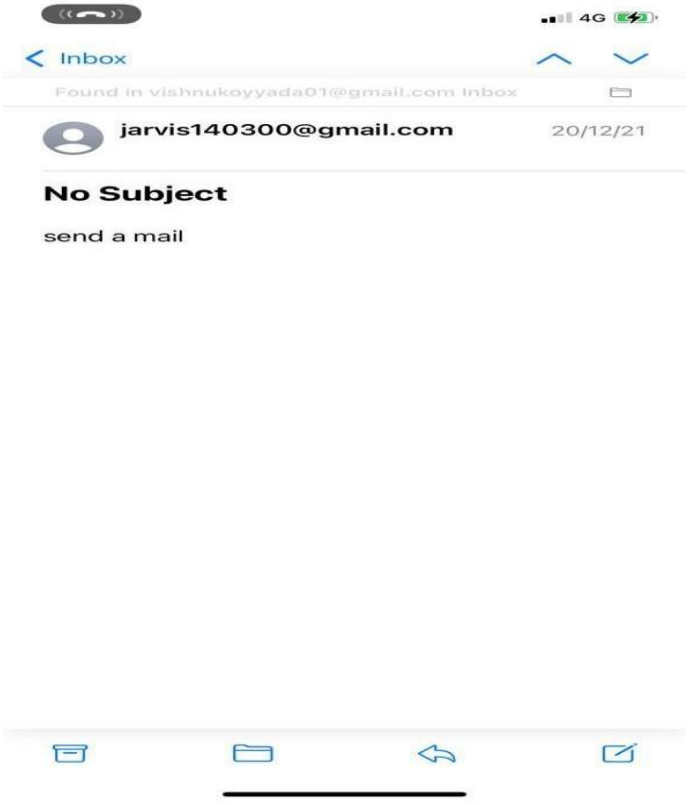
You can save screenshots:



Opening stackoverflow:



You can send mails:



XI. Conclusion and Future Work

Without a doubt, JARVIS is a very useful voice assistant because it saves the user's time through conversational engagements, as well as its effectiveness and efficiency. However, while working on this project, several limitations were discovered, as well as potential future upgrade opportunities, which are listed below.

Limitations

- Security is somewhere an issue, there is no voice command encryption in this project.
- Background voice can interfere
- Misinterpretation because of accents and may cause inaccurate results.
- JARVIS cannot be called externally anytime like other traditional assistants like Google Assistant can be called just by saying, “Ok Google!”

Scope For Future Work

- Make JARVIS to learn more on its own and develop a new skill in it.
- JARVIS android app can also be developed.
- Make more Jarvis voice terminals.
- Voice commands can be encrypted to maintain security.

XII. References

Websites referred

- www.stackoverflow.com
- www.pythonprogramming.net
- www.codecademy.com
- www.tutorialspoint.com
- www.google.co.in

- Designing Personal Assistant Software for Task Management using Semantic Web Technologies and Knowledge Databases, PurushothamBotla

- Python code for Artificial Intelligence: Foundations of Computational Agents, David L. Poole and Alan K. Mackwort