**Prototype-1**

Submitted in partial fulfilment of the requirements of the degree

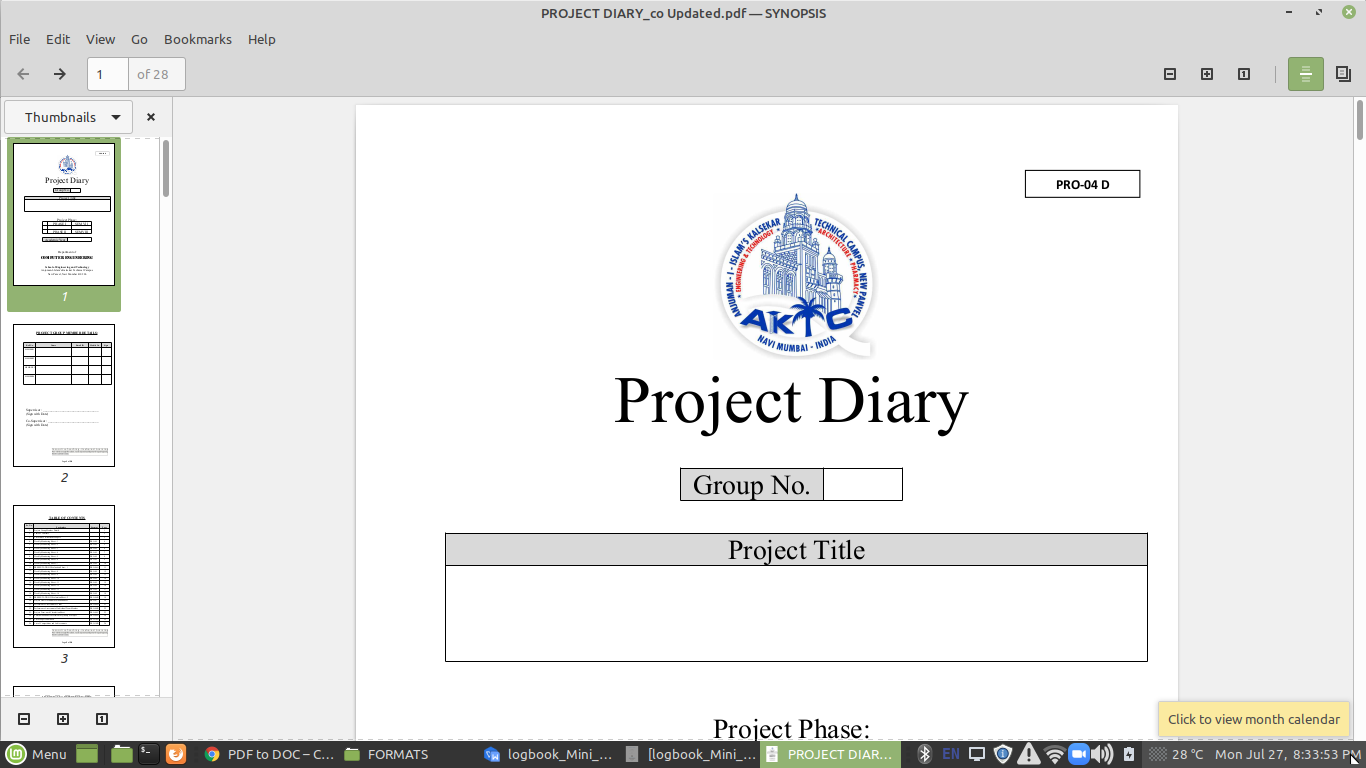
**BACHELOR OF ENGINEERING IN COMPUTER ENGINEERING**

By

**SHAIKH MOHD RAZA 20CO53**

Supervisor

**Prof. Samreen Banu Kazi**



Department of Computer Engineering

**Anjuman-I-Islam's Kalsekar Technical Campus**

**School of Engineering,**

Plot No. 2 & 3, Sector - 16, Near Thana Naka,

Khandagao, New Panvel, Navi Mumbai, Maharashtra 410206

**(AY 2021-22)**

**CERTIFICATE**

This is to certify that the Mini Project entitled “***Prototype-1***” is a Bonafede work of **Shaikh Mohd Raza 20CO53** submitted to the University of Mumbai in partial fulfilment of the requirement for the award of the degree of **“Bachelor of Engineering”** in **“Computer Engineering”** .

(**Prof. Samreen Banu Kazi**)

Supervisor

(**Prof. Tabrez Khan**) (**Dr. Abdul Razzak Honnutagi)**

Head of Department Director

**Mini Project Approval**

This Mini Project entitled “***Prototype-1***” by **Shaikh Mohd Raza 20C053** is approved for the degree of Bachelor of Engineering in Computer Engineering.

Examiners

1.............................................

(Internal Examiner Name & Sign)

2................................................

(External Examiner name & Sign)

Date:

Place:

**Contents**

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Abstract | ii |
|  |  | Acknowledgments | iii |
|  |  | List of Abbreviations | iv |
|  |  | List of Figures | v |
|  |  | List of Tables | vi |
|  |  | List of Symbols | vii |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
| 1 |  | Introduction |  |
|  | 1.1 | Introduction | 1 |
|  | 1.2 | Motivation | 2 |
|  | 1.3 | Problem Statement & Objectives |  |
|  | 1.4 | Organization of the Report |  |
|  |  |  |  |
| 2 |  | Literature Survey |  |
|  | 2.1 | Survey of Existing System |  |
|  | 2.2 | Limitation Existing system or research gap |  |
|  | 2.3 | Mini Project Contribution |  |
|  |  |  |  |
| 3 |  | Proposed System (eg New Approach of Data Summarization ) |  |
|  | 3.1 | Introduction |  |
|  | 3.2 | Algorithm and Process Design |  |
|  | 3.3 | Details of Hardware & Software |  |
|  | 3.4 | Experiment and Results |  |
|  | 3.5 | Conclusion and Future work. |  |
|  |  |  |  |
|  |  |  |  |
|  |  | References |  |

|  |  |
| --- | --- |
|  | **Abstract** |
|  |  |

This is a Voice Assistant using Python programming language. We can use this as a small project to boost their programming skills and understanding logic. Voice assistants are software agents that can interpret human speech and respond via synthesized voices. Apple’s Siri, Amazon’s Alexa, Microsoft’s Cortana, and Google’s Assistant are the most popular voice assistants and are embedded in smartphones or dedicated home speakers. Users can ask their assistants questions, control home automation devices and media playback via voice, and manage other basic tasks such as email, to-do lists, and calendars with verbal commands. This column will explore the basic workings and common features of today’s voice assistants. It will also discuss some of the privacy and security issues inherent to voice assistants and some potential future uses for these devices. As voice assistants become more widely used, librarians will want to be familiar with their operation and perhaps consider them as a means to deliver library services and materials.

# ACKNOWLEDGEMENT

I would like to take the opportunity to express my sincere thanks to my guide

**Prof. Samreen Banu Kazi ,**Assistant Professor, Department of Computer Engineering, AIKTC,

School of Engineering, Panvel for his invaluable support and guidance throughout

my project research work. Without his kind guidance & support this was not possi-

ble.

I am grateful to him/her for his timely feedback which helped me track and schedule

the process effectively. His/her time, ideas and encouragement that he gave is help

me to complete my project efficiently.

We would like to express deepest appreciation towards **DR. ABDUL RAZAK**

**HONNUTAGI**, Director, AIKTC, Navi Mumbai, **Prof. TABREZ KHAN**, Head of

Department of Computer Engineering and **Prof. JAVED SHAIKH**, Project Coordi-

nator whose invaluable guidance supported us in completing this project.

At last we must express our sincere heartfelt gratitude to all the staff members of

Computer Engineering Department who helped me directly or indirectly during this

course of work

**PROTOTYPE-1**

**CHAPTER 1**

**1.1Introduction**

In today's era, almost all tasks are digitalized. We have Smartphones in our hands and it is nothing less than having the world at your fingertips. These days we aren't even using fingers. We just speak of the task and it is done. There exist systems where we can say Text Dad, "I'll be late today." And the text is sent. That is the task of a Virtual Assistant. It also supports specialized tasks such as booking a flight or finding the cheapest book online from various e-commerce sites and then providing an interface to book orders are helping automate search, discovery, and online order operations.

Virtual Assistants are software programs that help you ease your day-to-day tasks, such as showing weather reports. creating reminders, making shopping lists, etc. They can take commands via text (online chatbots) or by voice. Voice-based intelligent assistants need an invoking word or wake word to activate the listener, followed by the command.

This system is designed to be used efficiently on desktops. Personal assistant software improves user productivity by managing routine tasks of the user and by providing information from online sources to the user. JARVIS is effortless to use. Call the wake word 'JARVIS' followed by the command And within seconds, it gets executed.

Voice searches have dominated text search. Web Searches conducted via mobile devices have only just overtaken those carried out using a computer and the analysts are already predicting that 50% of searches will be via voice by 2022. Virtual assistants are turning out to be smarter than ever. Allow your intelligent assistant to make email work for you Detect intent, pick out important information, automate processes, and deliver personalized responses

This project was started on the premise that there is a sufficient amount of openly available data and information on the web that can be utilized to build a virtual assistant that has access to making intelligent decisions for routine user activities

**1.2 Motivation**

Purpose of virtual assistant is to being capable of voice interaction, music playback, making to-do lists, setting alarms, streaming podcasts, playing audiobooks, and providing weather, traffic, sports, and other real-time information, such as news. Virtual assistants enable users to speak natural language voice commands in order to operate the device and its apps.

There is an increased overall awareness and a higher level of comfort demonstrated specifically by millennial consumers. In this ever-evolving digital world where speed, efficiency, and convenience are constantly being optimized, it's clear that we are moving towards less screen interaction.

Voice assistants will continue to offer more individualized experiences as they get better at differentiating between voices. However, it's not just developers that need to address the complexity of developing for voice as brands also need to understand the capabilities of each device and integration and if it makes sense for their specific brand. They will also need to focus on maintaining a user experience that is consistent within the coming years as complexity becomes more of a concern. This is because the visual interface with voice assistants is missing Users simply cannot see or touch a voice interface

The mass adoption of artificial intelligence in users' everyday lives is also fuelling the shift towards voice. The number of IoT devices such as smart thermostats and speakers are giving voice assistants more utility in a connected user's life. Smart speakers are the number one way we are seeing voice being used. Many industry experts even predict that nearly every application will integrate voice technology in some way in the next 5 years.

The use of virtual assistants can also enhance the system of IoT (Internet of Things). Twenty years from now. Microsoft and its competitors will be offering personal digital assistants that will offer the services of a full-time employee usually reserved for the rich and famous.

* 1. **Problem Statement & Objectives**

**Problem Statement:**

Usually, the user needs to manually manage multiple sets of applications to complete one task. For example, a user trying to make a travel plan needs to check for airport codes for nearby airports and then check travel sites for tickets between combinations of airports to reach the destination. There is a need for a system that can manage tasks effortlessly.

We already have multiple virtual assistants. But we hardly use it. There is a number of people who have issues with voice recognition. These systems can understand English phrases but they fail to recognize our accent. Our way of pronunciation is way distinct from theirs. Also, they are easy to use on mobile devices than on desktop systems. There is a need for a virtual assistant that can understand English in an Indian accent and work on the desktop system.

When a virtual assistant is not able to answer questions accurately, it's because it lacks the proper context or doesn't understand the intent of the question. Its ability to answer questions relevantly only happens with rigorous optimization, involving both humans and machine learning. Continuously ensuring solid quality control strategies will also help manage the risk of the virtual assistant learning undesired behaviours. They require a large amount of information to be fed in order for it to work efficiently.

A virtual assistant should be able to model complex task dependencies and use these models to recommend optimized plans for the user. It needs to be tested for finding optimum paths when a task has multiple sub-tasks and each sub-task can have its own sub-tasks. In such a case there can be multiple solutions to paths, and it should be able to consider user preferences, other active tasks, and priorities in order to recommend a particular plan.

**Objectives:**

The main objective of building personal assistant software (a virtual assistant) is using semantic data sources available on the web, user-generated content, and providing knowledge from knowledge databases. The main purpose of an intelligent virtual assistant is to answer questions that users may have. This may be done in a business environment, for example, on the business website, with a chat interface. On the mobile platform, the intelligent virtual assistant is available as a call-button operated service where a voice asks the user "What can I do for you?" and then responds to verbal input.

Virtual assistants can tremendously save you time. We spend hours in online research and then make the report in our terms of understanding. JARVIS can do that for you. Provide a topic for research and continue with your tasks while JARVIS does the research. Another difficult task is to remember test dates, birthdates, or anniversaries. It comes with a surprise when you enter the class and realize it is a class test today. Just tell JARVIS in advance about your tests and she reminds you well in advance so you can prepare for the test.

One of the main advantages of voice searches is their rapidity. In fact, voice is reputed to be four times faster than a written search: whereas we can write about 40 words per minute, we are capable of speaking around 150 during the same period of time 15. In this respect, the ability of personal assistants to accurately recognize spoken words is a prerequisite for them to be adopted by consumers.

* 1. **Organization of the Report**

The report has been organised and provided in 3 sections.

First section is all about the introduction of the proposed project.

Includes sub points as Introduction, Motivation, Problem Statement, Objective and Organisation of Report. This section will give you overall ideas related to the project concept.

Second section is based on the Literature Survey. It gives information regarding already existing projects related to our proposed idea. We have provided the title and description of three such papers. Limitation of those projects are mentioned as well with contribution of our proposed idea to overcome limitations.

Third section focuses on information related to the Proposed System.

It contains Introduction, System Architecture of project, Algorithm- explaining working of program, information of hardware and software required and at the last screenshots of result with conclusion and future work.

**CHAPTER 2**

**2 Literature Survey:**

We did a survey on the existing system before proposing this idea and we got to know that there are some limitations in those system which our proposed project can overcome. There already exist a number of desktop virtual assistants. A few examples of current virtual assistants available in market are discussed in this section along with the tasks they can provide and their drawbacks.

**2.1 Survey of Existing System:**

SIRI from Apple:

SIRI is personal assistant software that interfaces with the user thru voice Interface, recognizes commands and acts on them. It learns to adapt to user’s speech and thus improves voice recognition over time. It also tries to converse with the user when it does not identify the user request. It integrates with calendar, contacts and music library applications on the device and also integrates with GPS and camera on the device. It uses location, temporal, social and task based contexts, to personalize the agent behaviour specifically to the user at a given point of time.

Supported Tasks :

• Call someone from my contacts list

•Launch an application on my iPhone

• Send a text message to someone

• Set up a meeting on my calendar for 9am tomorrow

• Set an alarm for 5am tomorrow morning

• Play a specific song in my iTunes library

• Enter a new note

ReQall:

ReQall is personal assistant software that runs on smartphones running Apple iOS or Google Android operating system. It helps user to recall notes as well as tasks within a location and time context. It records user inputs and converts them into commands, and monitors current stack of user tasks to proactively suggest actions while considering any changes in the environment. It also presents information based on the context of the user, as well as filter information to the user based on its learned understanding of the priority of that information.

Supported Tasks:

• Reminders

• Email

• Calendar, Google Calendar

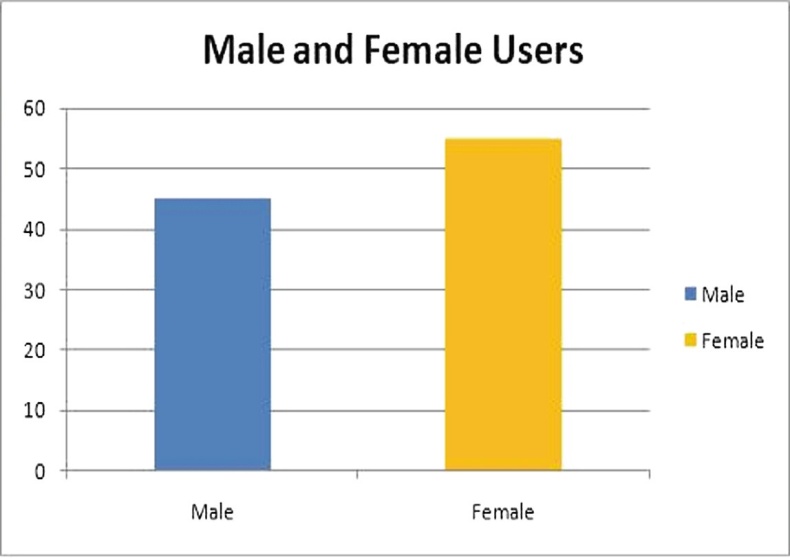
• Outlook

• Evernote

• Facebook, LinkedIn

• News Feeds

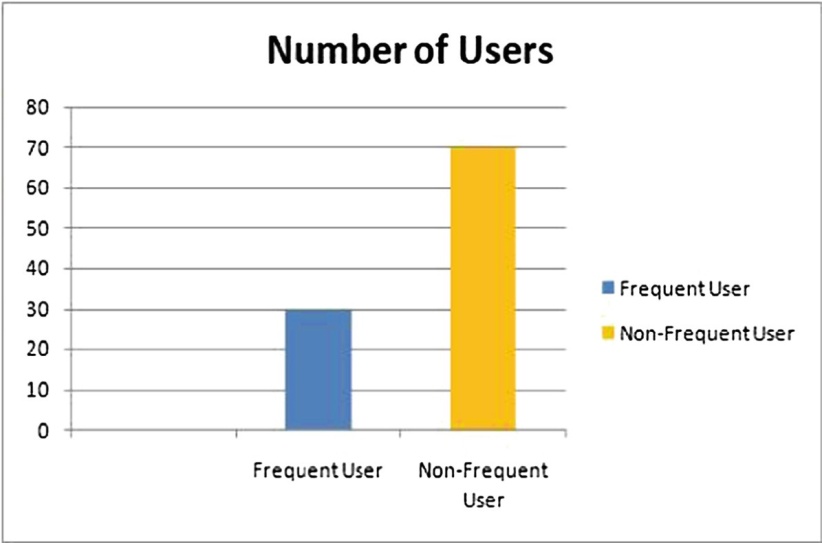
In survey we have observe that the ratio of male and female user aur 70,30 which means number of female user are more compared to female users we have also seen that the number of active user are less compared to the non active user and this is because of the features provided by the voice assistant are limited so people it doesn't attract the user



1. Number of Male and Female Users

In the second survey we have observe that the there is doubt in the users mind while using this type of assistance, that there data is safe or not it is safe or

leaked due to which the ratio of active users reduced. The user which actively use the assistant want update in their assistance like they need to use it offline mode they need various type of new features which is not provided by the assistants currently due to which the again the number of user reduce



1. Number of Users

**2.2 Limitation Existing system or research gap:**

Drawback of SIRI:

SIRI does not maintain a knowledge database of its own and its understanding comes from the information captured in domain models and data models.

Drawback of ReQall:

It will take some time to put all of the to-do items in – you could spend more time putting the entries in than actually doing the revision.

While working on our project we had one advantage in our project that is we can use a voice assistant without an internet connection but while elaborating on our project and exceeding it we have found one issue that it will work only when you are connected to the internet because prototype 1 can send emails and WhatsApp messages too and hence it requires an internet connection to fetch the details of the user to log in successfully with security as of now prototype 1 can be used only on the terminal as we have not built an application

**2.3Mini Project Contribution**

Talking about the contribution every member in our grp contributed equally. Every one joins the meeting on time and helps each other on writing the codes making the flow chart, making the ppt and making the report. Everyone did the coding together by keeping Google meet and sharing their screens and doing their parts of codes, we also made the flow chart by everyone’s contribution like one is colouring one is writing the statements one is making the shapes and one is writing the comments, in report writing everyone in the group picked the topics and completed their part, and in ppt making everyone took some slides and worked on it

**CHAPTER 3**

**Proposed System**

**3.1 Introduction:**

Talking about the proposed system we have programmed our project in such way that it will understand the human natural language and it will respond to the user with a particular result

**3.2 Algorithm and Process Design:**

Pyttsx:

Pyttsx stands for Python Text to Speech. It is a cross-platform Python wrapper for text-to-speech synthesis. It is a Python package supporting common text-to-speech engines on Mac OS X, Windows, and Linux. It works for both Python2.x and 3.x versions. Its main advantage is that it works offline.

Speech Recognition:

This is a library for performing speech recognition, with support for several engines and APIs, online and offline. It supports APIs like Google Cloud Speech API, IBM Speech to Text, Microsoft Bing Voice Recognition etc.

Different functions and attributes of Recognition used:

1. Microphone() : It is advisable to specify the microphone during the program to avoid any glitches. In our project we’ve specified it as source and our local system’s microphone will be enabled whenever it is called.

2. pause\_threshold : pause\_threshold seconds of non-speaking or there is no more audio input. The ending silence is not included. The timeout parameter is the maximum number of seconds that this will wait for a phrase to start before giving up and throwing a speech recognition.

3. energy\_threshold : It defines what audio level and above should be considered speech. Now, this default energy threshold works most of the time. If your environment is sufficiently quiet, it will be able to recognize you talking without problems.

4. recognize\_google() : Now you've created an instance of the Recognizer class we'll use the recognize\_google() method on it to access the Google web speech API and turn spoken language into text. recognize\_google() requires an argument audio\_data otherwise it will return an error. US English is the default language.

5. listen() : Used to listen to the phrase said by the user in which we pass different arguments such as phrase\_time\_limit to wait for the user to finish his sentence and also we need to define the source

After configuring the speech to text, all the speech that user says will be stored in query variable in the text format. In our project, with the help of this query we are going to perform different functions by using amazing modules provided by Python.

**Algorithm:**

• After execution of our program, Prototype-1 begins by introducing itself.

• It waits for the user commands and listens to to command, once it is done listening it begins to analyze the given command and work accordingly.

• For example: When the user gives the command to read him/her today’s news, it analyses the command and looks up for top 5 news and read it to the user.

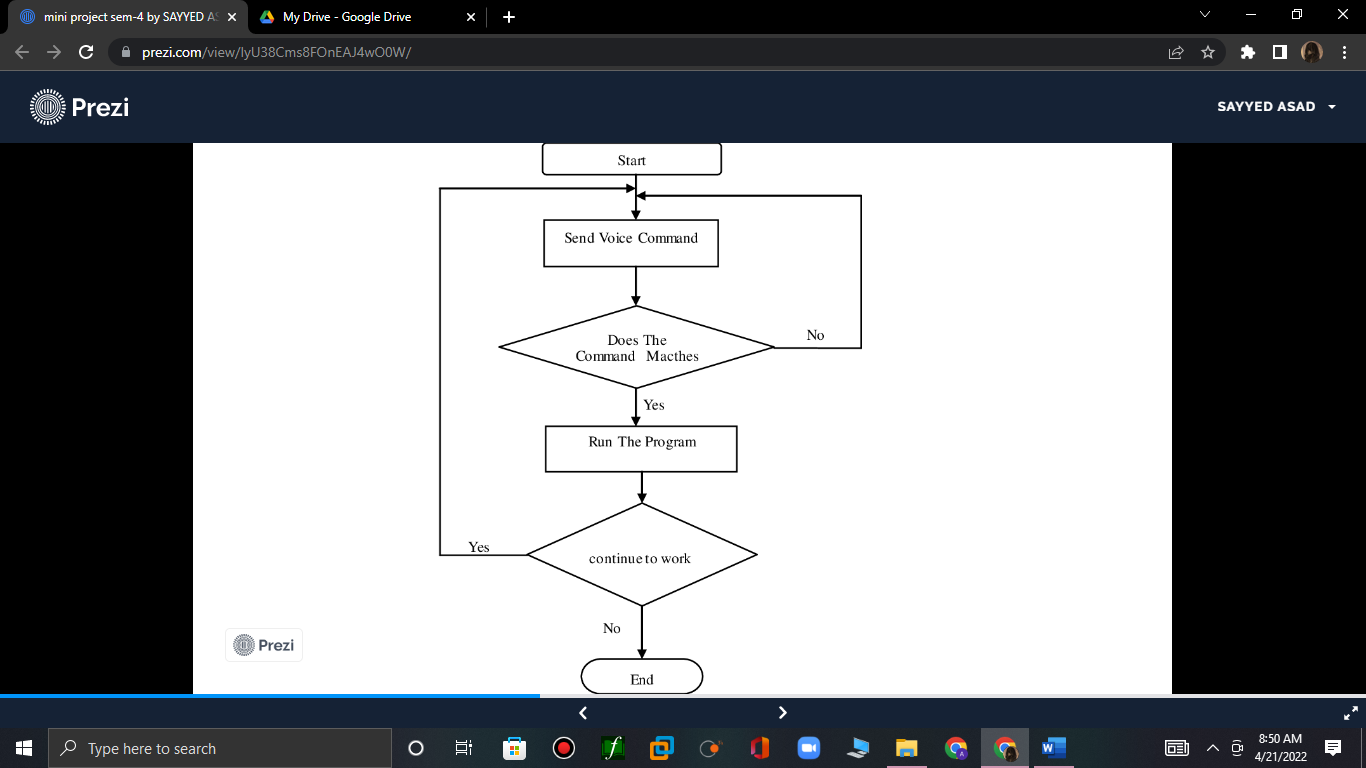
• Prototype-1 compares the commands given by the user the conditions and when the commands match up it moves forward and performs the given task.

• But for cases when the commands doesn’t match up or is not quite understood, Prototype-1 asks the user to repeat the sentence and starts listening to the user again.

• Prototype-1 does not stop listening and recognizing the task unless it has been said to sleep.

• After executing the commands and presenting the output and goes back and waits for its next instruction.

• But when the user asks it to go to sleep , it stops the program and shuts down.

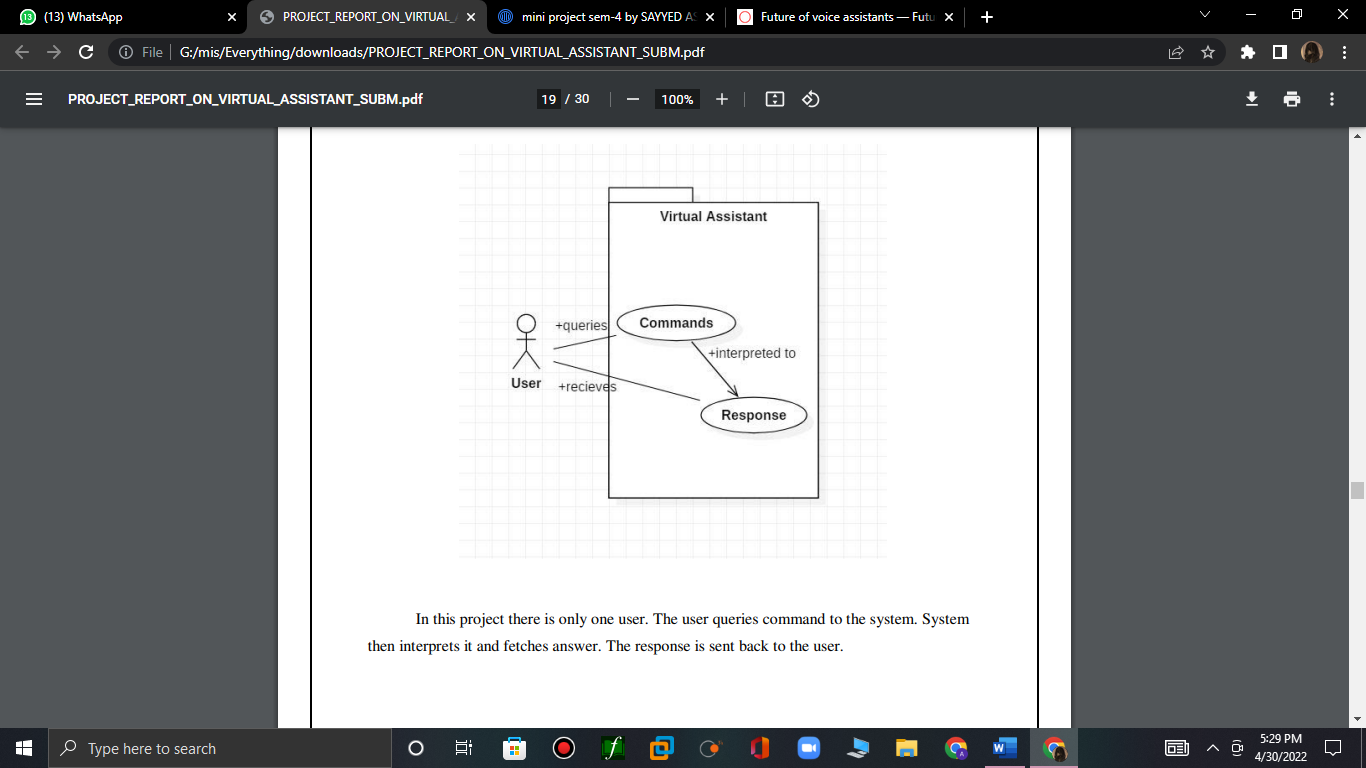
****

**Process Design:**

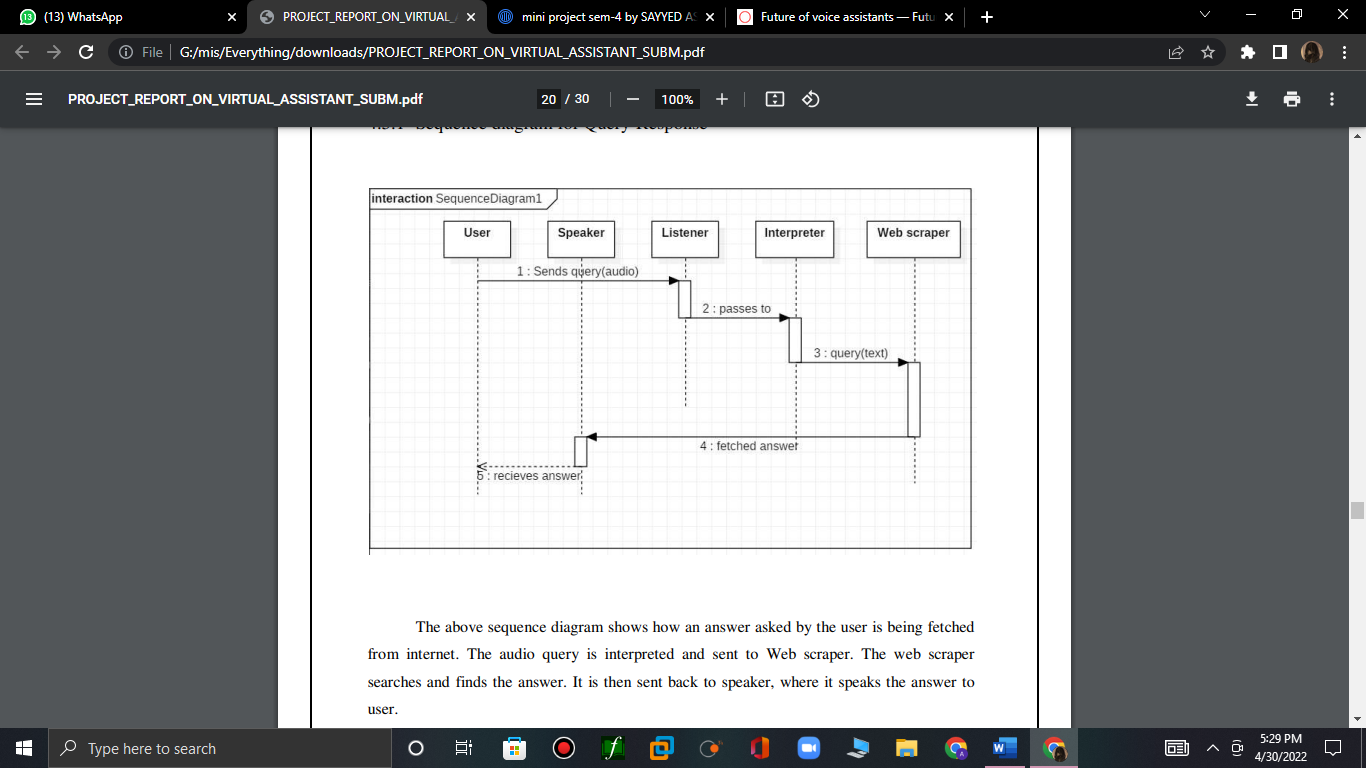
Here is the graphical re presentation of our project:

1)USE CASE DIAGRAM:

In this project there is only one user. The user queries command to the system. System then interprets it and fetches answer. The response is sent back to the user.



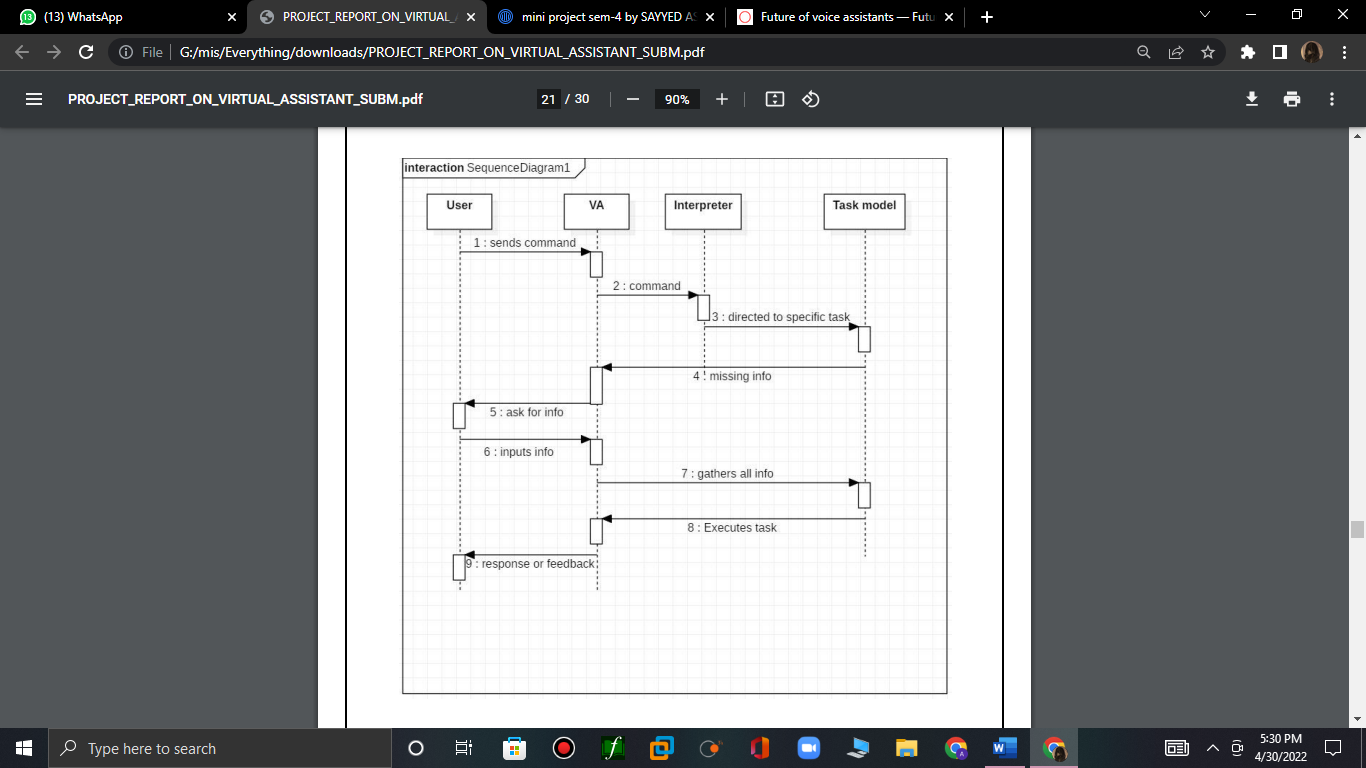
2) SEQUENCE DIAGRAM-1:



The above sequence diagram shows how an answer asked by the user is being fetched from internet. The audio query is interpreted and sent to Web scraper. The web scraper searches and finds the answer. It is then sent back to speaker, where it speaks the answer to user

3) SEQUENCE DIAGRAM-2:

The user sends command to virtual assistant in audio form. The command is passed to the interpreter. It identifies what the user has asked and directs it to task executer. If the task is missing some info, the virtual assistant asks user back about it. The received information is sent back to task and it is accomplished. After execution feedback is sent back to user.



**3.3 Details of Hardware & Software**

**Hardware:**

Device name DESKTOP-LTU6EUH

Processor Intel(R) Core(TM) i5-7200U CPU @ 2.50GHz 2.70 GHz

Installed RAM 8.00 GB

System type 64-bit operating system, x64-based processor

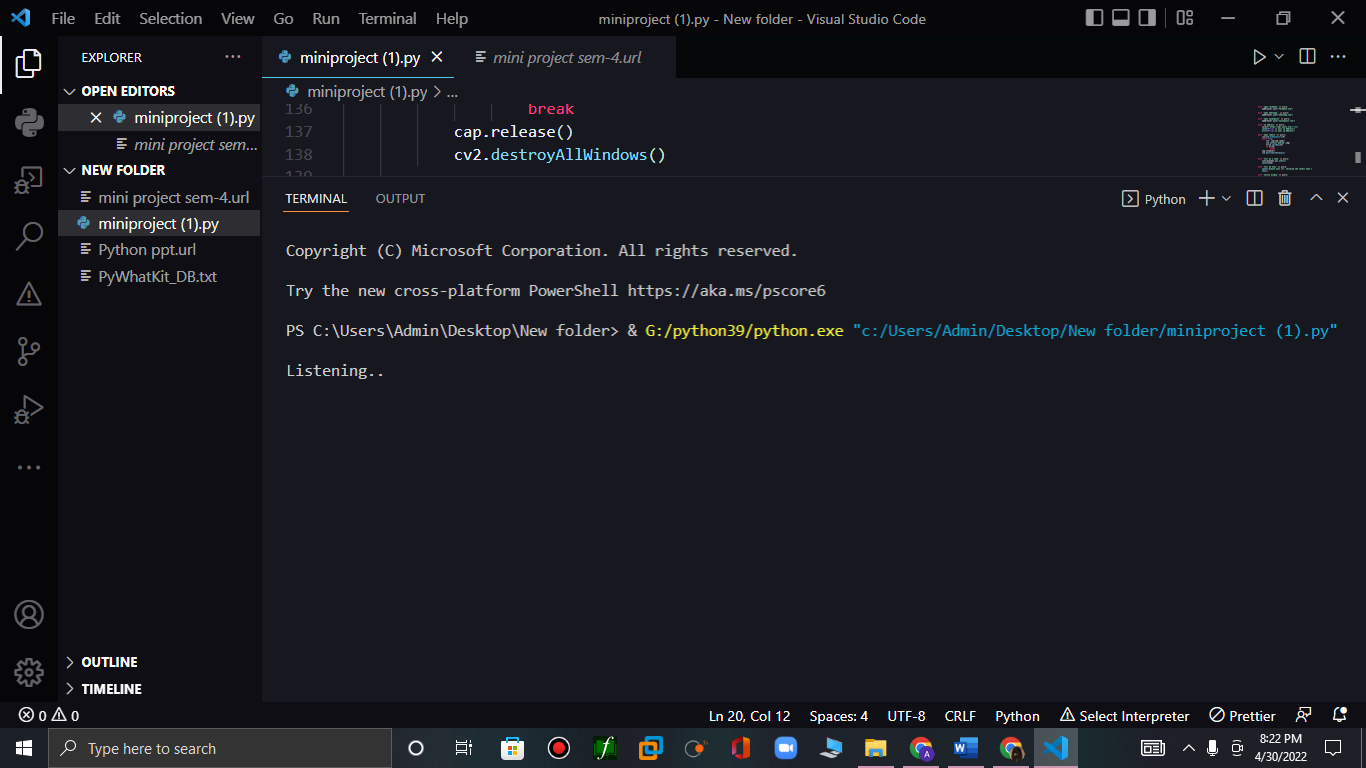
**Software:**

Python:

Python is an OOPs (Object Oriented Programming) based, high level, interpreted programming language. It is a robust, highly useful language focused on rapid application development (RAD). Python helps in easy writing and execution of codes. Python can implement the same logic with as much as 1/5th code as compared to other OOPs languages. Python provides a huge list of benefits to all. The usage of Python is such that it cannot be limited to only one activity. Its growing popularity has allowed it to enter into some of the most popular and complex processes like Artificial Intelligence (AI), Machine Learning (ML), natural language processing, data science etc. Python has a lot of libraries for every need of this project. For JIA, libraries used are speechrecognition to recognize voice, Pyttsx for text to speech, selenium for web automation etc. Python is reasonably efficient. Efficiency is usually not a problem for small examples. If your Python code is not efficient enough, a general procedure to improve it is to find out what is taking most the time, and implement just that part more efficiently in some lower-level language. This will result in much less programming and more efficient code (because you will have more time to optimize) than writing everything in a low-level language.

Visual Studio Code.

**3.4 Experiment and Result**

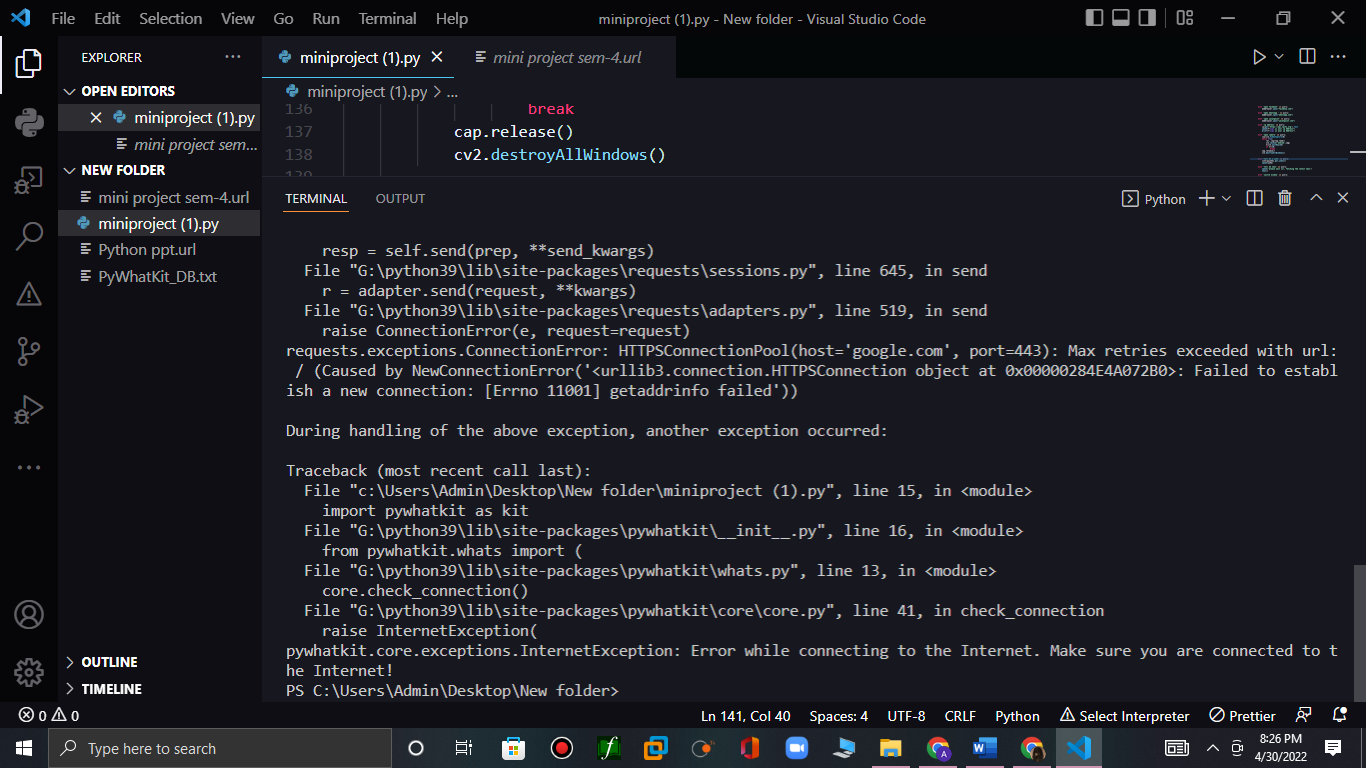
to run our program we just need to click on the run button and make sure that we have a stable internet connection 

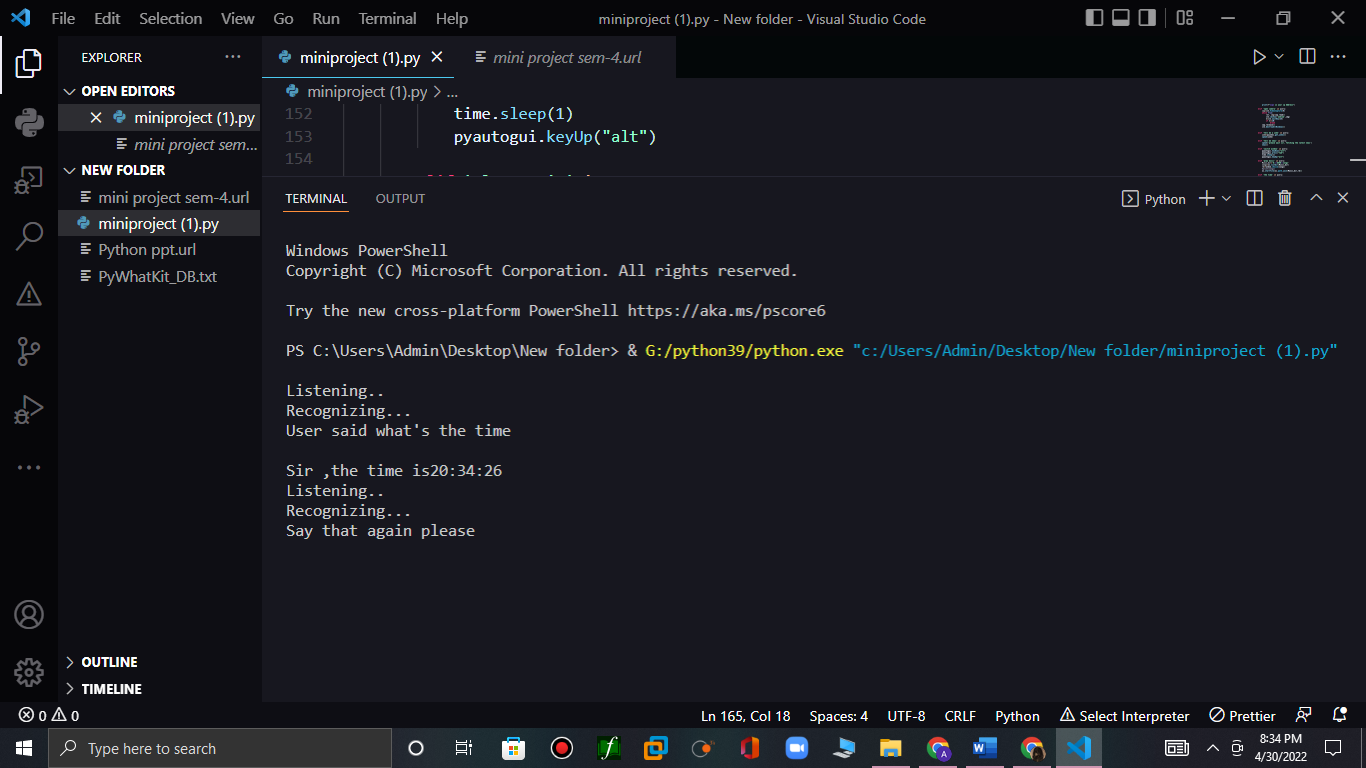
if we don’t have a internet connection then it will show this error

check\_connection

raise InternetException(

pywhatkit.core.exceptions.InternetException: Error while connecting to the Internet. Make sure you are connected to the Internet!



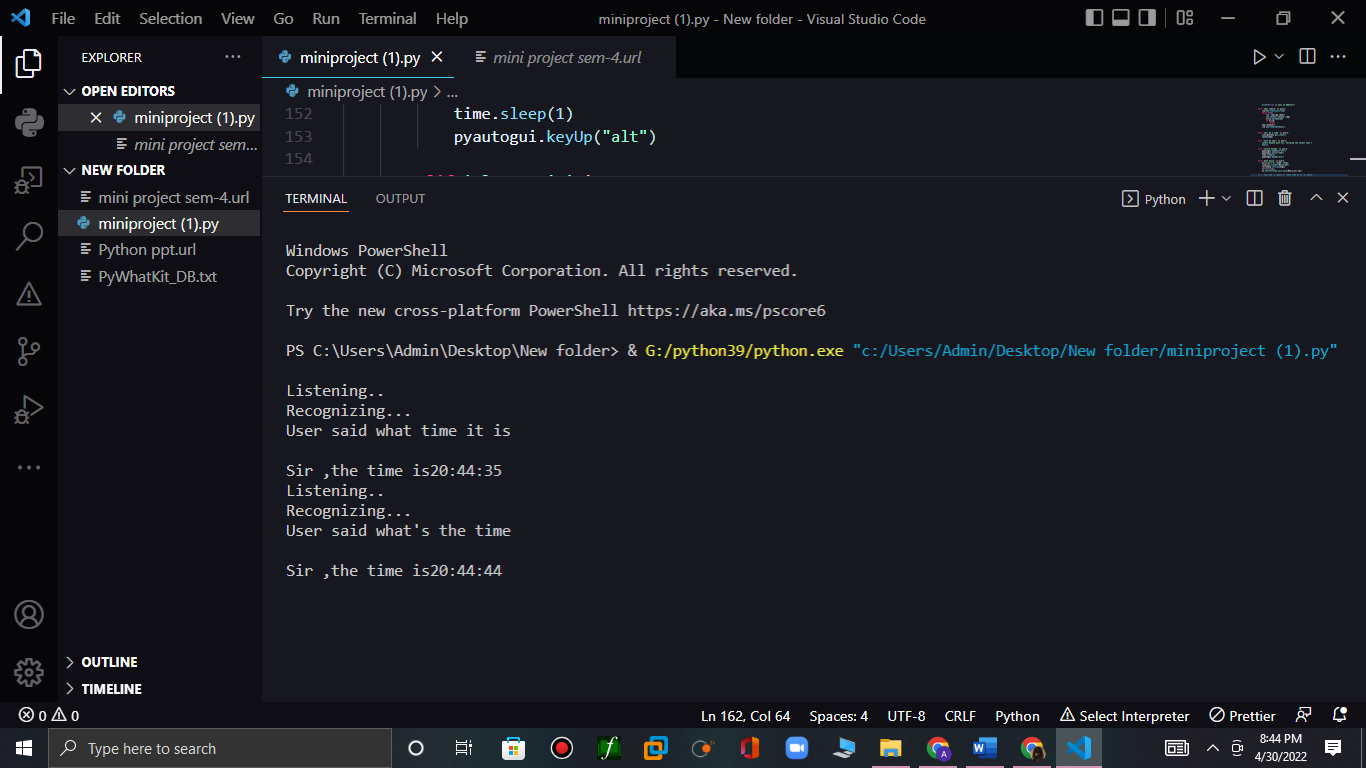
In the above experiment we have used time command when we ask our prototype time it will tell us the exact time and will print it 

Here first of all it will listen the command or query then it will process it and if the command matches then it will show the result

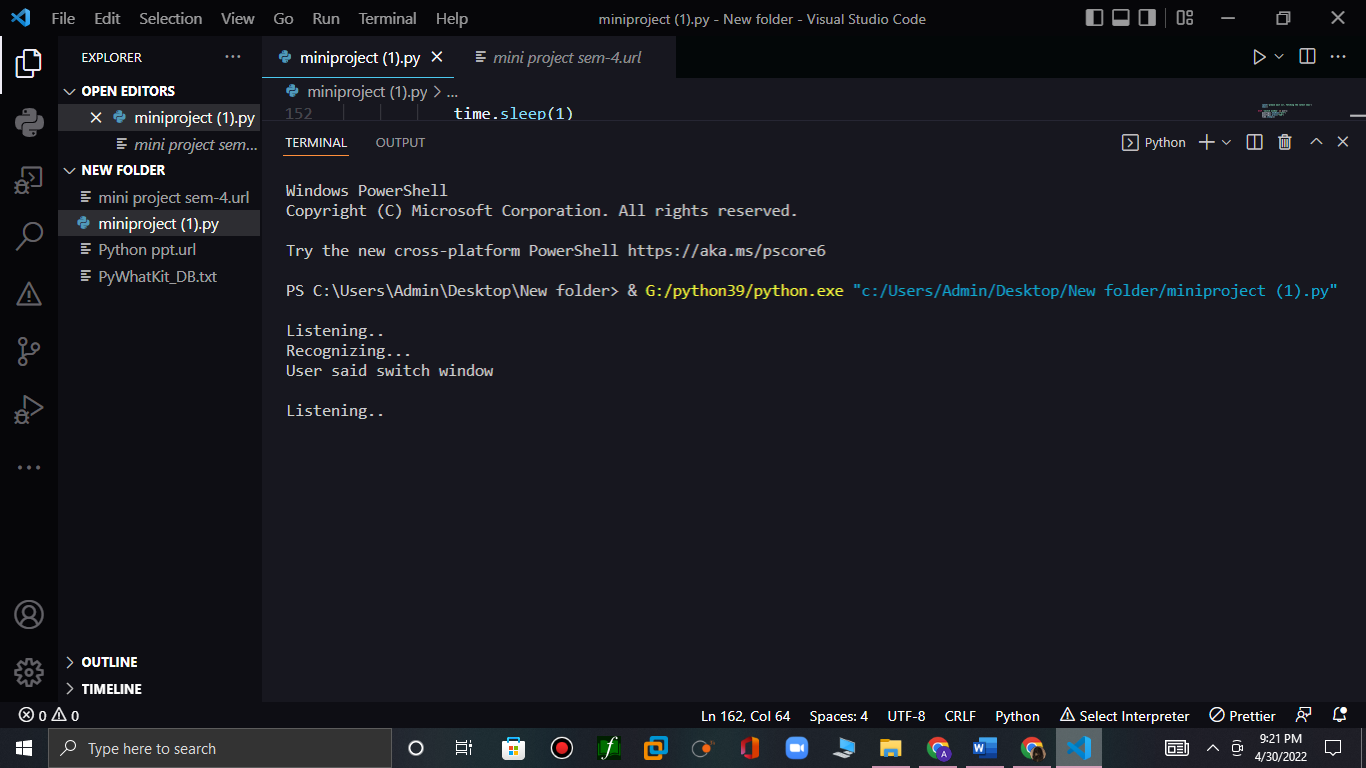
It doesn’t stops here it will take commands continuously and ask user to say any query until the user say close the program or terminate the program.

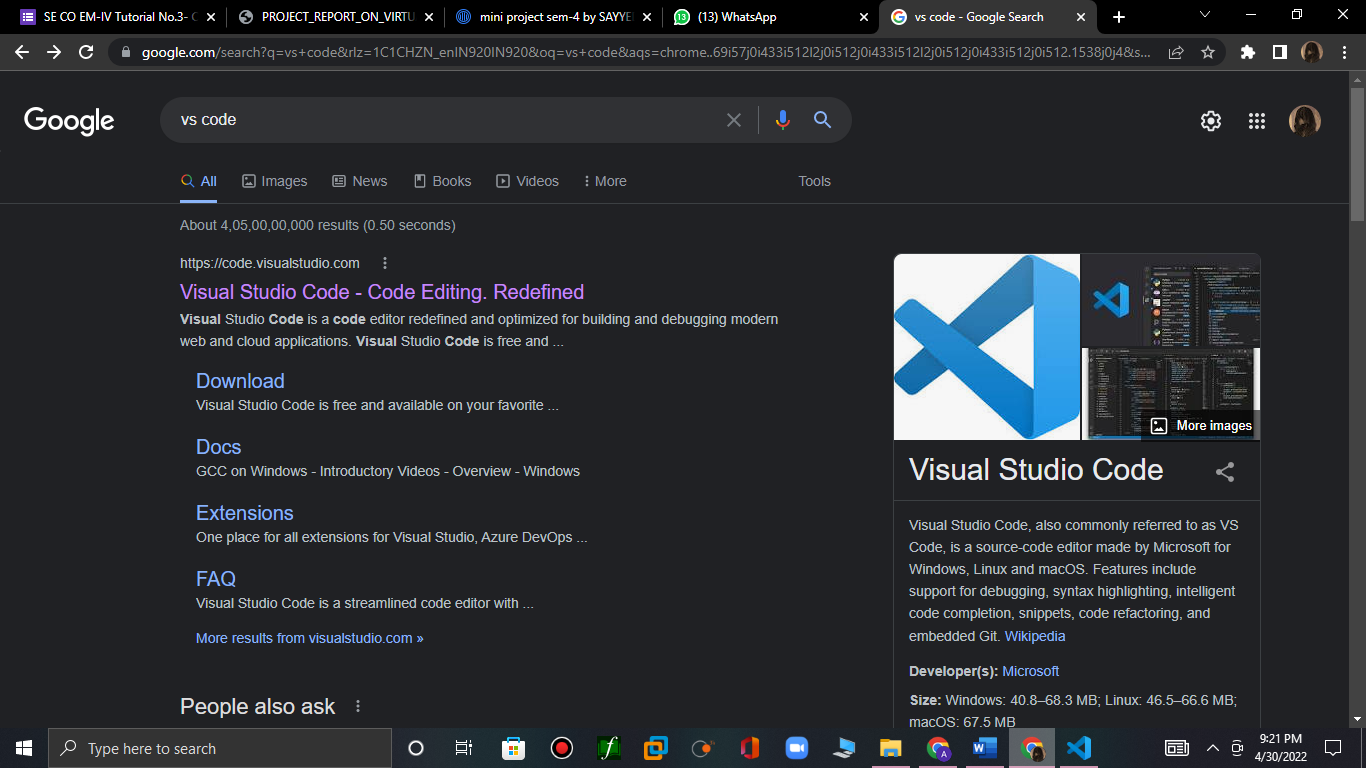
There are different peoples with different language, everyone has different accents and they can give different commands so we have programed our project in this way that it will understand every language.

Some people can say tell me the time or some can say what time it s or else both the commands are same but in our project we have used every possible commands with different way here is the another example

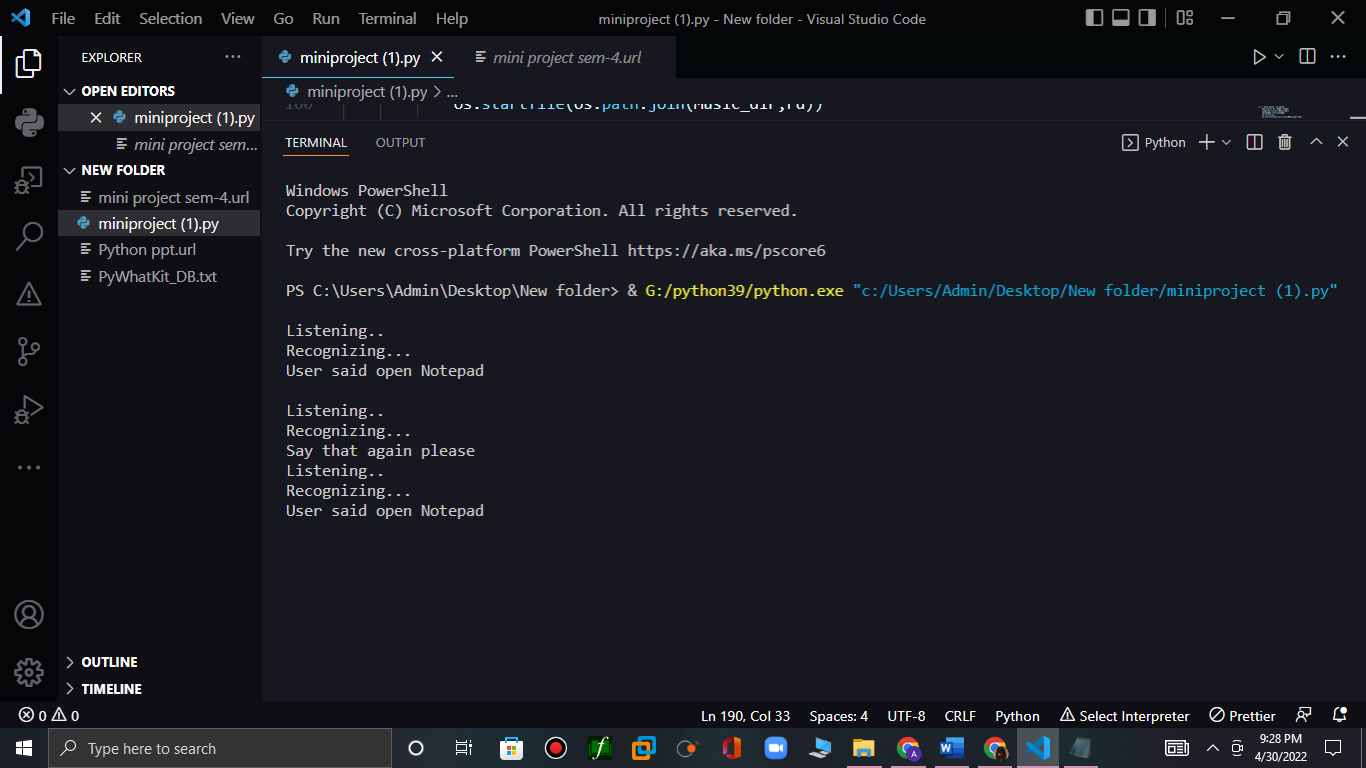


In above experiment we have used switch window command which can switch the window without touching or without any physical interaction

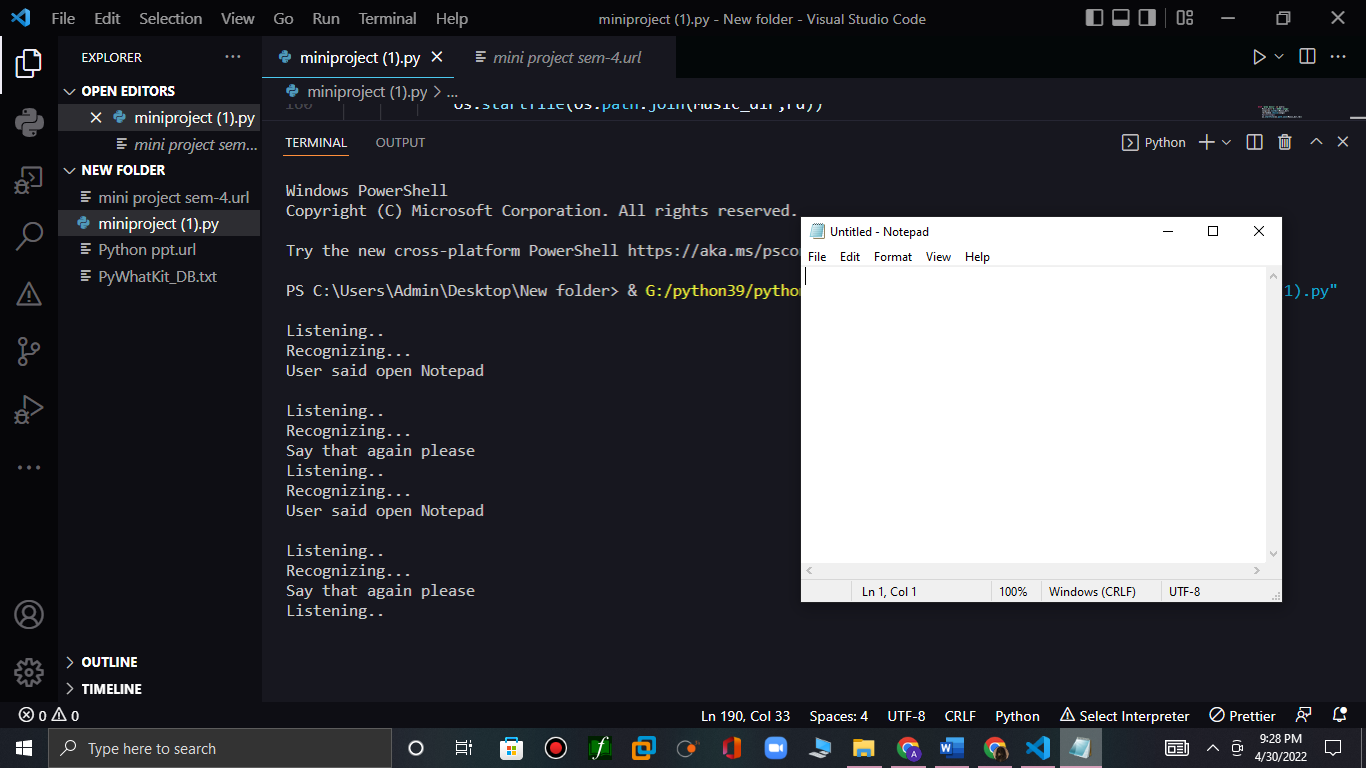




There are many commands for the following example we have used switch window command as you can see in the picture notepad is barely visible which means the notepad is opening

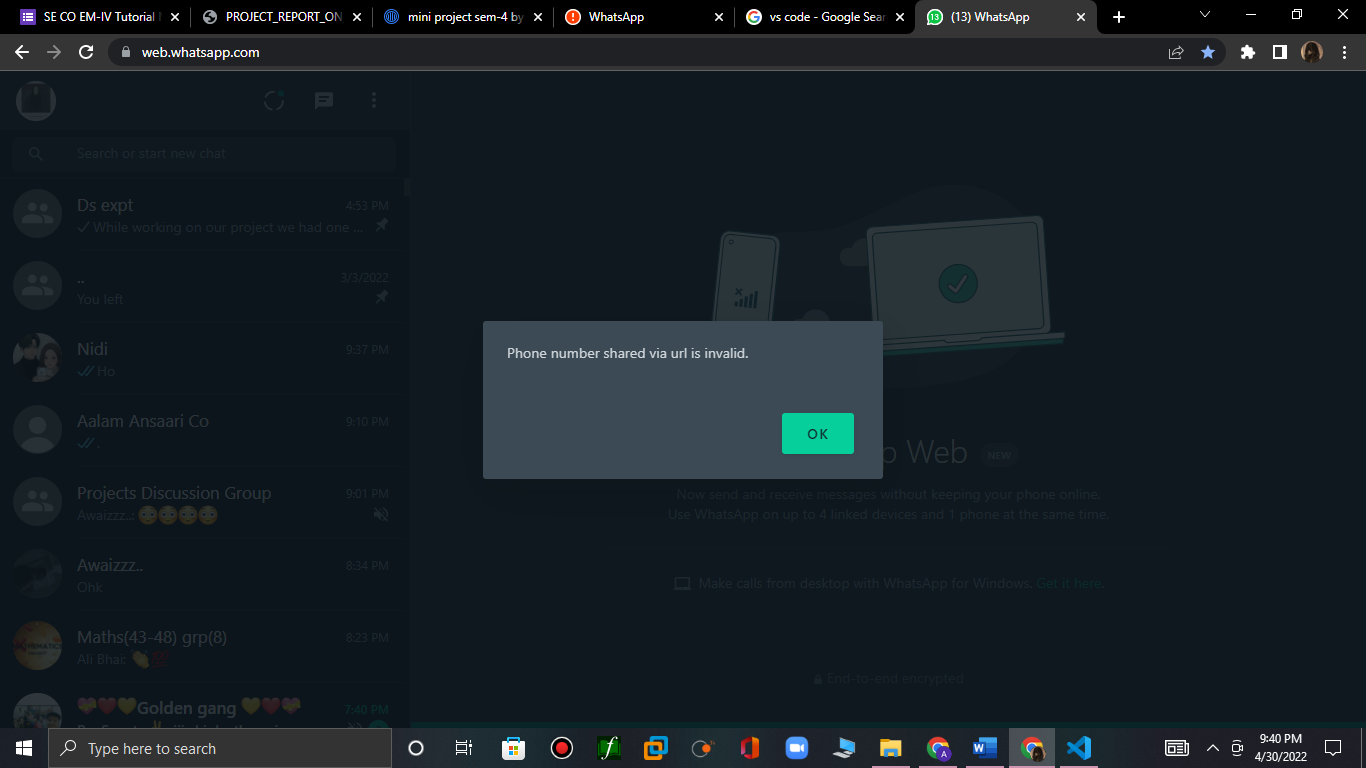


Now the note pad has opened successfully, similarly user can close it with close notepad command

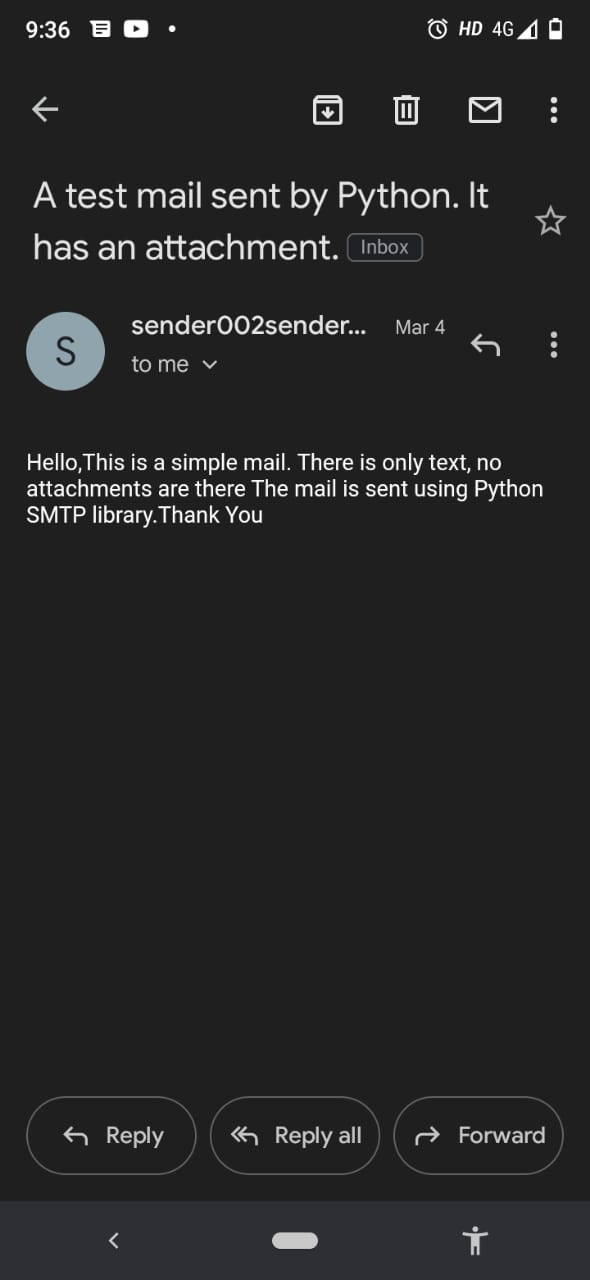


and user can say many more commands like open vs code, open camera, paly music from our default device or play song on youtube, sportify, gaana, etc and asking for news or joke, setting up alarm, open any site from Wikipedia, google etc

we can also send email and whatsapp message

user don’t need to log into web it can automatically log in into WhatsApp web and sends messages if the phone number is wrong then it will show the following error 

here is the snapshot of email:



**3.5 Conclusion and Future work**.

**Conclusion:**  
There can be no doubt that voice assistants are, and will continue to become, a great feat of human ingenuity and they are already creeping into our lives in some shape or form. With the eventual roll-out of 5G and the improvement in machine learning voice assistants may be setting themselves up to be a tool we cannot live without.

However, before we get to that stage, there are hurdles to cross which include heavy investment, improvement in the technology and confidence from consumers that this device that is in their lives does not pose a risk to their privacy.

Our mini project “prototype-1” understands the human natural language and come up with a particular result, it minimizes the human efforts and used to interact with IoT (Internet of things) also it is voice controlled we can just give any command and then it will show us one result if the query matches otherwise, it will user to speak again or to say that query again.

**Future Work:**

As of now our mini project “Prototype-1 ” work only on terminal which is a major drawback of our project the very first thing we are going to do is to make it available on internet or play store and to make the GUI of our project which makes it more interesting and interactive

Then in next semester we will try to make a small application “prototype-2” which is a chat bot for deaf peoples it will help then to interact and learn sign language and then we have “prototype-3” which is our major project for final semester.

In that project we will combine our current project and next projects with some hidden features.

**References**

* <https://www.google.com/search?q=vs+code&rlz=1C1CHZN_enIN920IN920&oq=vs+code&aqs=chrome..69i57j0i433i512l2j0i512j0i433i512l2j0i512j0i433i512j0i512.1538j0j4&sourceid=chrome&ie=UTF-8>
* https://www.python.org
* <https://www.ripublication.com/irph/ijert_spl17/ijertv10n1spl_80.pdf>
* <https://www.researchgate.net/profile/Amrita-Tulshan-2/publication/330165159_Survey_on_Virtual_Assistant_Google_Assistant_Siri_Cortana_Alexa_4th_International_Symposium_SIRS_2018_Bangalore_India_September_19-22_2018_Revised_Selected_Papers/links/5c75283c92851c6950418394/Survey-on-Virtual-Assistant-Google-Assistant-Siri-Cortana-Alexa-4th-International-Symposium-SIRS-2018-Bangalore-India-September-19-22-2018-Revised-Selected-Papers.pdf>