

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 **SpeechRecognition** 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.
10. It can have some basic conversation.

Now the basic question arises in mind that how it is an AI? The virtual assistant that I have created is like if it is not an A.I, but it is the output of a bundle of the statement. But fundamentally, the main purpose of A.I machines is that it can perform human tasks with the same efficiency or even more efficiently than humans. It is a fact that my virtual assistant is not a very good example of A.I., but it is an A.I.

Chapter 1: Introduction

1. INTRODUCTION

Artificial Intelligence when used with machines, it shows us the capability of thinking like humans. In this, a computer system is designed in such a way that typically requires interaction from human. 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 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.

As the voice assistant is using Artificial Intelligence hence the result that it is providing are highly accurate and efficient. The assistant 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. The assistant is no less than a human assistant but we can say that this is more effective and efficient to perform any task. The libraries and packages used to make this assistant focuses on the time complexities and reduces time.

The functionalities include , It can send emails, It can read PDF, It can send text on WhatsApp, It can open command prompt, your favorite IDE, notepad etc., It can play music, It can do Wikipedia searches for you, It can open websites like Google, YouTube, etc., in a web browser, It can give weather forecast, It can give desktop reminders of your choice. It can have some basic conversation.

Tools and technologies used are PyCharm IDE for making this project, and I created all py files in PyCharm. Along with this I used following modules and libraries in my project. pyttsx3, SpeechRecognition, Datetime, Wikipedia, Smtplib, pywhatkit, pyjokes, pyPDF2, pyautogui, PyQt etc. I have created a live GUI for interacting with the JARVIS as it gives a design and interesting look while having the conversation.

1.1 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 listen the command given by the user as per their requirements and perform 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 consume 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.

1.2 PROPOSED SYSTEM

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. Jarvis is different from other traditional voice assistants in terms that it is specific to desktop and user does not need to make account to use this, it does not require any internet connection while getting the instructions to perform any specific task.

The IDE used in this project is PyCharm. All the python files were created in PyCharm and all the necessary packages were easily installable in this IDE. For this project following modules and libraries were used i.e. pyttsx3, SpeechRecognition, Datetime, Wikipedia, Smtplib, pywhatkit, pyjokes, pyPDF2, pyautogui, PyQt etc. I have created a live GUI for interacting with the JARVIS as it gives a design and interesting look while having the conversation.

With the advancement JARVIS 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, It can send emails, It can read PDF, It can send text on WhatsApp, It can open command prompt, your favorite IDE, notepad etc., It can play music, It can do Wikipedia searches for you, It can open websites like Google, YouTube, etc., in a web browser, It can give weather forecast, It can give desktop reminders of your choice. It can have some basic conversation.

Chapter 2: System Design

2.1. DATA FLOW

The data flow for JARVIS is as follow:

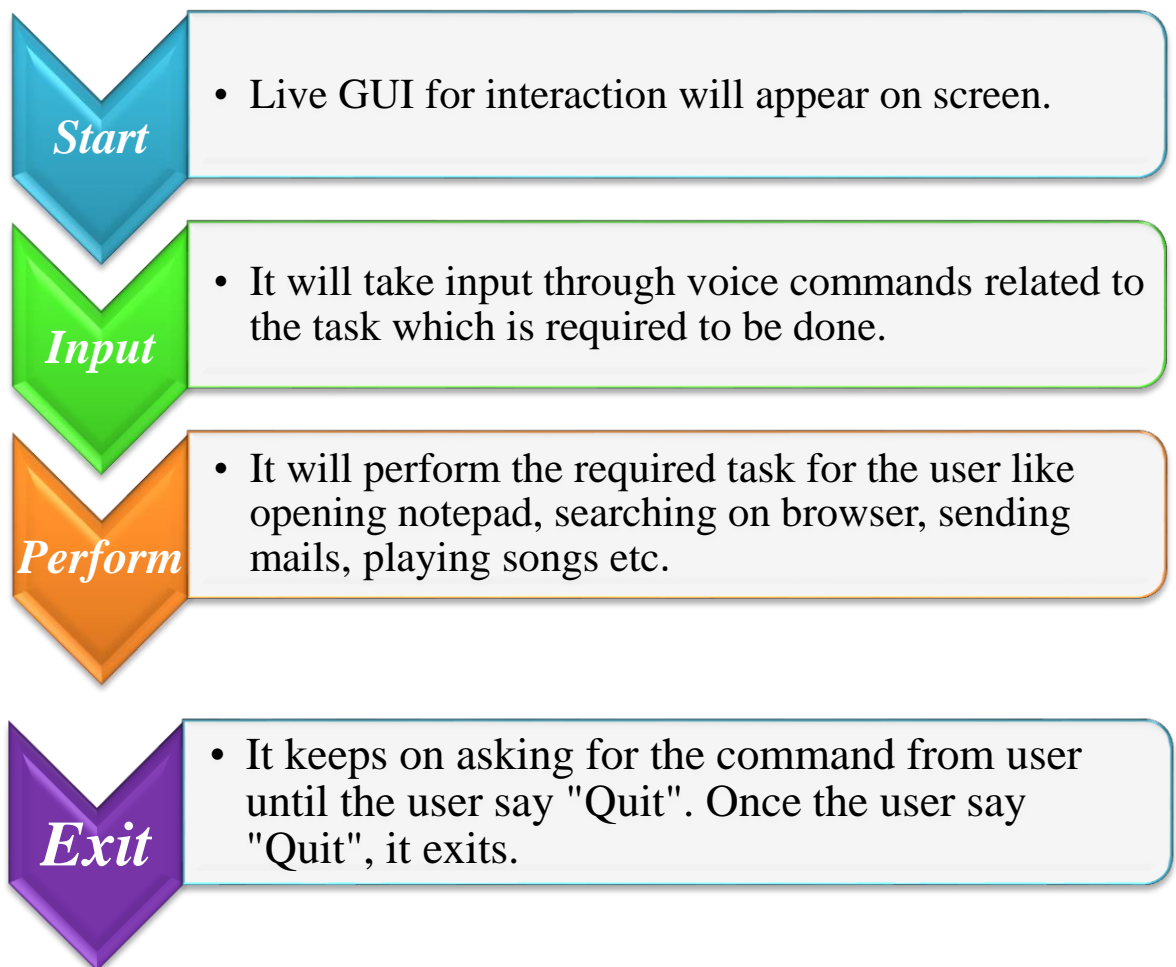


Figure 2.1 Data flow for JARVIS

The system is designed using the concept of Artificial Intelligence and with the help of necessary packages of Python. Python provides many libraries and packages to perform the tasks, for example pyPDF2 can be used to read PDF. The details of these packages are mentioned in Chapter 3 of this report.

The data in this project is nothing but user input, whatever the user says, the assistant performs the task accordingly. The user input is nothing specific but the list of tasks which a user wants to get performed in human language i.e. English.

Chapter 3: Software Details

The IDE used in this project is PyCharm. All the python files were created in PyCharm and all the necessary packages were easily installable in this IDE. For this project following modules and libraries were used i.e. pyttsx3, SpeechRecognition, Datetime, Wikipedia, Smtplib, pywhatkit, pyjokes, pyPDF2, pyautogui, PyQt etc. I have created a live GUI for interacting with the JARVIS as it gives a design and interesting look while having the conversation.

3.1. PYCHARM

It is an IDE i.e. Integrated Development Environment which has many features like it supports scientific tools(like matplotlib, numpy, scipy etc) web frameworks (example Django,web2py and Flask) refactoring in Python, integrated python debugger, code completion, code and project navigation etc. It also provides Data Science when used with Anaconda.

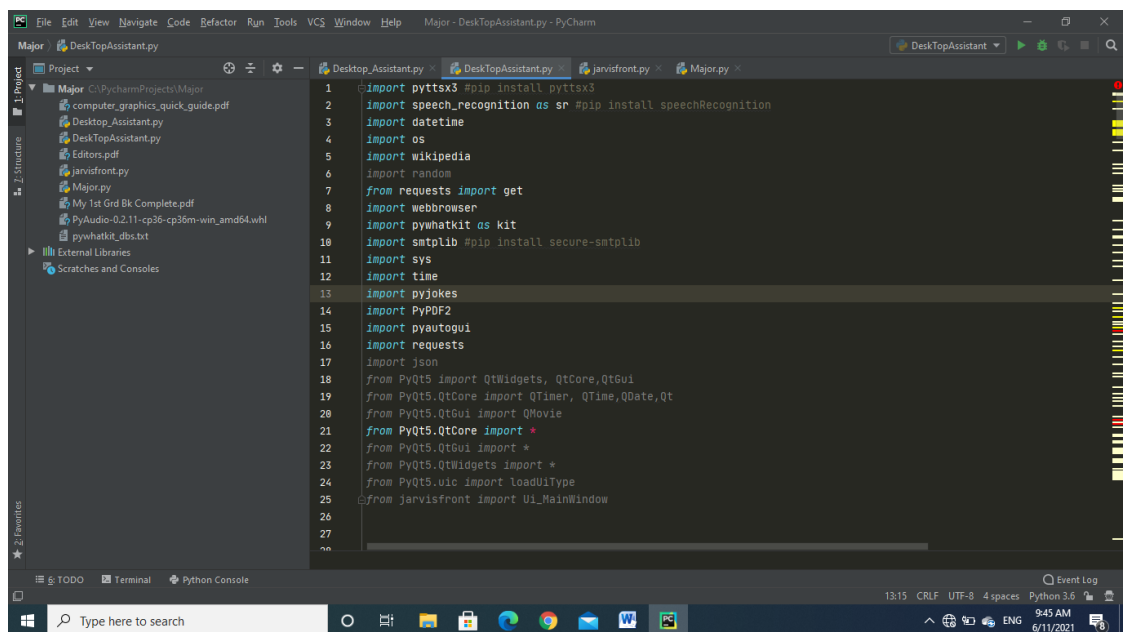


Figure 3.1 PyCharm IDE

3.2. 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.

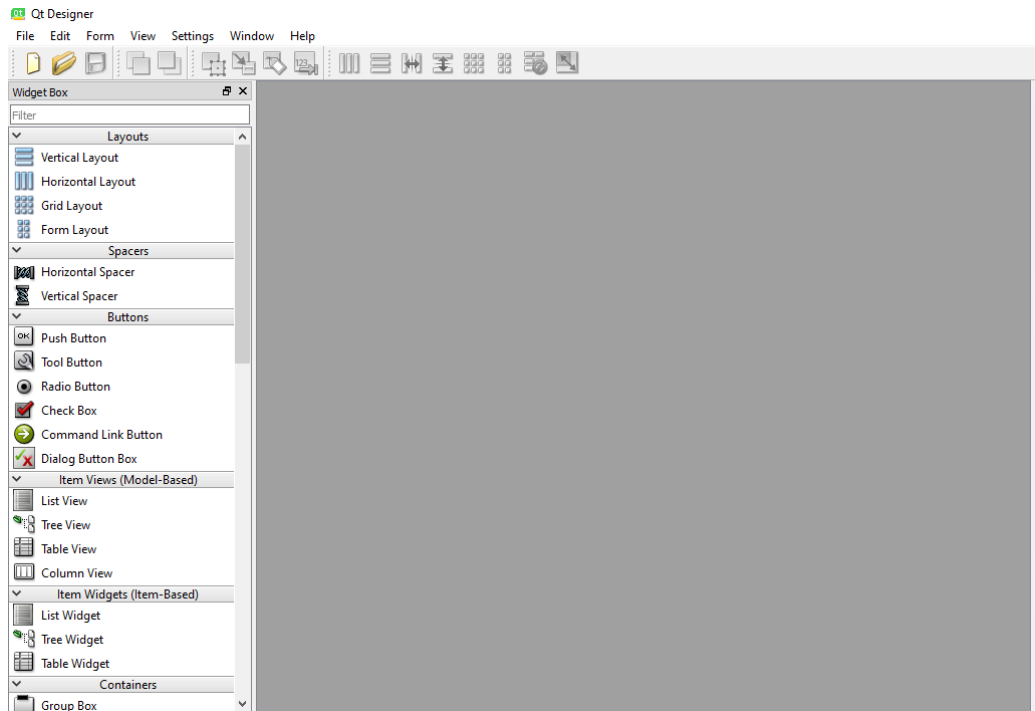


Figure 3.2 PyQt5

3.3. PYTHON LIBRARIES

In JARVIS following python libraries were used:

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

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

3.3.3. pywhatkit: It is python library to send WhatsApp message at a particular time with some additional features.

3.3.4. Datetime: This library provides us the actual date and time.

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

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

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

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

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

3.3.10. Pyautogui: It is a python libraries for graphical user interface.

3.3.11. os: It represents Operating System related functionality.

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

```
1  import pyttsx3 #pip install pyttsx3
2  import speech_recognition as sr #pip install speechRecognition
3  import datetime
4  import os
5  import wikipedia
6  import random
7  from requests import get
8  import webbrowser
9  import pywhatkit as kit
10 import smtplib #pip install secure-smtplib
11 import sys
12 import time
13 import pyjokes
14 import PyPDF2
15 import pyautogui
16 import requests
17 import json
18 from PyQt5 import QtWidgets, QtCore, QtGui
19 from PyQt5.QtCore import QTimer, QTime, QDate, Qt
20 from PyQt5.QtGui import QMovie
21 from PyQt5.QtCore import *
22 from PyQt5.QtGui import *
23 from PyQt5.QtWidgets import *
24 from PyQt5.uic import loadUiType
25 from jarvisfront import Ui_MainWindow
```

Figure 3.3 Imported Modules

Chapter 4: Implementation Work Details

JARVIS, a desktop assistant is a voice assistant that can perform many daily tasks of desktop like playing music, opening your favorite IDE with the help of a single voice command. Jarvis is different from other traditional voice assistants in terms that it is specific to desktop and user does not need to make account to use this, it does not require any internet connection while getting the instructions to perform any specific task.

4.1. REAL LIFE APPLICATION

4.1.1. Saves time: JARVIS is a desktop voice assistant which works on the voice command offered to it, it can do voice searching, voice-activated device control and can let us complete a set of tasks.

4.1.2. Conversational interaction It makes it easier to complete any task as it automatically do it by using the essential module or libraries of Python, in a conversational interaction way. Hence any user when instruct any task to it, they feel like giving task to a human assistant because of the conversational interaction for giving input and getting the desired output in the form of task done.

4.1.3. Reactive nature: The desktop assistant is reactive which means it know human language very well and understand the context that is provided by the user and gives response in the same way, i.e. human understandable language, English. So user finds its reaction in an informed and smart way.

4.1.4. Multitasking: The main application of it can be its multitasking ability. It can ask for continuous instruction one after other until the user “QUIT” it.

4.1.5. No Trigger phase: It asks for the instruction and listen the response that is given by user without needing any trigger phase and then only executes the task.

4.2. 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 follows:

4.2.1. LIBRARIES AND PACKAGES

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

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

4.2.2.3. pywhatkit: It is python library to send WhatsApp message at a particular time with some additional features.

4.2.2.4. Datetime: This library provides us the actual date and time.

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

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

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

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

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

4.2. 2.10. Pyautogui: It is a python librariy for graphical user interface.

4.2.2.11. os: It represents Operating System related functionality.

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

4.2.2. FUNCTIONS

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

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

4.2.2.3. 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.

Chapter 5: Source Code and Commands

Chapter 6: Input/Output Screenshot

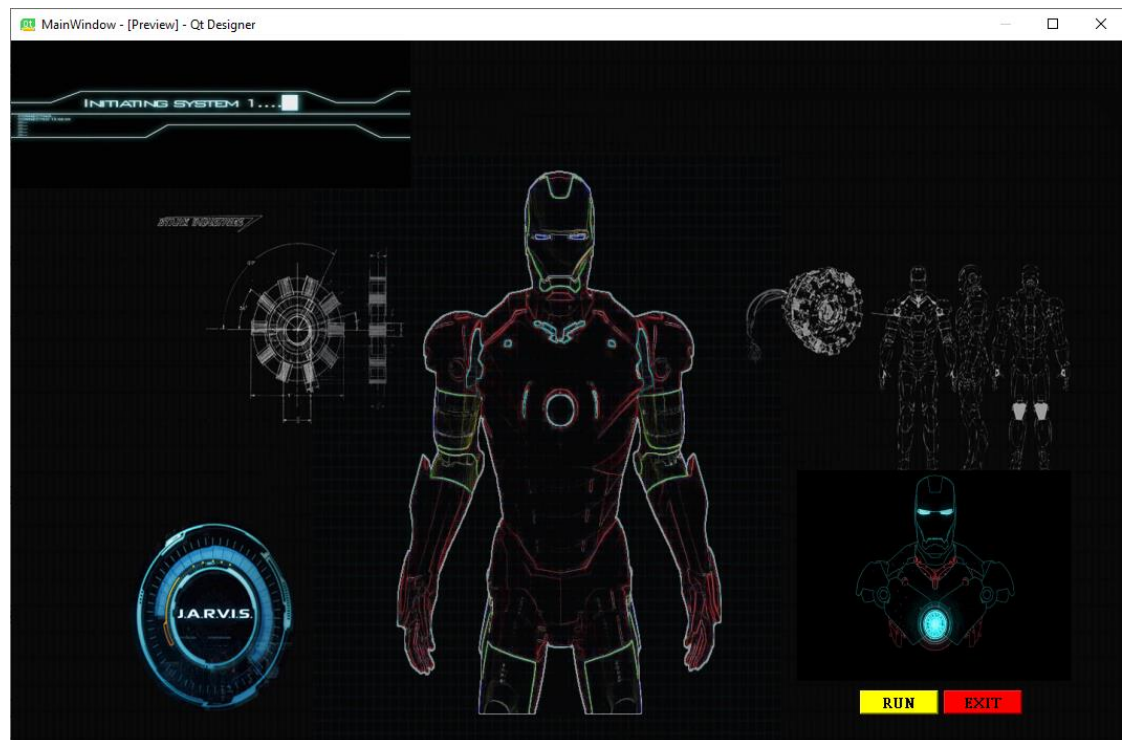


Figure 6.1 Live GUI of JARVIS

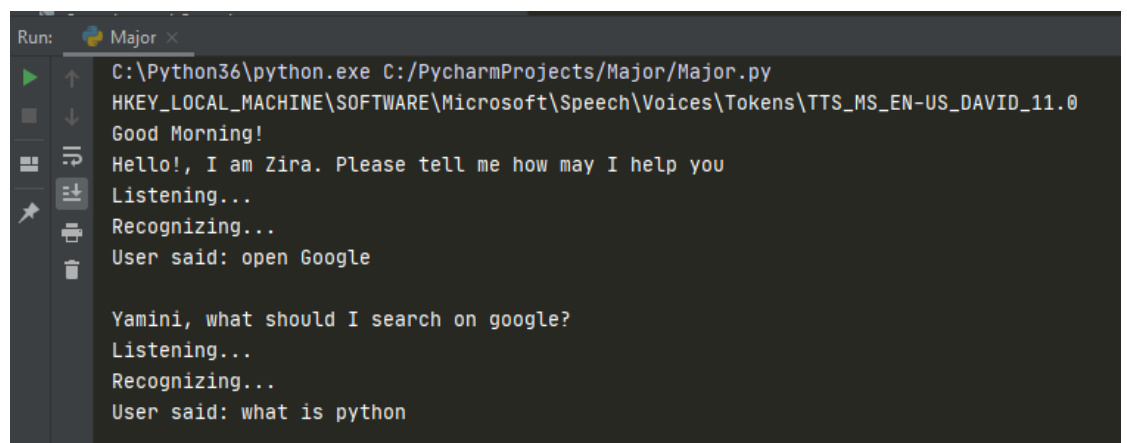


Figure 6.2 Input for Google search

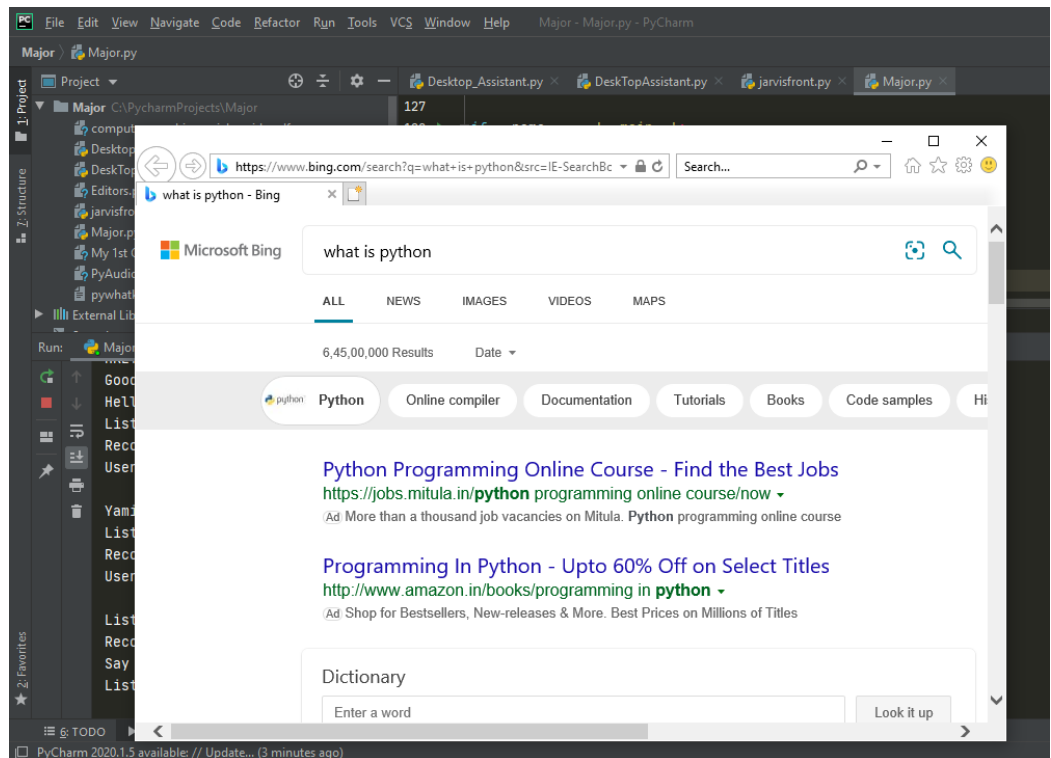


Figure 6.3 Output for Google search

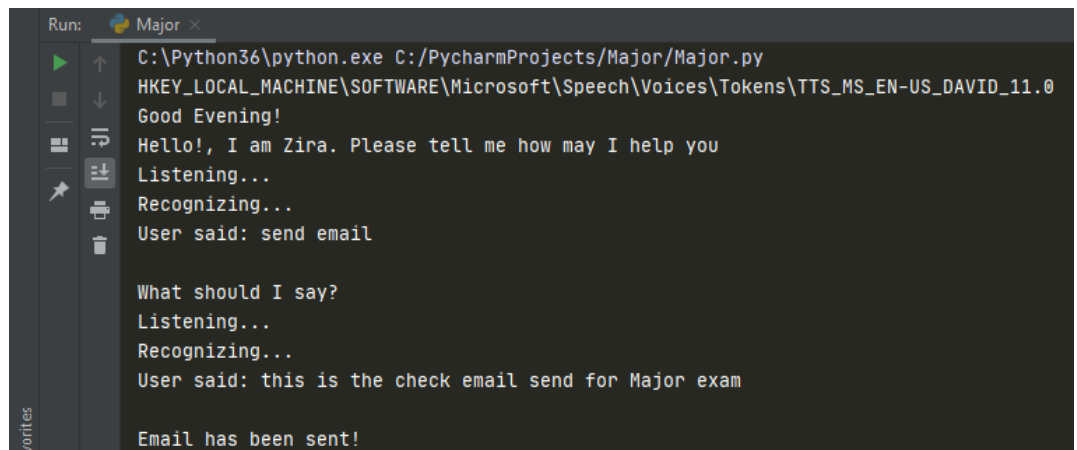


Figure 6.4 Input to send Email

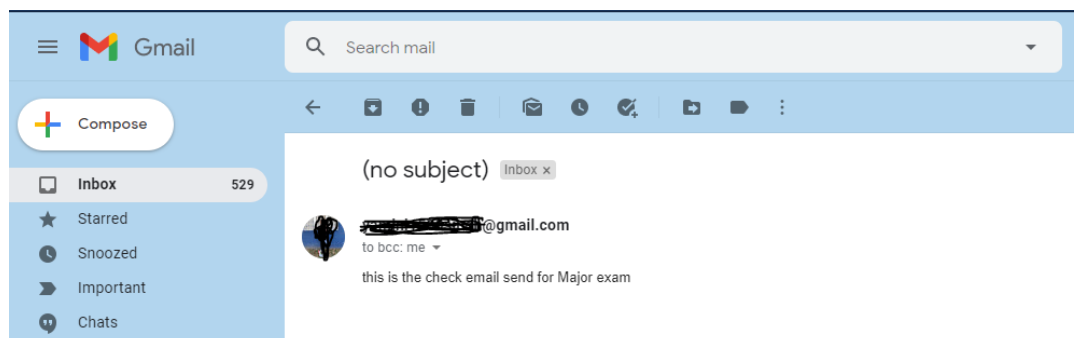


Figure 6.5 Output to send Email

```
Run: Major x
C:\Python36\python.exe C:/PycharmProjects/Major/Major.py
HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\Speech\Voices\Tokens\TTS_MS_EN-US_DAVID_11.0
Good Evening!
Hello!, I am Zira. Please tell me how may I help you
Listening...
Recognizing...
User said: open YouTube

Yamini, what should I search on YouTube?
Listening...
Recognizing...
User said: calma song
```

Figure 6.6 Input for YouTube search

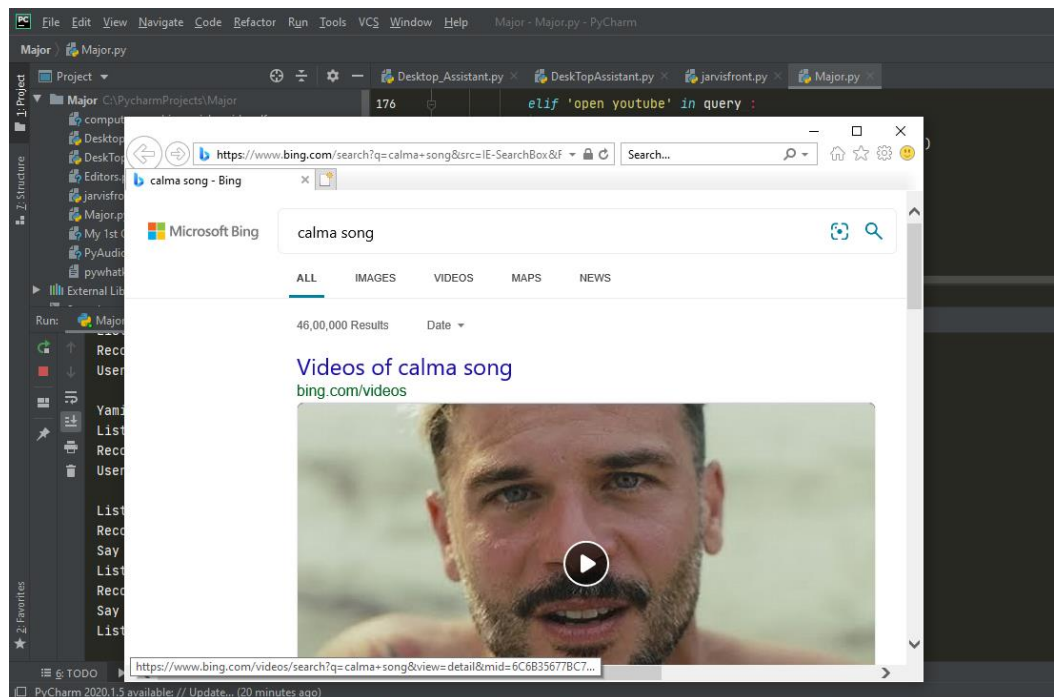


Figure 6.7 Output for YouTube search

```
Run: Major x
C:\Python36\python.exe C:/PycharmProjects/Major/Major.py
HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\Speech\Voices\Tokens\TTS_MS_EN-US_DAVID_11.0
Good Evening!
Hello!, I am Zira. Please tell me how may I help you
Listening...
Recognizing...
User said: play music
```

Figure 6.8 Input to play music

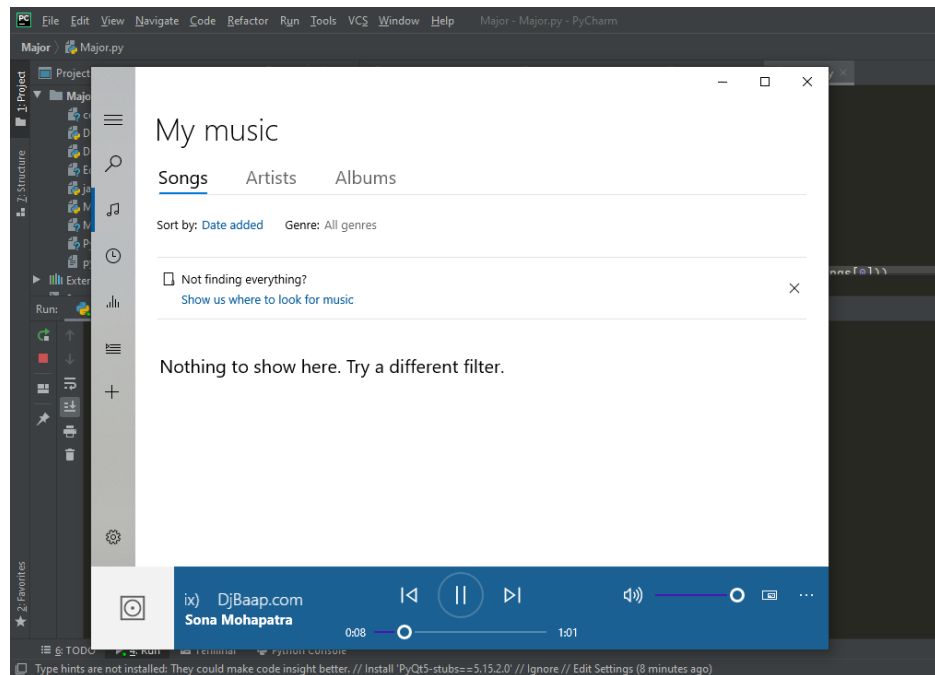


Figure 6.9 Output to play music

```

Listening...
Recognizing...
Say that again please...
Listening...
Recognizing...
User said: open command prompt

```

Figure 6.10 Input to open cmd

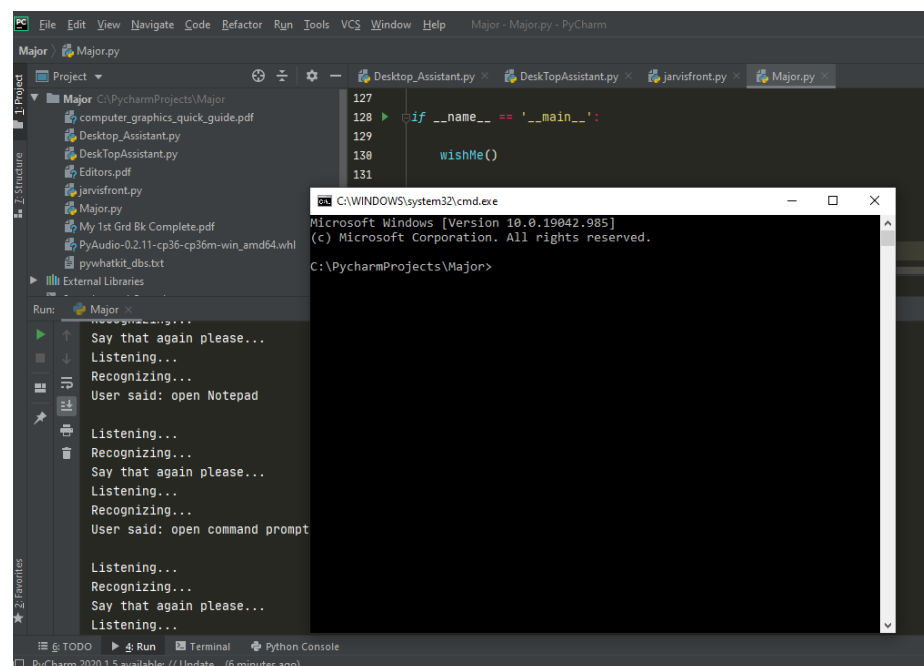


Figure 6.11 Output to open cmd

```
Run: Major x
C:\Python36\python.exe C:/PycharmProjects/Major/Major.py
HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\Speech\Voices\Tokens\TTS_MS_EN-US_DAVID_11.0
Good Evening!
Hello!, I am Zira. Please tell me how may I help you
Listening...
Recognizing...
User said: search Python on Wikipedia

Searching Wikipedia...
Yamini, According to Wikipedia....
Python is an interpreted high-level general-purpose programming language. Python's design philosophy
```

Figure 6.12 Input and output for Wikipedia search

```
Run: Major x
C:\Python36\python.exe C:/PycharmProjects/Major/Major.py
HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\Speech\Voices\Tokens\TTS_
Good Evening!
Hello!, I am Zira. Please tell me how may I help you
Listening...
Recognizing...
User said: open Microsoft Office
```

Figure 6.13 Input to open Microsoft Office

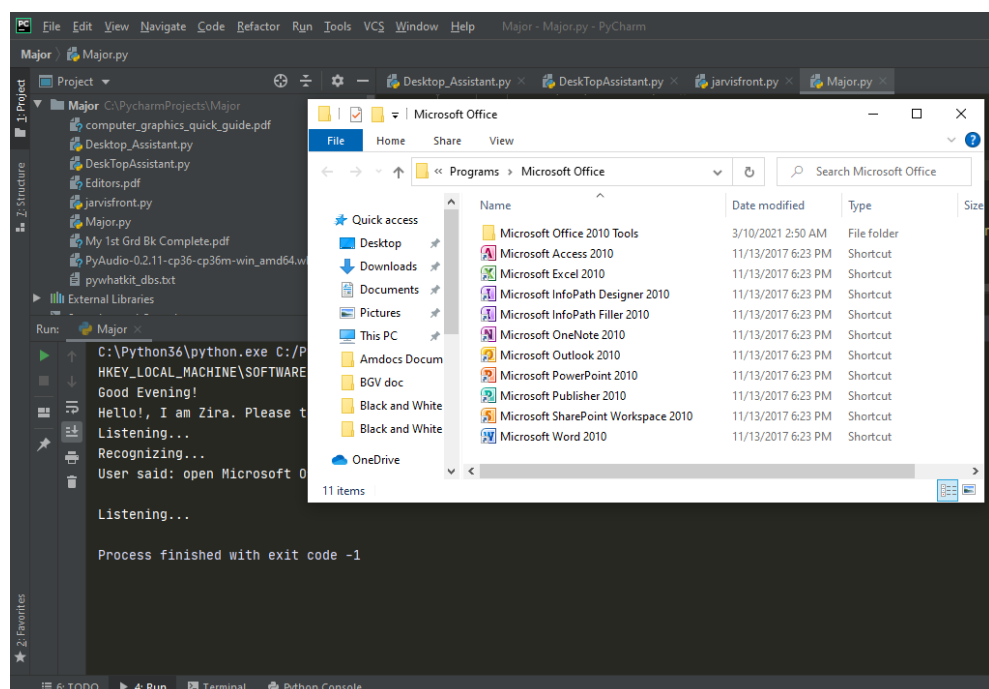


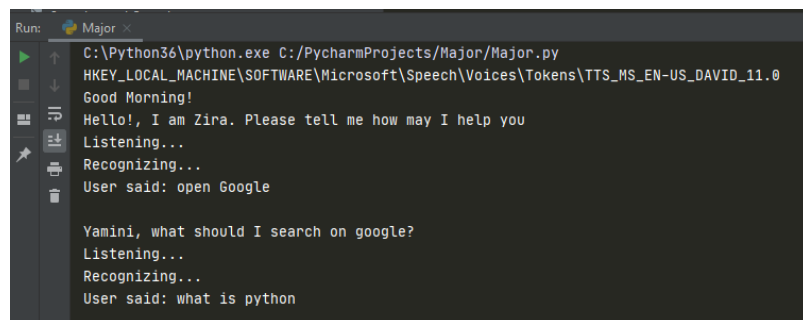
Figure 6.14 Output to open Microsoft Office

Chapter 7: System testing

The system testing is done on fully integrated system to check whether the requirements are matching or not. The system testing for JARVIS desktop assistant focuses on the following four parameters:

7.1. FUNCTIONALITY

In this we check the functionality of the system whether the system performs the task which it was intended to do. To check the functionality each function was checked and run, if it is able to execute the required task correctly then the system passes in that particular functionality test. For example to check whether JARVIS can search on Google or not, as we can see in the figure 7.1, user said “Open Google”, then Jarvis asked, ”What should I search on Google?” then user said, “What is Python”, Jarvis open Google and searched for the required input.



```
Run: Major x
C:\Python36\python.exe C:/PycharmProjects/Major/Major.py
HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\Speech\Voices\Tokens\TTS_MS_EN-US_DAVID_11.0
Good Morning!
Hello!, I am Zira. Please tell me how may I help you
Listening...
Recognizing...
User said: open Google

Yamini, what should I search on google?
Listening...
Recognizing...
User said: what is python
```

Figure 7.1 Input through voice commands

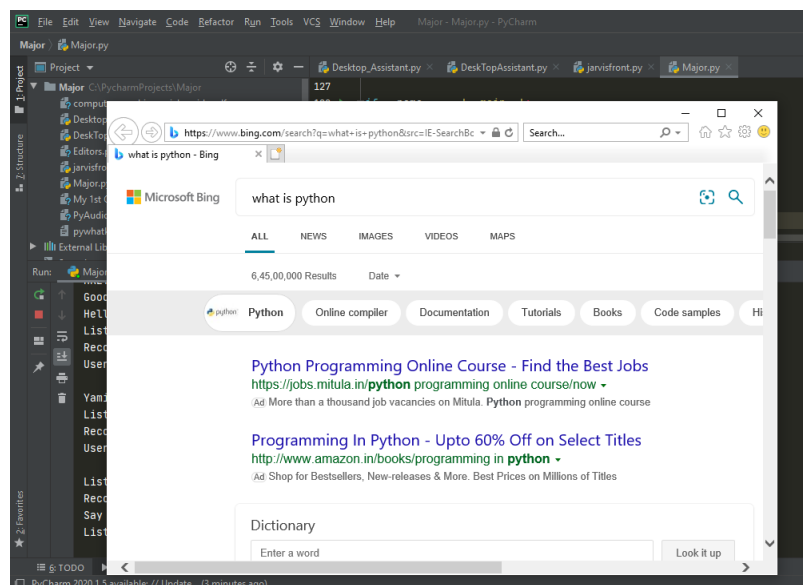


Figure 7.2 Output

7.2. USABILITY

Usability of a system is checked by measuring the easiness of the software and how user friendly it is for the user to use, how it responds to each query that is being asked by the user.

It makes it easier to complete any task as it automatically do it by using the essential module or libraries of Python, in a conversational interaction way. Hence any user when instruct any task to it, they feel like giving task to a human assistant because of the **conversational interaction** for giving input and getting the desired output in the form of task done.

The desktop assistant is **reactive** which means it know human language very well and understand the context that is provided by the user and gives response in the same way, i.e. human understandable language, English. So user finds its reaction in an informed and smart way.

The main application of it can be its **multitasking** ability. It can ask for continuous instruction one after other until the user “QUIT” it. It asks for the instruction and listen the response that is given by user without needing any **trigger phase** and then only executes the task.

7.3. SECURITY

The security testing mainly focuses on vulnerabilities and risks. As JARVIS is a local desktop application, hence there is no risk of data breaching through remote access. The software is dedicated to a specific system so when the user logs in, it will be activated.

7.4. STABILITY

Stability of a system depends upon the output of the system, if the output is bounded and specific to the bounded input then the system is said to be stable. If the system works on all the poles of functionality then it is stable.

Chapter 8: Individual Contribution

The project titled “**A.I. DESKTOP VOICE ASSISTANT: JARVIS**” was designed by me individually. From installing of all the packages, importing, creating all the necessary functions, designing GUI in PyQt and connecting that live GUI with the backend, was all done by me individually.

I, myself have done all the research before making this project, designed the requirement documents for the requirements and functionalities, wrote synopsis and all the documentation, code and made the project in such a way that it is deliverable at each stage. I have created the front end (.ui file) of the project using PyQt designer, the front end comprises of a live GUI and is connected with the .py file which contains all the classes and packages of the .ui file. The live GUI consists of moving GIFs which makes the front end attractive and user friendly.

I have written the complete code in Python language and in PyCharm IDE from where it was very easy to install the packages and libraries, I have created the functions like takeCommand(), wishMe() and taskExecution() which has the following functionalities, like takeCommand() which is used to take the command as input through microphone of user and returns the output as string, wishMe() that greets the user according to the time like Good Morning, Good Afternoon and Good Evening and taskExecution() 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.

While making this project I realized that with the advancement JARVIS 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, It can send emails, It can read PDF, It can send text on WhatsApp, It can open command prompt, your favorite IDE, notepad etc., It can play music, It can do Wikipedia searches for you, It can open websites like Google, YouTube, etc., in a web browser, It can give weather forecast, It can give desktop reminders of your choice. It can have some basic conversation.

At last, I have updated my report and completed it by attaching all the necessary screen captures of inputs and outputs, mentioning the limitations and scope in future of this project.

Chapter 9: Conclusion

JARVIS is a very helpful voice assistant without any doubt as it saves time of the user by conversational interactions, its effectiveness and efficiency. But while working on this project, there were some limitations encountered and also realized some scope of enhancement in the future which are mentioned below:

9.1. LIMITATIONS

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

9.2. SCOPE FOR FUTURE WORK

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