##### VOICE ASSISTANT USING PYTHON

**A PROJECT REPORT**

###### ***Submitted by***

###### 

**Dakshal Dalsania (20BCE10172)**

**Utkarsh Tiwari (20BCE10284)**

**Yash Kabra (20BCE10343)**

**Sanket Kabra (20BCE10447)**

*in partial fulfillment for the award of the degree*

*of*

##### **BACHELOR OF TECHNOLOGY**

*in*

# 

# COMPUTER SCIENCE AND ENGINEERING

****

**SCHOOL OF COMPUTING SCIENCE AND ENGINEERING**

**VIT BHOPAL UNIVERSITY**

**KOTHRI KALAN, SEHORE**

**MADHYA PRADESH - 466114**

April 2022

**BONAFIDE CERTIFICATE**

Certified that this project report titled **“VOICE ASSISTANT USING PYTHON”** is the bonafide work of “**Dakshal Dalsania (Reg No: 20BCE10172), Utkarsh Tiwari (Reg No: 20BCE10284), Yash Kabra (Reg No: 20BCE10343), Sanket Kabra (Reg No: 20BCE10447)”** who carried out the project work under my supervision. Certified further that to the best of my knowledge the work reported at this time does not form part of any other project/research work based on which a degree or award was conferred on an earlier occasion on this or any other candidate.

**PROGRAM CHAIR PROJECT GUIDE**

Dr. Sandeep Mal Dr. Murugeswari

School of Computer Science and Engineering School of Computer Science and Engineering

VIT BHOPAL UNIVERSITY VIT BHOPAL UNIVERSITY

The Project Exhibition II Examination is held on **23rd April, 2022.**

**ACKNOWLEDGEMENT**

First and foremost I would like to thank the Lord Almighty for His presence and immense blessings throughout the project work.

I wish to express my heartfelt gratitude to Dr. Sandeep Mal, Head of the Department, School of Computer Science for much of his valuable support and encouragement in carrying out this work.

I would like to thank my internal guide Ms. Dr. Murugeswari, for continually guiding and actively participating in my project, giving valuable suggestions to complete the project work.

I would like to thank all the technical and teaching staff of the School of Computer Science, who extended directly or indirectly all support.

Last, but not least, I am deeply indebted to my parents who have been the greatest support while I worked day and night for the project to make it a success.

**ABSTRACT**

Today technological advancement is increasing day by day. Earlier only there was a computer system in which we could do only a few tasks. But now machine learning, artificial intelligence, deep learning, and few more technologies have made computer systems so advanced that we can perform any type of task. In such an era of advancement, if people are still struggling to interact using various input devices, then it’s not worth it. For this reason, we developed a voice assistant using python which allows the user to run any type of command in linux without interaction with the keyboard. The main task of voice assistant is to minimize the use of input

devices like keyboard, mouse etc. It will also reduce the hardware space and cost.

**TABLE OF CONTENTS**

| **CHAPTER NO.** | **TITLE** | **PAGE NO.** |
| --- | --- | --- |
|  | List of Abbreviations  List of Figures and Graphs  List of Tables  Abstract | iii  iv  v  vi |
| 1 | **CHAPTER-1:**  **PROJECT DESCRIPTION AND OUTLINE** Introduction 1.2 Motivation for the work  1.3 [About Introduction to the project  including techniques]  1.5 Problem Statement  1.6 Objective of the work  1.7 Organization of the project  1.8 Summary | 1  .  .  . |
| 2 | **CHAPTER-2:**  **RELATED WORK INVESTIGATION**  2.1 Introduction  2.2 <Core area of the project>  2.3 Existing Approaches/Methods  2.3.1 Approaches/Methods -1  2.3.2 Approaches/Methods -2  2.3.3 Approaches/Methods -3  2.4 <Pros and cons of the stated Approaches/Methods >  2.5 Issues/observations from investigation  2.6 Summary |  |
| 3 | **CHAPTER-3:**  **REQUIREMENT ARTIFACTS**  3.1 Introduction  3.2 Hardware and Software requirements  3.3 Specific Project requirements  3.3.1 Data requirement  3.3.2 Functions requirement  3.3.3 Performance and security requirement  3.3.4 Look and Feel Requirements  3.3.5 ………  3.4 Summary |  |
| 4 | **CHAPTER-4:**  **DESIGN METHODOLOGY AND ITS NOVELTY**  4.1 Methodology and goal  4.2 Functional modules design and analysis  4.3 Software Architectural designs  4.4 Subsystem services  4.5 User Interface designs  4.5 ………………..  4.6 Summary |  |
| 5 | **CHAPTER-5:**  **TECHNICAL IMPLEMENTATION & ANALYSIS**  5.1Outline  5.2 Technical coding and code solutions  5.3 Working Layout of Forms  5.4 Prototype submission  5.5 Test and validation  5.6 Performance Analysis(Graphs/Charts)  5.7 Summary |  |
| 6 | **CHAPTER-6:**  **PROJECT OUTCOME AND APPLICABILITY**  6.1Outline  6.2 key implementations outlines of the System  6.3 Significant project outcomes  6.4 Project applicability on Real-world applications  6.4 Inference |  |
| 7 | **CHAPTER-7:**  **CONCLUSIONS AND RECOMMENDATION**  7.1Outline  7.2 Limitation/Constraints of the System  7.3 Future Enhancements  7.4 Inference |  |
|  | Appendix A  Appendix B  References  ***Note: List of References should be written as per IEEE/Springer reference format. (Specimen attached)*** |  |

**PROJECT DESCRIPTION AND OUTLINE**

***Introduction:***

In this era of technology everything that human beings can do is being replaced by machines. One of the main reasons is change in performance. In today’s world we train our machines to think like humans and do their task by themselves. Therefore, there came a concept of virtual assistant.

A virtual assistant is a digital assistant that uses voice recognition features and language processing algorithms to recognize voice commands of the user and perform relevant tasks as requested by the user. Based on specific commands given by the user a virtual assistant is capable of filtering out the ambient noise and returning relevant information.

Virtual Assistants are completely software based but nowadays they are integrated in different devices and also some of the assistants are designed specifically for single devices like Alexa.

Due to drastic change in technology now it’s a. high time to train our machine with the help of machine learning, deep learning, neural networks. Today we can talk to our machine with the help of Voice Assistant.

The key here is voice. A voice/virtual assistant is a digital assistant that uses voice recognition, speech synthesis, natural language processing (NLP) and also AI to provide a service through a particular application. Speech recognition is the process of converting audio into text. This is commonly used in voice assistants like Alexa, Siri, etc. Python provides an API called Speech Recognition to allow us to convert audio into text for further processing.

***Problem Statement:***

We all want to make our life much easier and faster and so we use small changes in our life which lead to big changes. One such change that we want nowadays is Personal voice assistance. Artificial Intelligence personal assistants have become plentiful over the last few years. Applications such as Siri, Bixby, Ok Google, Alexa and Cortana make mobile device users’ daily routines that much easier. It also makes it easier for us to do small tasks of our daily life easily and fastly like writing notes, playing music, weather info etc. it proves to be very helpful to the visually impaired people.

***Objective of the Work:***

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.

Virtual assistants can tremendously save you time. Another difficult task is to remember test dates, birthdays or anniversaries. It comes with a surprise when you enter the class and realize it is a class test today.

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 time15. In this respect, the ability of personal assistants to accurately recognize spoken words is a prerequisite for them to be adopted by consumers.

**RELATED WORK INVESTIGATION**

Each company developer of the intelligent assistant applies his own specific methods and approaches for development, which in turn affects the final product. One assistant can synthesize speech more qualitatively, another can more accurately and without additional explanations and corrections perform tasks, others can perform a narrower range of tasks, but most accurately and as the user wants. Obviously, there is no universal assistant who would perform all tasks equally well. The set of characteristics that an assistant has depends entirely on which area the developer has paid more attention to. Since all systems are based on machine learning methods and use for their creation huge amounts of data collected from various sources and then trained on them, an important role is played by the source of this data, be it search systems, various information sources or social networks. The amount of information from different sources determines the nature of the assistant, which can result as a result. Despite the different approaches to learning, different algorithms and techniques, the principle of building such systems remain approximately the same.The main technologies are voice activation, automatic speech recognition, Teach-To-Speech, voice biometrics, dialogue manager, natural language understanding and named entity recognition.

**REQUIREMENT ARTIFACTS**

System Analysis is about complete understanding of existing systems and finding where the existing system fails. The solution is determined to resolve issues in the proposed system. It defines the system. The system is divided into smaller parts. Their functions and inter relation of these modules are studied in system analysis. The complete analysis is followed below.

***Problem definition:***

Usually, a 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 are a number of people who have issues in 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 easier to use on mobile devices than 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 bad behaviors. They require a large amount of information to be fed in order for it to work efficiently.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, priorities in order to recommend a particular plan.

***HARDWARE AND SOFTWARE REQUIREMENTS:***

The software is designed to be light-weighted so that it doesn’t be a burden on the machine running it. This system is being built keeping in mind the generally available hardware and software compatibility. Here are the minimum hardware and software requirements for virtual assistants.

**Hardware:**

* Pentium-pro processor or later.
* RAM 512MB or more.

**Software:**

* Windows 7(32-bit) or above.
* Python 2.7 or later
* Chrome Driver

***SPECIFIC PROJECT REQUIREMENTS:***

Personal assistant software is required to act as an interface into the digital world by understanding user requests or commands and then translating into actions or recommendations based on agent’s understanding of the world.JIA focuses on relieving the user of entering text input and using voice as primary means of user input. Agent then applies voice recognition algorithms to this input and records the input. It then uses this input to call one of the personal information management applications such as task list or calendar to record a new entry or to search about it on search engines like Google, Bing or Yahoo etc. Focus is on capturing the user input through voice,recognizing the input and then executing the tasks if the agent understands the task. Softwaretakes this input in natural language, and so makes it easier for the user to input what he or she desires to be done.Voice recognition software enables hands free use of the applications, letting users to query or command the agent through voice interface. This helps users to have access to the agent while performing other tasks and thus enhances the value of the system itself. JIA also has ubiquitous connectivity through Wi-Fi or LAN connection, enabling distributed applications that can leverage other APIs exposed on the web without a need to store them locally.Virtual assistants must provide a wide variety of services. These include:

* Providing information such as weather, facts from e.g. Wikipedia etc.
* Set an alarm or make to-do lists and shopping lists.
* Remind you of birthdays and meetings.
* Play music from streaming services such as Saavn and Gaana.
* Play videos, TV shows or movies on televisions, streaming from e.g. Netflix orHotstar.
* Book tickets for shows, travel and movies.

**Feasibility Study:**

Feasibility study can help you determine whether or not you should proceed with your project. It is essential to evaluate cost and benefit. It is essential to evaluate the cost and benefit of the proposed system. Five types of feasibility study are taken into consideration.

1. **Technical feasibility:**

It includes finding out technologies for the project, both hardware and software. For a virtual assistant, the user must have a microphone to convey their message and a speaker to listen when the system speaks. These are very cheap nowadays and everyone generally possesses them. Besides, the system needs internet connection.While using JIA, make sure you have a steady internet connection. It is also not an issue in this era where almost every home or office has Wi-Fi.

1. **Operational feasibility:**

It is the ease and simplicity of operation of the proposed system.System does not require any special skill set for users to operate it. In fact, it is designed to be used by almost everyone. Kids who still don’t know how to write can read out problems for the system and get answers.

1. **Economical feasibility:**

Here, we find the total cost and benefit of the proposed system over the current system. For this project, the main cost is documentation cost. Useralso would have to pay for a microphone and speakers. Again, they are cheap and available. As far as maintenance is concerned, JIA won’t cost too much.

1. **Organizational feasibility:**

This shows the management and organizational structure of the project. This project is not built by a team. The management tasks are all to be carried out by a single person. That won’t create any management issues and will increase the feasibility of the project.

1. **Cultural feasibility:**

It deals with compatibility of the project with the cultural environment. Virtual assistant is built in accordance with the general culture. The project is named JIA so as to represent Indian culture without undermining local beliefs.This project is technically feasible with no external hardware requirements. Also it is simple in operation and does not cost training or repairs. Overall feasibility study of the project reveals that the goals of the proposed system are achievable. Decision is taken to proceed with the project.

**DESIGN METHODOLOGY AND ITS NOVELTY**

***METHODOLOGY AND GOAL:***

**Speech Recognition module:**

The system uses Google’s online speech recognition system for converting speech input to text. The speech input Users can obtain texts from the special corpora organized on the computer network server at the information center from the microphone is temporarily stored in the system which is then sent to Google cloud for speech recognition. The equivalent text is then received and fed to the central processor.

**Python Backend:**

The python backend gets the output from the speech recognition module and then identifies whether the command or the speech output is an API Call and Context Extraction. The output is then sent back to the python backend to give the required output to the user.

**API calls:**

API stands for Application Programming Interface. An API is a software intermediary that allows two applications to talk to each other. In other words, an API is a messenger that delivers your request to the provider that you’re requesting it from and then delivers the response back to you.

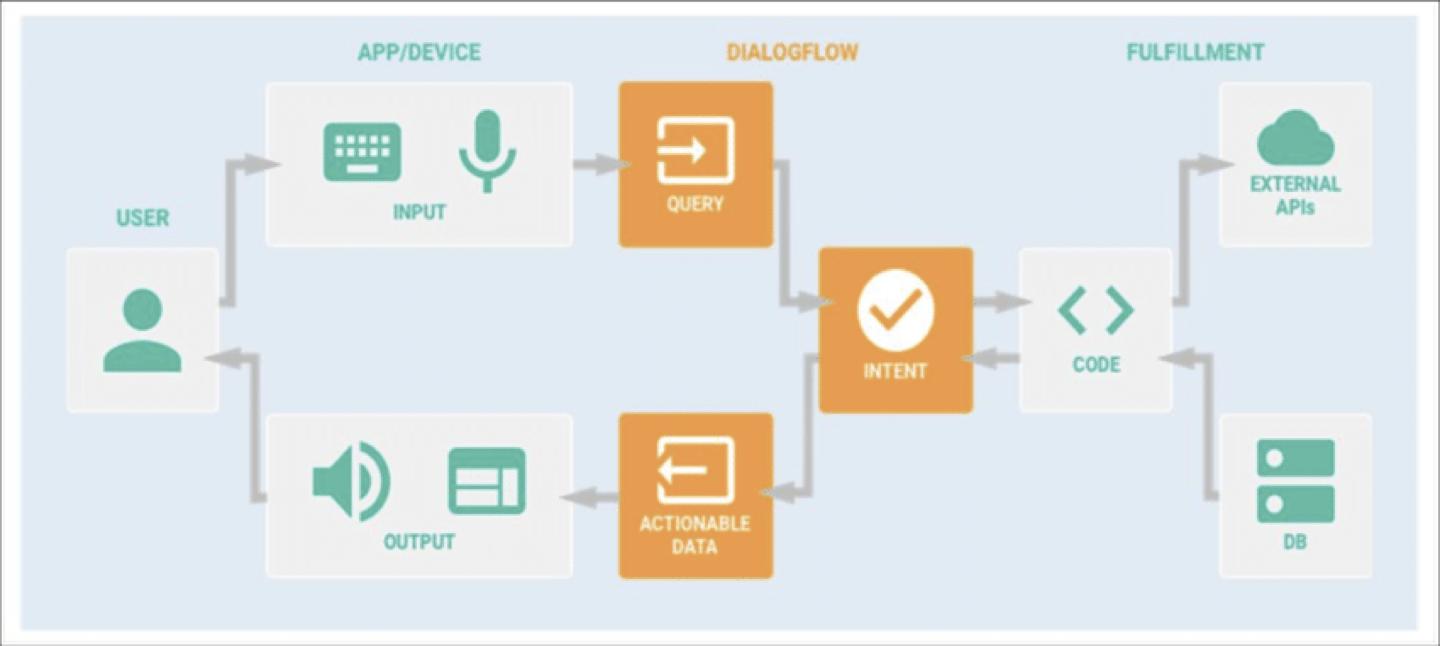
**Content Extraction:**

Context extraction (CE) is the task of automatically extracting structured information from unstructured and/or semi-structured machine-readable documents. In most cases, this activity concerns processing human language texts using natural language processing (NLP). Recent activities in multimedia document processing like automatic annotation and content extraction out of images/audio/video could be seen as context extraction TEST RESULTS.

**Text-to-speech module:**

Text-to-Speech (TTS) refers to the ability of computers to read text aloud. A TTS Engine converts written text to a phonemic representation, then converts the phonemic representation to waveforms that can be output as sound. TTS engines with different languages, dialects and specialized vocabularies are available through third-party publishers.

***SOFTWARE ARCHITECTURAL DESIGN:***

****

**PROJECT OUTCOMES AND APPLICABILITY**

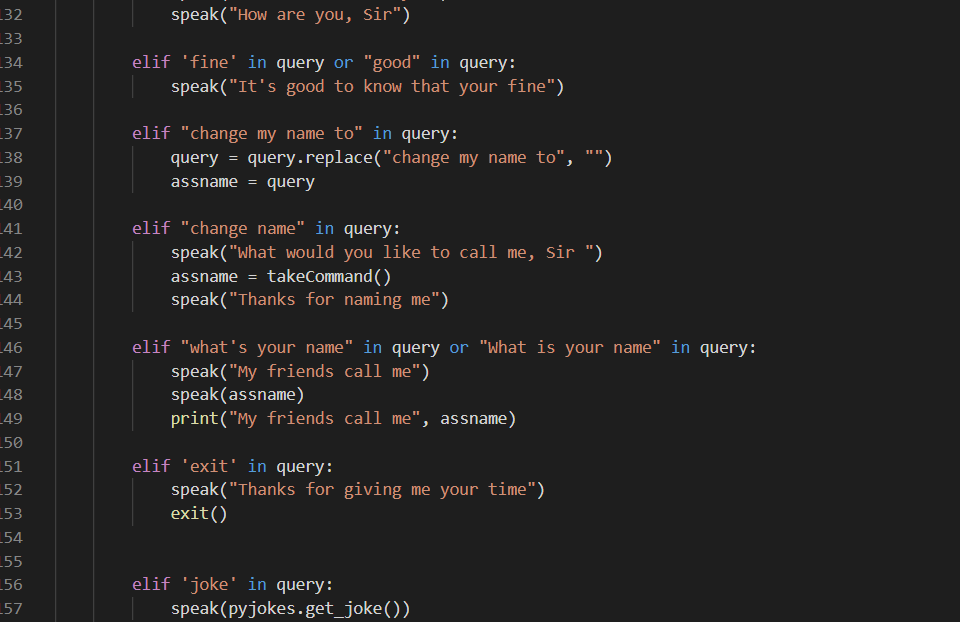
We have employed this idea by means of Python, Machine Learning and AI. Our main aim is to assist the users in their tasks with the help of their voice commands. After translating the command which the user has given, the assistant will display output based on the query input of the user.

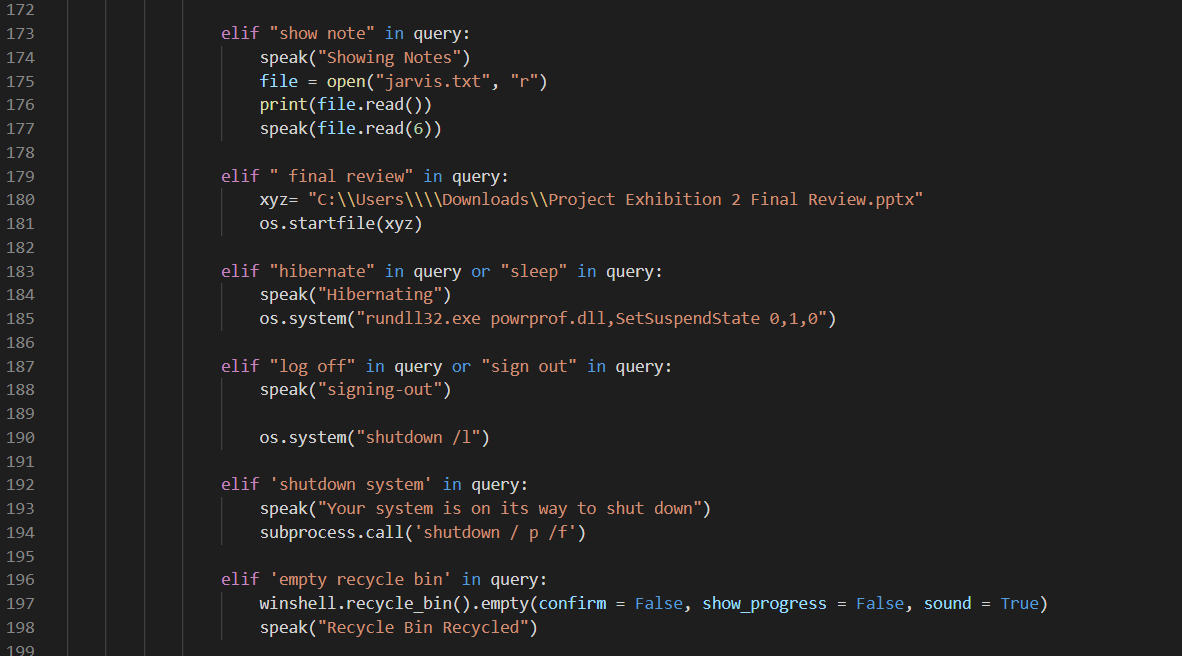
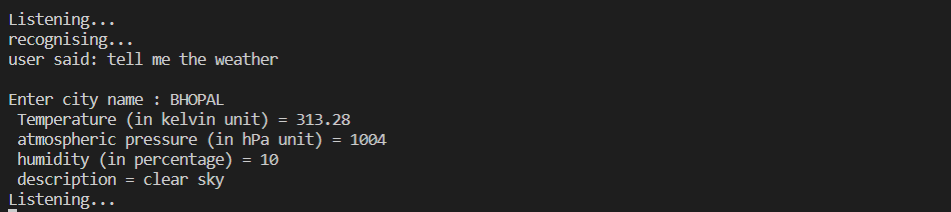
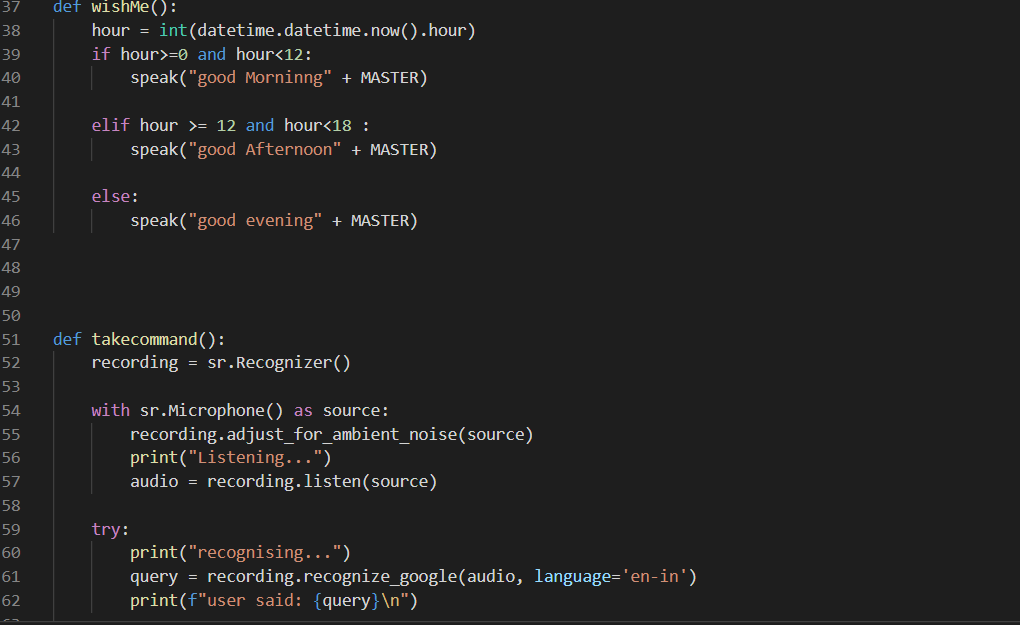
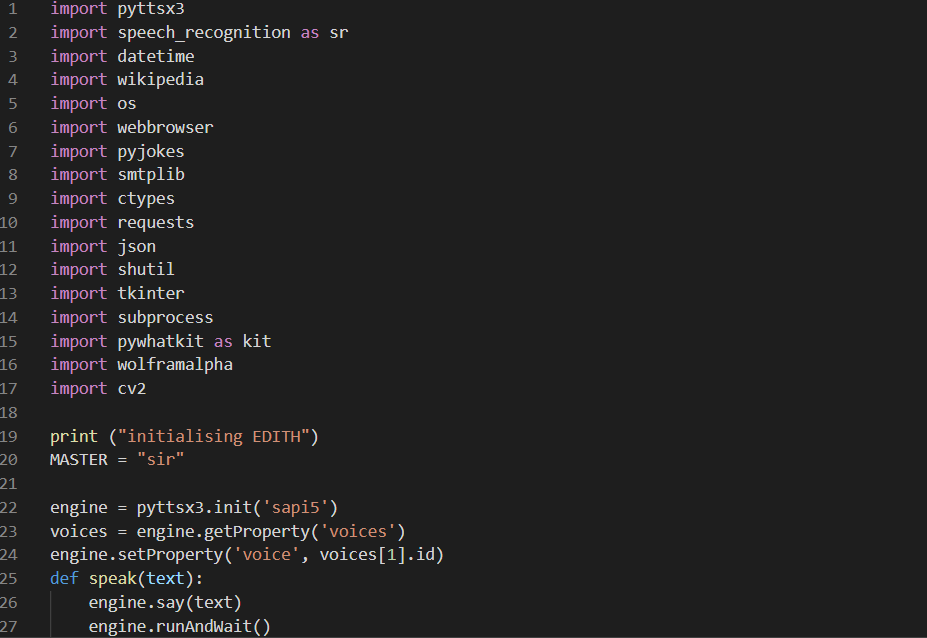
**CONCLUSIONS AND RECOMMENDATION**

In this paper “Virtual Assistant Using Python” we discussed the design and implementation of Digital Assistance. The project is built using open source software modules with PyCharm community backing which can accommodate any updates shortly. The modular nature of this project makes it more flexible and easy to add additional features without disturbing current system functionalities.

This assistant currently works as an application based and performs basic tasks like weather updates, stream music, search Wikipedia, open desktop applications, etc. The functionality of the current system is limited to working on application based only.

It not only works on human commands but also give responses to the user based on the query being asked or the words spoken by the user such as opening tasks and operations. It is greeting the user the way the user feels more comfortable and feels free to interact with the voice assistant. The application should also eliminate any kind of unnecessary manual work required in the user life of performing every task. The entire system works on the verbal input rather than the next one.

***Screenshots of the code (desktop voice assistant)***

******

**REFERENCES**

1. [**https://www.geeksforgeeks.org/voice-assistant-using-python/**](https://www.geeksforgeeks.org/voice-assistant-using-python/)
2. [**https://www.analyticsvidhya.com/blog/2020/11/build-your-own-desktop-voice-assistant-in-python/**](https://www.analyticsvidhya.com/blog/2020/11/build-your-own-desktop-voice-assistant-in-python/)
3. [**https://ijirt.org/master/publishedpaper/IJIRT152099\_PAPER.pdf**](https://ijirt.org/master/publishedpaper/IJIRT152099_PAPER.pdf)
4. [**https://www.irjet.net/archives/V7/i9/IRJET-V7I9678.pdf**](https://www.irjet.net/archives/V7/i9/IRJET-V7I9678.pdf)