PERSONAL AI USING PYTHON

MINI PROJECT -II REPORT

Submitted by

RAJESH K S (621321205038)

SIVA HARISH R (621321205054)

KAVIRAJ D (621321205303)

in partial fulfillment for the award of the degree

of

BACHELOR OF TECHNOLOGY

in

INFORMATION TECHNOLOGY

KONGUNADU COLLEGE OF ENGINEERING AND TECHNOLOGY

(AUTONOMOUS)

THOTTIAM

ANNA UNIVERSITY:: CHENNAI 600 025

MAY 2024

KONGUNADU COLLEGE OF ENGINEERING AND TECHNOLOGY (AUTONOMOUS)
NAMAKKAL- TRICHY MAIN ROAD, THOTTIAM, TRICHY
COLLEGE VISION & MISSION STATEMENT

VISION

"To become an Internationally Renowned Institution in Technical Education, Research, and Development by Transforming the Students into Competent Professionals with Leadership Skills and Ethical Values."

MISSION

- Providing the Best Resources and Infrastructure.
- Creating Learner-Centric Environment and continuous Learning.
- Promoting Effective Links with Intellectuals and Industries.
- Enriching Employability and Entrepreneurial Skills.
- Adapting to Changes for Sustainable Development.

KONGUNADU COLLEGE OF ENGINEERING AND TECHNOLOGY (AUTONOMOUS)
NAMAKKAL- TRICHY MAIN ROAD, THOTTIAM, TRICHY
DEPARTMENT OF INFORMATION TECHNOLOGY

VISION

To produce competent IT professionals, researchers and entrepreneurs with moral values in the field of Information Technology.

MISSION

- Enrich the students' programming and computing skills through best teachinglearning processes, laboratory practices and through project based learning.
- Inculcate real world challenges, emerging technologies and endeavour the students to become entrepreneurs or make them employable.
- Inculcating moral and ethical values to serve the society and focus on students' overall development

PROGRAM EDUCATIONAL OBJECTIVES (PEOs)

- **PEO I:** Graduates shall become IT professionals with specialization in Software Engineering, Networking, Data Mining and Cloud computing.
- **PEO II:** Graduates shall build IT solutions through analysis, design and development of software and firmware solutions for real-world problems and social issues.
- **PEO III:** Graduates shall have professional ethics, team spirit, life-long learning, good oral and written communication skills and adopt corporate culture, core values and leadership skills.

PROGRAM SPECIFIC OUTCOMES (PSOs)

- **PSO1: Professional skills:** Students shall understand, analyze and develop IT applications in the field of Data Mining/Analytics, Cloud Computing, Networking etc., to meet the requirements of industry and society.
- **PSO2: Competency:** Students shall qualify at the State, National and International level competitive examination for employment, higher studies and research.

.

.

PROGRAM OUTCOMES (POs)

Engineering Graduates will be able to:

1. **Engineering knowledge**: Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.
2. **Problem analysis**: Identify, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions using the first principles of mathematics, natural sciences, and engineering sciences.
3. **Design/development of solutions**: Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for public health and safety, and cultural, societal, and environmental considerations.
4. **Conduct investigations of complex problems**: Use research-based knowledge and research methods including design of experiments, analysis, and interpretation of data, and synthesis of the information to provide valid conclusions.
5. **Modern tool usage**: Create, select, and apply appropriate techniques, resources,and modern engineering and IT tools including prediction and modelling to complex engineering activities with an understanding of the limitations.
6. **The engineer and society**: Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice.
7. **Environment and sustainability**: Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.
8. **Ethics**: Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.
9. **Individual and team work**: Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.
10. **Communication**: Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.
11. **Project management and finance**: Demonstrate knowledge and understandingof the engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.
12. **Life-long learning**: Recognize the need for, and have the preparation and abilityto engage in independent and life-long learning in the broadest.

.

KONGUNADU COLLEGE OF ENGINEERING AND TECHNOLOGY

(AUTONOMOUS)

ANNA UNIVERSITY :: CHENNAI 600 025

BONAFIDE CERTIFICATE

Certified that this mini project report "**PERSONAL AI USING PYTHON**" is the bonafide work of **"RAJESH K S (621321205038), SIVA HARISH R (621321205054),** and **KAVIRAJ D (621320205303)"** who carried out the mini project work under my supervision.

SIGNATURE

HEAD OF THE DEPARTMENT

Mr.N.PREMKUMAR, M.E., (Ph.D)

Associate Professor,

Department of Information Technology,

Kongunadu College of Engineering and Technology,Thottiam,Trichy-621 215

SIGNATURE

SUPERVISOR

Mr.J.SATHISHKUMAR,M.TECH., (Ph.D)

Assistant Professor,

Department of Information Technology,

Kongunadu College of Engineering and Technology,Thottiam,Trichy-621 215

Submitted for the Mini-project-II viva-voce examination held on _____

INTERNAL EXAMINER **EXTERNAL EXAMINER**

ACKNOWLEDGEMENT

We wish to express our sincere gratitude to our beloved chairman **Dr.PSK.R.PERIASWAMY** for providing the facilitation to make us extensively internationalized to meet global competence with almost perfection.

We would like to express our sincere thanks to our Principal **Dr.R.ASOKAN,M.S.,M.Tech.,Ph.D.,** for the facilities and the encouragement given to the progression and completion of this project.

We proudly render our immense gratitude to **Mr.N.PREMKUMAR, M.E., (Ph.D)** Associate Professor and Head of the Department, Department of Information Technology for his effective leadership, encouragement and supportive guidance to this project.

We proudly render our thanks to our Mini project-II Coordinator **Mr.R.V.VISWANATHAN , M.E.,(Ph.D).,** Assistant Professor, Department of Information Technology for his valuable ideas, encouragement and supportive guidance throughout this project.

We wish to extend our heartfelt regard and sincere thanks to our Mini project-II guide **Mr.J.SATHISHKUMAR,M.TECH.,(Ph.D).,**Assistant Professor, Department of Information Technology for his motivation, continuous encouragement and expert guidance throughout this project.

We wish to extend our sincere thanks to all faculty members of Information Technology Department for their valuable suggestions, kind cooperation and encouragement on successful completion of this project.

Finally, we profound gratitude to the almighty for his presence and who gave me the bravery, privilege, power and belief to complete this project.

ABSTRACT

The Virtual Assistant with AI is a computer-based program that can assist its user with a variety of tasks. The program is built with Python extension files that allow it to understand and respond to user commands by utilizing the power of Artificial Intelligence (AI) and Natural Language Processing (NLP). The Virtual Assistant with AI is intended to function as an intelligent personal assistant for the user, performing a variety of tasks such as setting reminders, scheduling appointments, sending messages and much more. The program employs advanced algorithms to analayse user commands and provide accurate real-time responses. The AI-powered Virtual Assistant is a tremendously powerful tool that can help users save time and be more productive in their daily lives. Early As we all know, how life is interlinked with the technology and the use of AI. AI-powered voice assistants have become an integral part of our lives, intertwining technology and daily tasks. A Personal Virtual Assistant allows a user to command or ask questions in the same manner that they would do with another human and are even capable of doing some basic tasks like opening apps, doing Wikipedia searches without opening a browser, playing music etc, with just a voice command. This project presents the development of a personal desktop assistant using Python, aiming to provide convenience, automation, and assistance to users in their computer-related activities. The assistant incorporates features such as voice recognition, natural language processing, and integration with external APIs to enhance its functionality and user experience

TABLE OF CONTENT

LIST OF FIGURES

LIST OF ABBREVIATIONS

NLP	**N**atural **L**anguage **P**rocessing
GUI	**G**raphical **U**ser **I**nterface
VAPA	**V**oice-**A**ctivated **P**ersonal **A**ssistant
TTS	**T**ext-**T**o-**S**peech
UAT	**U**ser **A**cceptance **T**esting
NLTK	**N**atural **L**anguage **T**oolkit

CHAPTER 1

INTRODUCTION

A virtual assistant in Python is a computer program that performs tasks based on user commands or interactions. It can provide information, execute specific functions, automate processes, managedata, and interact with other applications or systems.The system utilizes various libraries, such as collect the preprocess data., wikipedia, and Python, to perform speech to text conversion, converting tasks into audio signals. The Text-to-Speech Engineconverts the text into phonemic representation, which is then output as waveforms. If your personal Al needs to learn from user interactions, consider incorporating machine learning techniques.

This could involve sentiment analysis, recommendation systems, or other personalized features. It can provide information, execute specific functions, automate processes, manage data, and interact with other applications or systems. Python's libraries and frameworks, such as and natural language processing, enable the creation of virtual assistants with voice and text-based interfacesArtificial intelligence (AI) is used by virtual assistants known as "personal assistants" to streamlineobligations and facilitate herbal interactionsArtificial intelligence (AI) is the simulation of human intelligence processes by computers, primarily computer systems. It encompasses robots, natural language processing, and machine learning.

AI gives machines the ability to learn, make choices, and communicate with people. It is used in virtual assistants, medical diagnostics, and other fields, changing businesses and everyday lives. This innovation provides a hands-free assistant experience using Pyttsx3, which redefines how users interact with technology through voice-activated Google search, minus real-time news feeds, and easy, new-age access to YouTube, WhatsApp and Gmail for technical communication.

Artificial intelligence (AI) is the simulation of human intelligence processes by computers, primarily computer systems. It encompasses robots, natural language processing, and machine learning. AI gives machines the ability to learn, make choices, and communicate with people. It is used in virtual assistants, medical diagnostics, and other fields, changing businesses and everyday lives.

The way we interact with our computers has changed dramatically thanks to desktop assistants driven by AI. These intelligent systems make use of cutting-edge automation, machine learning, and natural language processing techniques to offer consumers smooth and effective support. Users may carry out a variety of operations using voice commands or text inputs thanks to desktop assistants like Siri, Cortana, and Google Assistant. Setting reminders, sending messages, conducting web searches, scheduling appointments, controlling smart home gadgets, and more are examples of these activities.

Working with voice recognition capability is made easy and straightforward by the Python SpeechRecognition package. It enables you to translate spoken language into text so you may create software that can comprehend and analyze human speech. It is used to integrate voice recognition features into your apps. It facilitates the use of numerous voice-to-text services by supporting a variety of speech-recognition engines and APIs. pyttsx3 is a multiplatform TTS engine that works with a variety of speech engines and platforms. It doesn't need an internet connection and is simple to operate.

1.1 OVERVIEW OF PROJECT

The program includes several features, such as Fluid Google search, email functionality through Gmail integration, real- time news feed extraction, and dynamic - Thanks to weather reports, clever use of familiar systems like WhatsApp and YouTube , this Python-based AI personal assistant is a perfect example of how AI Determine what tasks and functionalities you want your personal AI to perform. This could range from simple tasks like setting reminders and answering questions to more complex tasks like playing music, sending emails, or controlling smart home devices.

SpeechRecognition For adding speech recognition capabilities.Flask or Django For building a web interface for your personal AI. Depending on your preference, you can create a graphical user interface (GUI) using frameworks like Tkinter or PyQt, or a web-based interface using Flask or Django. Test your personal AI thoroughly to identify and fix any bugs or issues. Solicit feedback from users to improve its usability and effectiveness. Ensure that your personal AI respects user privacy and implements necessary security measures, especially if it handles sensitive information or interacts with external services.

Once your personal AI is ready, you can deploy it on your local machine or a server for personal use. You can also explore options for deploying it on cloud platforms or integrating it with other devices or services. Keep refining and updating your personal AI based on user feedback and emerging technologies to enhance its capabilities and usability over time. Deploy your personal AI for personal use. This could involve running it on your local machine, hosting it on a server, or deploying it to a cloud platform. Ensure that the deployment environment meets any requirements, such as hardware specifications or network connectivity.

1.2 OBJECTIVES OF PROJECT

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 making the report in our terms of understanding . Provide a topic for research and continue with your tasks while the assistant 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 class test today. Just tell assistant in advance about your tests and she reminds you well in advance so you can prepare for the test. In this respect, the ability of personal assistants to accurately recognize spoken words is a prerequisite for them to be adopted by consumers.

To provide better services also better human machine interactions.The objective is to build a program that can perform tasks and provide information to users throughvoice commands or keyboard input.It aims to provide services like humans like managing schedules, searching online, playing music and more. Al's primary goal is to make computers act wise like humans.This means teaching them to see things, understand language, think, learn, and solve problems likeours. 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 Al assistants is missing. Users simply cannot see or touch a voice interface

To provide better services also better human machine interactions. The objective is to build a program that can perform tasks and provide information to users through voice commands or keyboard input. It aims to provide services like humans like managing schedules, searching online, playing music and more. Al's primary goal is to make computers act wise like humans.

This means teaching them to see things, understand language, think, learn, and solve problems likeours. Al assistants will continue to offer more individualized experiences as they get better at differentiating between human skills. However, it's not just developers that need to address the complexity of developing for human idea 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 Al assistants is missing. Users simply cannot see or touch a voice interface. A voice assistant, also known as an intelligent personal assistant or a connected speaker, is a new type of device that is based on natural language speech recognition and is offered by popular companies like Apple, Amazon, and Google. We got inspired by that and created one our self.

CHAPTER 2

LITERATURE SURVEY

VOICE ACTIVATED PERSONAL ASSISTANT

Author: Aarthi Eswara Moorthy , Kim Phoung L. Vu

Abstract

Voice interface is becoming a common feature in mobile devices such as tablets and smartphones. Moreover, voice recognition technology is touted to mature and become the default method to control of a variety of interfaces, including mobile devices. Thus, it is critical to understand the factors that influence the use of voice activated applications in the public domain. The present study examined how the perceived acceptability of using the Voice-Activated Personal Assistant (VAPA) in smartphones influences its reported use. Participants were U.S.

smartphone users recruited from Amazon Mechanical Turk. Results showed that participants preferred using the VAPA in a private location, such as their home, but even in that environment, they were hesitant about using it to input private or personally identifying information in comparison to more general, non-private information. Participants' perceived social acceptability of using the VAPA to transmit information in different contexts could explain these preferred usage patterns.

Prior studies, in which participants evaluated intelligent assistive systems, have mainly focused on the efficiency and user preference of voice input when compared to other methods of information input such as text entry and direct manipulation. For example, Cox, Cairns, Walton, and Lee ound that user preference was higher and information input was faster with voice input method compared to multitap or predictive text-entry when composing a text message in "hands-busy" and "eyes-busy" situations

Introduction

Voice recognition technology is developing rapidly, and it is anticipated to become not only the default input method in smartphones but also in automobiles and other home appliances such as the TV. Because of the unique characteristics associated with voice input, including the overt verbalization of commands, privacy and acceptability concerns will influence the use and adoption of voice-based human-machine interfaces.

Traditional models from the technology adoption literature, such as the Technology Acceptance Model , the Unified Theory of Acceptance and Use of Technology , and Mobile Phone Technology Acceptance Model , identify factors that determine and moderate technology usage. However, the focus of these models has been on use of stationary desktop-based software or basic mobile phone voice services. extensive research on Voice-Activated Personal Assistants (VAPA) user preferences has not been conducted.

Siri, the voice-activated intelligent personal assistant that debuted with iPhone 4S in 2011 is thought to be the first implemented application, in which “voice recognition, information management, artificial intelligence, task fulfillment, and user interface cooperate in a way the general public finds usable and productive enough to adopt on a global scale of tens of millions devices” . However, there are also many concerns with using Siri and other VAPAs of its class. Notably, many users are concerned with the propriety of public use of the VAPA in front of strangers

Methodology

VAPA usage preferences in various locations using Amazon Mechanical Turk (AMT). However, only “quality” responses from 76 participants were used for data analysis. “Quality” was defined as correctly selecting pre-specified answers for quality control questions.

A little more than half of the participants were male (55%). Most participants were below 35 years (78%), had some college education (84%) and reported being Caucasian (84%). An online survey titled “Smartphone Usage Preferences Survey” was designed using the survey development website SurveyMonkey. This generic title, which omitted mention of voice assistants, was intended to keep participants partially blind to the purpose of the study and thus prevent biased responses towards the voice assistant.

Merits

- Reducing the age sensitivity of the facial features
- The proposed models’ performance when addressing cross-age recognition in a large and challenging experimental dataset

Demerits

- Requiring significant resources in terms of time and computing power.
- Result in poor generalization to unseen data

JARVIS - AI VOICE ASSISTANT

Author:R. Sharma and A. Dwivedi

Abstract

JARVIS, a virtual embedded voice assistant that includes cutting-edge technology based on gTTS and Python in developing a personalized assistant. JARVIS integrates the functionality of AIML and, together with Google, the industry leader, a text-to-speechplatform and thus male/female voices into the gTTS libraries powered by the Marvel world. This is often the result of adopting thedynamic base Pyttsx Pythons considered wise in contiguous phases of gTTS, facilitating the establishment of essentially fine-tuneddialogues between assistants management and users. It will help end users in their daily activities like general human speech, query

search in Google, Bing or Yahoo, video search, image retrieval, live weather, word meaning, predict and remind users of scheduledevents and tasks. This is often the sole result of over-contributing by multiple contributors, such as AIML's usability and ability todynamically merge with platforms like Python [pyttsx] and gTTS [Google Text to Speech] results in the same JARVIS standardstructure showing general reusability and almost zero or no maintainability.

Introduction

AI voice assistant, also known as a virtual or digital assistant, is a device that uses voice recognition technology,natural language processing, and Artificial Intelligence (AI)to respond to people. Through technology, the device aggregates user messages, breaks them down, rates them,and gives meaningful feedback in return. Artificialintelligence can bring real conversations. Virtual assistants,understand natural language voice commands and performs tasks for users.

These tasks, previously performed by a personal assistant or secretary, include dictation, reading text messages or exchanging email messages aloud, schedule appointments for end users. The AI assistant can also perform other activities, such as sending messages, answering phone calls, and getting directions. It also helps to read news and weather updates, open Google, You Tube, Stack Overflow, etc. , answer any questions, web scraping, play mu-sic, etc.

Although this definition emphasizes the digital style of a virtual assistant, the term virtual assistant or virtual personal assistant is additionally unremarkably wontto describe contract employees United Nations agency workfrom home and perform body tasks un-remarkably performed by executives, assistant or secretary. Digital assistants can also be compared with other form of consumer-facing AI programming known as responsive advisors. Sensible adviser programs are topic-oriented, whereas virtual assistants are task-oriented.

"Virtual assistants are typically cloud-based programs that require internet-connected devices and/or applications to function". The technologies that power virtual assistants require vast amounts of knowledge, powering the platforms, as well as machine learning, language communication processes, and speech recognition arena. There are dedicated

devices to provide virtual assistance.

The most stylish on the market from Amazon, Google and Microsoft having Alexa, Google Siri and Cortana as AI voice assistants respectively given by each company. In a word, in order to clearly recognize the user's

speech command and get a response from the system, we should consider that the speech recognition system contains the whole process by which the application system directs

the generation. voice signal to text data and some important meanings, forms of speech.

Methodology

The speech recognition system is the core of the voice application system, which is capable of understanding the voice input given by the user, and at the same time operating the applications efficiently and generating voice feedback to the user. This system is an important component for users as a gateway to use their voice as an input component. As this program can also be controlled with your phone with help of an application 'WO-MIC', it just turns any android phone into a wireless microphone and helps in the reduction of unwanted noise in the environment. Using this application, which is Wikipedia's search engine, users can contact the wizard and the wizard will retrieve the data from the internet. The results are displayed in the console window in audible format, up to a limited number of line.

Merits

- This would upgrade its effectiveness and the wide range ability of producing responses.
- Further addition of more voices can also be done as an additional feature.

Demerits

- User input is speech only not in the text format.
- Accuracylevel is low and effective less.

JARVIS VOICE ASSISTANT

Author: Sai madhavi and Sudarshan Reddy R

Abstract

Developing a personal assistant is the primary objective of the project especially in view of windows-based operating system. Personal assistant namely JARVIS has been an idea inspired from existential virtual assistance like Cortana and Siri for the respective windows and iOS platforms, which carries out various tasks using the command line making it user friendly interface. Synchronous process which has speech recognition and processes and analyze the artificial speech dependent on humans as the base in this personal voice assistance, this software is capable of performing various tasks and provide services to the end user according to the requirement.

With the development of technology human dependency on the computers as increased drastically, people or users want their gadgets to be easy and more practical hence there has been a shift even in providing commands or input from text to speech. This shift has marked the beginning of speech recognition. Thereby the need for not just taking the input through speech arose even performing the task and command trough speech recognition has been the idea behind the Personal voice Assistant.

Introduction

Everything that a human being can do is being transformed by machines in this era. One of the key causes is the performance exchange. In today's society, we teach our machines to think like people and complete their missions on their own. As a result, the concept of a digital assistant emerged.

A digital assistant that understands voice commands and does relevant tasks as requested by the user using voice reputation features and language processing algorithms. A digital assistant is capable of filtering out ambient noise and returning relevant facts based on particular commands supplied by the person. Virtual Assistants are entirely software-based, but they've recently been integrated into a variety of devices, and some, such as Alexa, are created specifically for standalone devices Because of the rapid change in era, it is now more important than ever to educate our machines using machine learning, deep learning, and neural networks. With the help of Voice Assistant, we can now speak to our machines.

Nowadays, every significant company uses Voice Assistant so that their customers may use their voice to interact with their devices. As a result, with the Voice Assistant, we're progressing to the next level of development, where we'll be able to converse with our device. These types of digital assistants are extremely beneficial to the elderly, the visually and physically challenged, children, and others by ensuring that interacting with the system is no longer a difficult task for humans. Even blind people who are unable to see the system can engage with it by using their voice.

Methodology

In this section,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. The python backend parses the voice recognition module's output to determine whether the command or speech output is an API Call, Context Extraction, or System Call. The output is then sent back to the python backend to provide the user with the desired results.

The acronym for Application Programming Interface (API) is Application Programming Interface. An API is a software interface that allows two apps to communicate with one another.To put it another way, an API is the messenger that sends your request to the provider and then returns the result. The capacity of computers to read text aloud is referred to as text-to-speech (TTS). Written text is converted to a phonemic representation, which is subsequently converted to waveforms that can be generated as sound by a TTS Engine. Third-party publishers offer TTS engines in a variety of languages, dialects, and specialist vocabularies.

Merits

Future work includes Making the voice assistance to work on all the versions of the python in all the windows version and Linux versions and including some function that helps in doing payments while online shopping.

FUNDAMENTALS OF SPEECH RECOGNITION

Author: Hasini Gunasinghe And Elisa Bertino

Abstract

Speech recognition technology has become an integral part of various applications, ranging from virtual assistants to automated customer service systems. This paper provides an overview of the fundamentals of speech recognition, covering its basic concepts, underlying techniques, and applications. The process of converting speech signals into digital representations, including the use of signal processing techniques such as Fourier analysis and digital filtering. It then explores the various components of a typical speech recognition system, including feature extraction, acoustic modeling, language modeling, and decoding.

The paper also delves into the key techniques employed in speech recognition, such as Hidden Markov Models (HMMs), Gaussian Mixture Models (GMMs), and Deep Learning approaches like Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs). It highlights the strengths and limitations of each approach and discusses recent advancements in the field, such as end-to-end speech recognition systems.

Introduction

Automatic recognition of speech by machine has been a goal of research for more than four decades and has inspired such science fiction wonders as the computer HAL in Stanley Kubrick's famous movie 2001-A Space Odyssey and the robot R2D2 in the George Lucas classic Star Wars series of movies. However, in spite of the glamour of designing an intelligent machine that can recognize the spoken word and comprehend its meaning, and in spite of the enormous research efforts spent in trying to create such a machine, we are far from achieving the desired goal of a machine that can understand spoken discourse on any subject by all speakers in all environments.

Thus, an important question in this book is, What do we mean by "speech recognition by machine." Another important question is, How can we build a series of bridges that will enable us to advance both our knowledge as well as the capabilities of modern speech-recognition systems so that the "holy grail" of conversational speech recognition and understanding by machine is attained? Because we do not know how to solve the ultimate challenge of speech recognition, our goal in this book is to give a series of presentations on the fundamental principles of most modern, successful speech-recognition systems so as to provide a framework from which researchers can expand the frontier.

Instead we will provide the theoretical background and justification for each topic discussed so that the reader is able to understand why the techniques have provedvaluable and how they can be used to advantage in practical situations.

Successful speech-recognition systems require knowledge and expertise from a wide range of disciplines, a range far larger than any single person can possess. Therefore, it is especially important for a researcher to have a good understanding of the fundamentals of speech recognition (so that a range of techniques can be applied to a variety of problems), without necessarily having to be an expert in each aspect of the problem.

Methodology

Signal processing the process of extracting relevant information from the speech signal in an efficient, robust manner. Included in signal processing is the form of spectral analysis used to characterize the time-varying properties of the speech signal as well as various types of signal preprocessing (and postprocessing) to make the speech signal robust to the recording environment (signal enhancement). Physics (acoustics) the science of understanding the relationship between the physical speech signal and the physiological mechanisms (the human vocal tract mecha- nism) that produced the speech and with which the speech is perceived (the human hearing mechanism).

Pattern recognition-the set of algorithms used to cluster data to create one or more prototypical patterns of a data ensemble, and to match (compare) a pair of patterns on the basis of feature measurements of the patterns. Linguistics the relationships between sounds (phonology), words in a language (syntax), meaning of spoken words (semantics), and sense derived from meaning (pragmatics). Included within this discipline are the methodology of grammar and language parsing.

Merits

Speech recognition technology enhances accessibility for individuals with disabilities, including those with mobility impairments or visual impairments.

Demerits

Speech recognition technology raises privacy concerns related to the collection and storage of voice data.

- Speech recognition systems may struggle with recognizing specialized vocabulary or domain-specific terminology, particularly in niche applications or industries.

DESKTOP VOICE ASSISTANT

Author: Vishal Kumar Dhanraj, Lokesh kriplani, Semal Mahajan

Abstract

The main goal of Artificial intelligence (AI) is the realization of natural dialogue between humans and machines. There are many IT companies have used the dialogue systems technology to establish various kinds of Virtual Personal Assistants (VPAs) based on their applications and areas for increasing interaction between human and machine, such as Microsoft's Cortana, Apple's Siri, Amazon Alexa, Google Assistant. As like Microsoft cortana we have created our own virtual personal assistant only for windows using python which is able to access on any windows explorer such as windows 7,8,10.

We use python as a programming language because it have a major libraries which is use to execute commands. By using python installer packages our personal virtual assistant recognize the user voice and process on it. Voice assistants are the great innovation in the field of AI that can change the way of living of the people in a different manner. The voice assistant was first introduced on smartphones and after the popularity it got. It was widely accepted by all. Initially, the voice assistant was mostly being used in smartphones and laptops but now it is also coming as home automation and smart speakers.

Introduction

Virtual assistant is used to run machine like laptop or PC's on your own command. Virtual assistant is an application program that understands natural language and voice commands to complete tasks for the users. The Users can ask their assistants' questions, control home automation devices, and media player.

Nowadays virtual assistant is very useful to human. It makes human life easier like operate PC's or laptop on only voice command. Virtual assistant is a less time consuming. By using virtual assistant we saves our time and contribute in other works. Virtual assistants are typically cloud-based program that requires internet connected devices. Virtual assistant is the flexibility to contract for just the services they need. For creating virtual assistant for your computer go from basics python. Virtual assistants are task-oriented.

Virtual assistant's ability to understand and perform requests. Virtual assistants is a software that understands verbal and written commands and completes task assigned by clients. Virtual assistants are able to interpret human speech and respond via synthesized voices. There are several voice assistants in market like Siri for apple TV remote, Google Assistant for pixel XL smartphones, Alexa as a smart speaker which is developed by using Raspberry Pi, Microsoft Cortana for windows 10. As like this all virtual assistants we also created a virtual assistant for windows. We use Artificial Intelligence technology for this project. Also use python as a programming language, because python offers a good major libraries.

For thissoftware use microphone as input device to receive voice requests from user and speaker as output device to give the output voice. This process is the combination of several different technologies like voice recognition, voice analysis and language processing. Virtual assistant use Natural Processing language to match user text or voice input to executable commands. When a user give a command to personal virtual assistant to perform a task, the natural language is convertedtheaudio signals into digital signals.

Methodology

Virtual assistants use NLP to match user text or voice input to executable commands. When a user asks a question to personal assistant to perform a task, then natural language audio signal is converted into executable command or digital data that can be analyzedby the software.

The system is using Google's online speech recognition system for converting speech input to text. Through this the users can speak and obtain the text in exchange of voice input from the special corpora organized on computer network server at the information center from the microphone which is temporarily stored in the system and then sent to google cloud for speech recognition. The same text is then received and sent to the voice

assistant program.

Merits

- Voice assistants provide hands-free interaction, allowing users to perform tasks while multitasking or when their hands are otherwise occupied.
- Voice assistants enable users with disabilities to interact with computers and devices more easily, enhancing accessibility for individuals with mobility or visual impairments.

Demerits

- Voice assistants typically store recordings of interactions, raising concerns about privacy and data security, especially if these recordings are accessed by unauthorized parties or used for targeted advertising.
- While voice recognition technology has improved significantly, it is not infallible and may struggle with accents, dialects, or background noise, leading to errors or misunderstandings.

CHAPTER 3

SYSTEM ANALYSIS

3.1 EXISTING SYSTEM

Existing system implements From the above literature survey, we have inferred that all the systemsexisting predict only particular diseases namely lung disease, breast cancer, heartdisease, diabetes by implementing various algorithms on the particular datasets. After implementing various algorithms, the most accurate one is selected and it is used for prediction of disease. Sometimes, we may get confused of what algorithm to use. Also, all the systems find only the particular disease and not the disease based on the symptoms.

spacy For natural language processing (NLP) tasks such as text classification, named entity recognition, etc.TensorFlow or PyTorch For building and training machine learning models.NLTK (Natural Language Toolkit) Another option for NLP tasks.Dialogflow or Rasa Frameworks specifically designed for building conversational AI agents. Integrate NLP capabilities to understand and process user input. This involves tasks like tokenization, part-of-speech tagging, entity recognition, and sentiment analysis. Libraries like spaCy or NLTK can be helpful here. Integrate NLP capabilities to understand and process user input. This involves tasks like tokenization, part-of-speech tagging, entity recognition, and sentiment analysis. Libraries like spaCy or NLTK can be helpful here.

Disadvantages

- Don’t you think with the advancements of AI, we are becoming lazy and entering our comfort zone?Well, the answer is, yes!Artificial Intelligence (AI) is making humans lazy by automating the majority of their tasks. Humans are prone to get addicted to these advancements, posing a threat to future generations.

3.2 PROPOSED SYSTEM

We are proposing a system in an efficient way of implementing a Personal AI assistant, Speech Recognition library has many in-built functions, that will let the assistant understand the command given by user and the response will be sent back to user in idea, with Text to read or write functions.When assistant captures the idea command given by user, the under lying algorithms will convert the voice into text. And according to the keywords present in the text (command given by user), respective action will be performed by the assistant.

This is made possible with the functions present in different libraries. Also, the assistant was able to achieve all the functionalities with help of some API's. We will be sending a request, and through the API, we're getting the respective output. API's like WOLFRAMALPHA, are very helpful in performing things like calculations, making small web searches, we have libraries like Random and many other libraries, each corresponding to a different technology. We used the library OS to implement Operating System related functionalities like Shutting down a system, or restarting a system.

Advantages

- When performing a task, we frequently make mistakes. This could be related to differences in an individual's intellectual abilities. With AI-based devices, however, this is not the case
- That's the advantage of AI-powered Digital Assistance.

CHAPTER 4

SYSTEM DESIGN

4.1 SYSTEM ARCHITECTURE

Designing a comprehensive system architecture for a personal AI using Python involves several interconnected components. At the core is the Natural Language Processing (NLP) module, responsible for understanding and generating human language inputs and outputs. This module interfaces with a Knowledge Graph, which organizes and represents the user's data and information, enabling efficient retrieval and reasoning. Concurrently, a Machine Learning (ML) pipeline, comprising various algorithms such as deep learning models, handles tasks like recommendation systems, classification, and predictive analytics.

These ML models are trained on a dataset stored in a scalable database, ensuring quick access to relevant information. Additionally, a Dialog Management system orchestrates conversations, maintaining context and coherence throughout interactions. To enhance user experience, modules for Speech Recognition and Synthesis enable hands-free communication, while Computer Vision algorithms process visual data for image recognition and understanding.

A secure Authentication and Authorization layer safeguards sensitive user data, controlling access to personalized features. Comprehensive monitoring and logging systems provide insights into system performance and user interactions, facilitating continuous improvement and optimization. Throughout the architecture, modular design principles and adherence to best practices ensure maintainability, extensibility, and robustness, empowering the personal AI system to adapt and evolve alongside the user's needs and preferences.

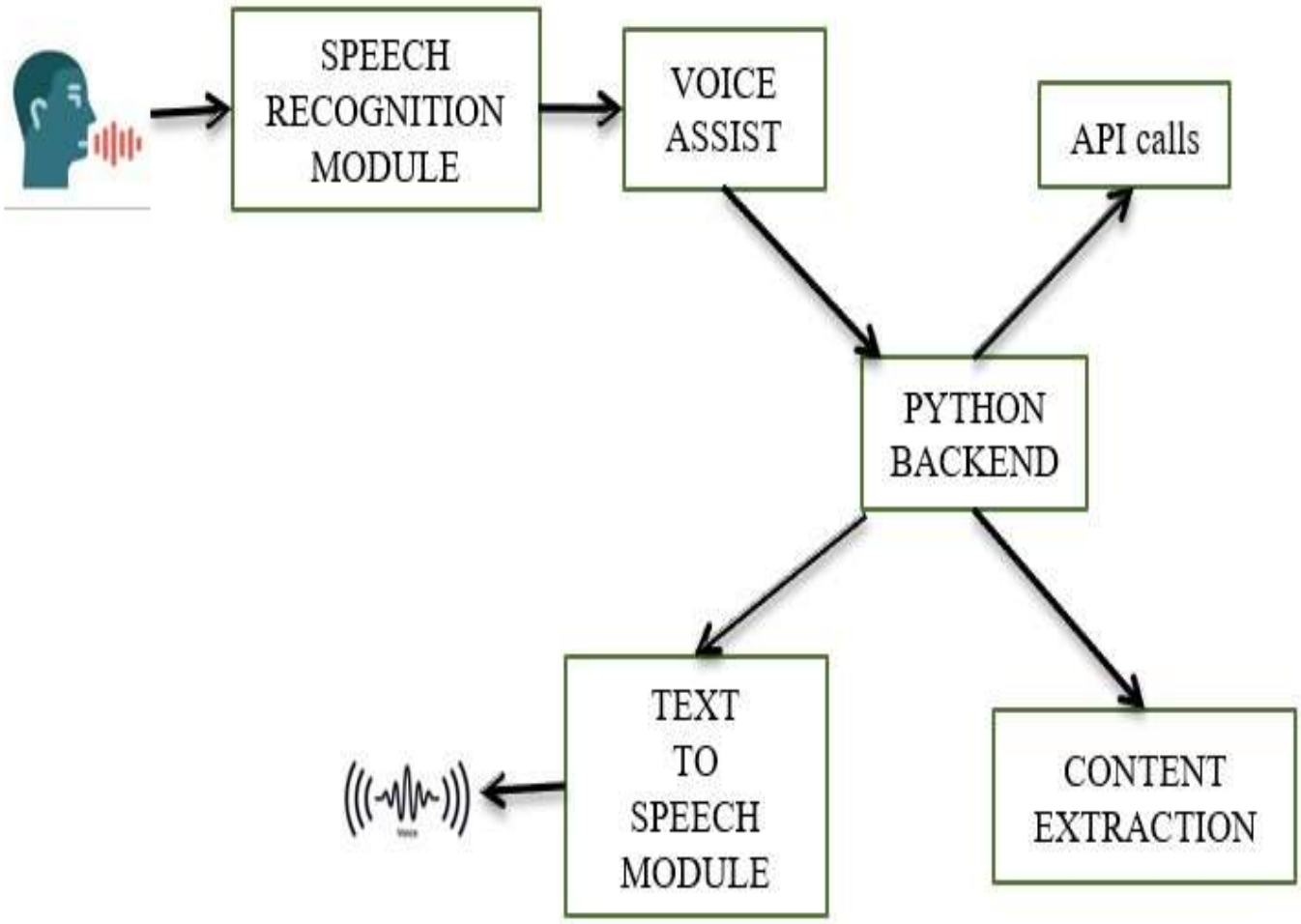

FIGURE 4.1 PROPOSED ARCHITECTURE DIAGRAM

4.2 UML DIAGRAM

4.2.1 Use Case Diagram

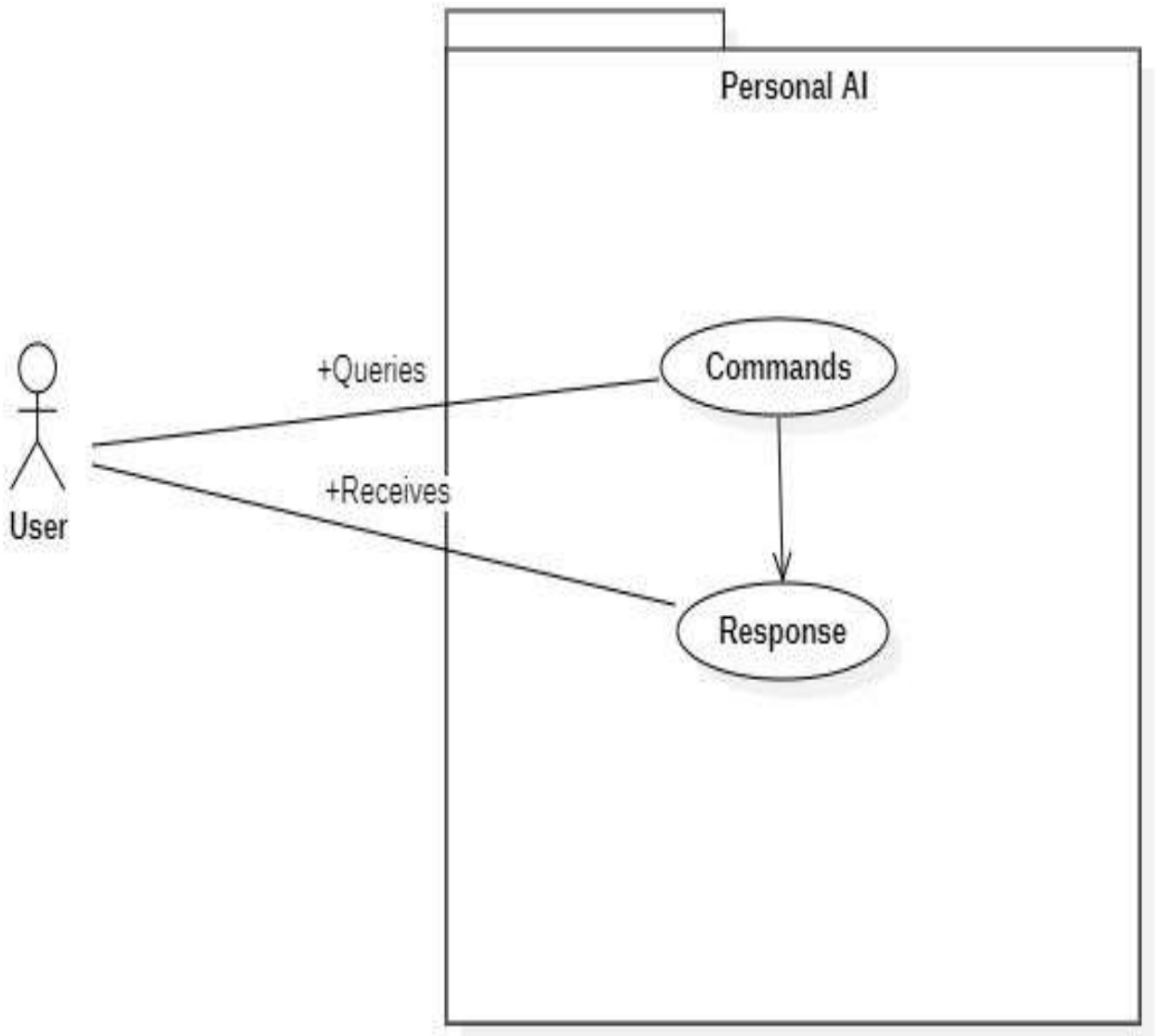

FIGURE 4.2.1 USE CASE DIAGRAM

4.2.2 Class Diagram

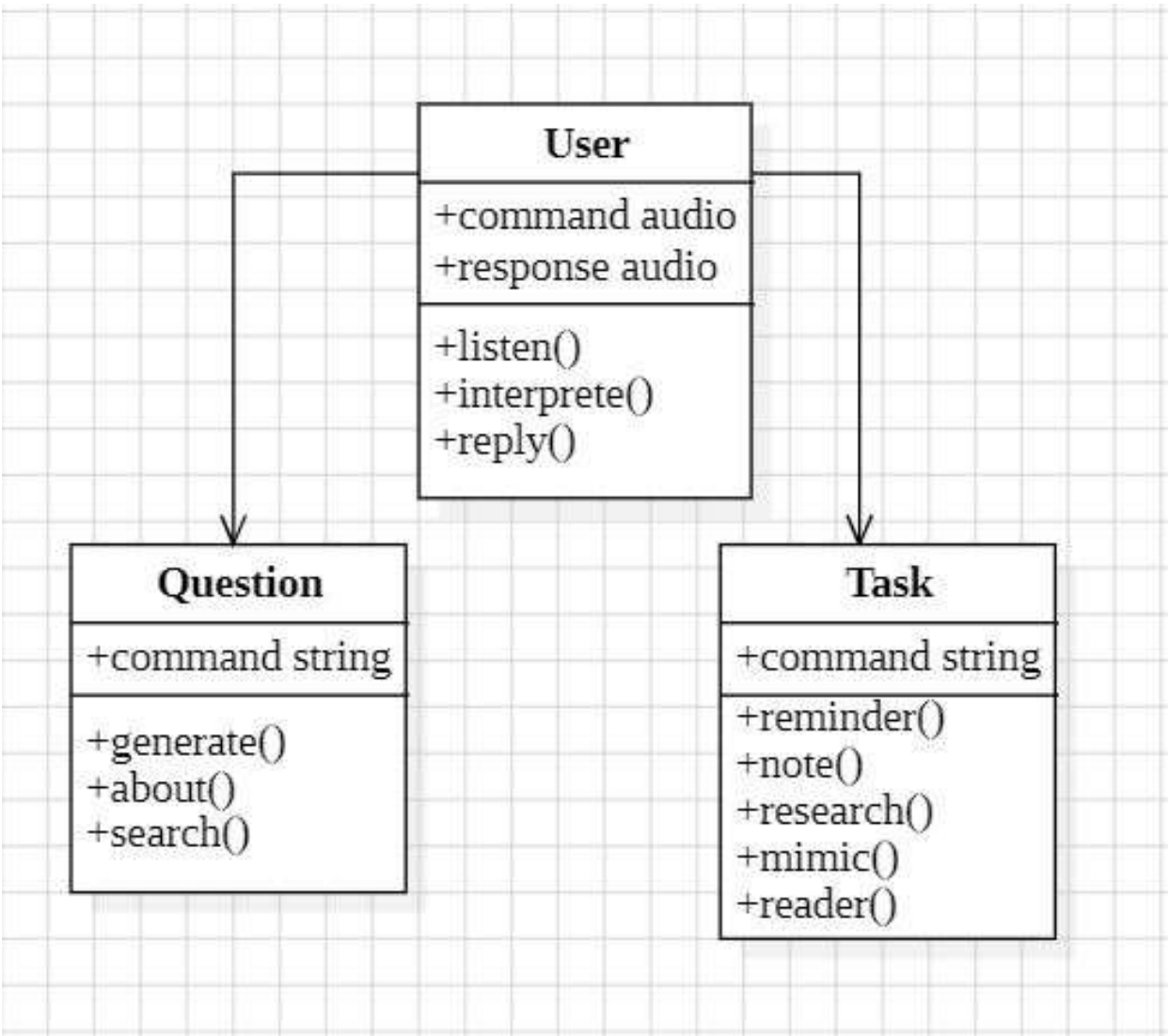

FIGURE 4.2.2 CLASS DIAGRAM

4.2.3 Sequence Diagram

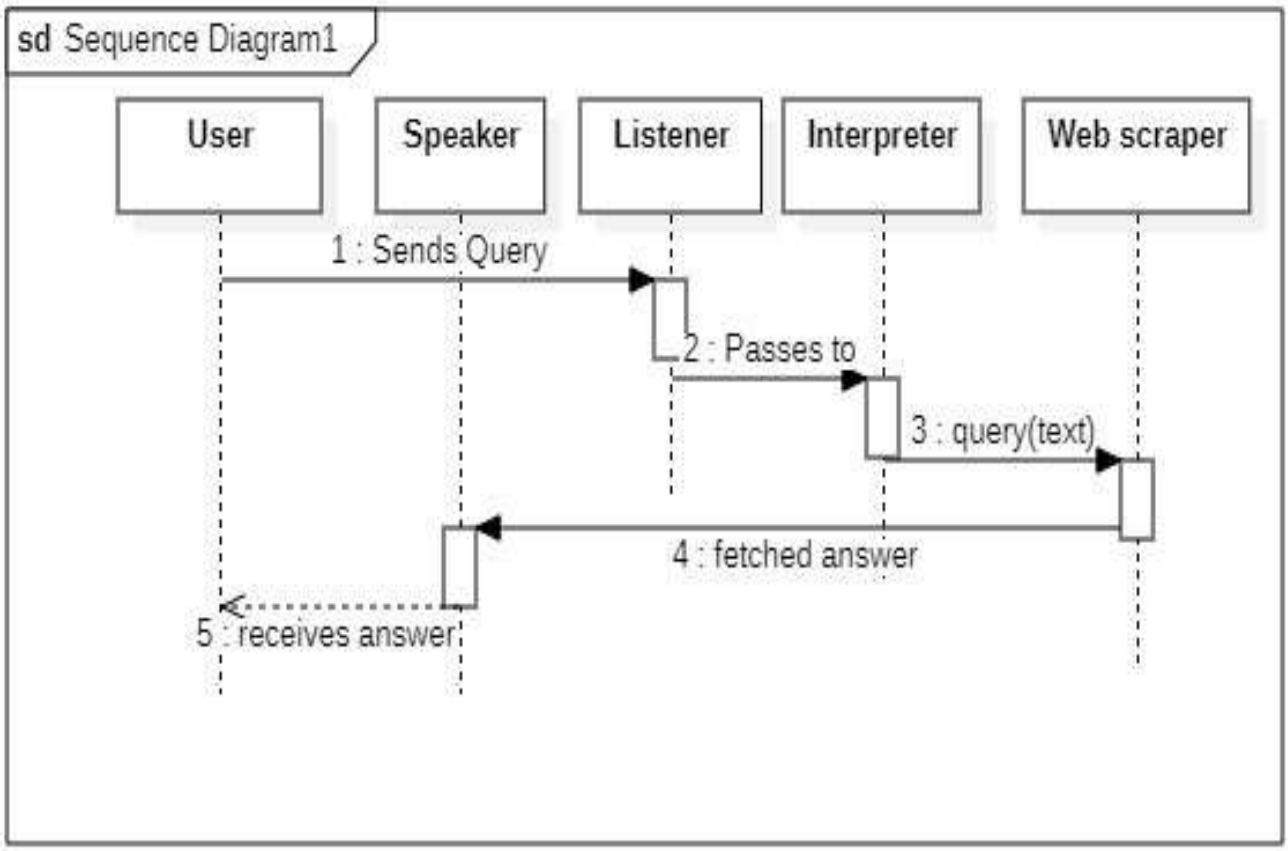

FIGURE 4.2.3 SEQUENCE DIAGRAM

4.3 DATA FLOW DIAGRAM

A data flow diagram is a depiction of the inputs and outputs of a process and the entirety of execution. . They are often elements of a formal methodology such as structured systems analysis and design method. A data-flow diagram has no control flow -there are no decision rules and no loops.

Level 0

The level 0 DFD shows how the system is divided into 'sub-systems' (processes), each of which deals with one or more of the data flows to or from an external agent, and which together provide all of the functionality of the system as a whole.

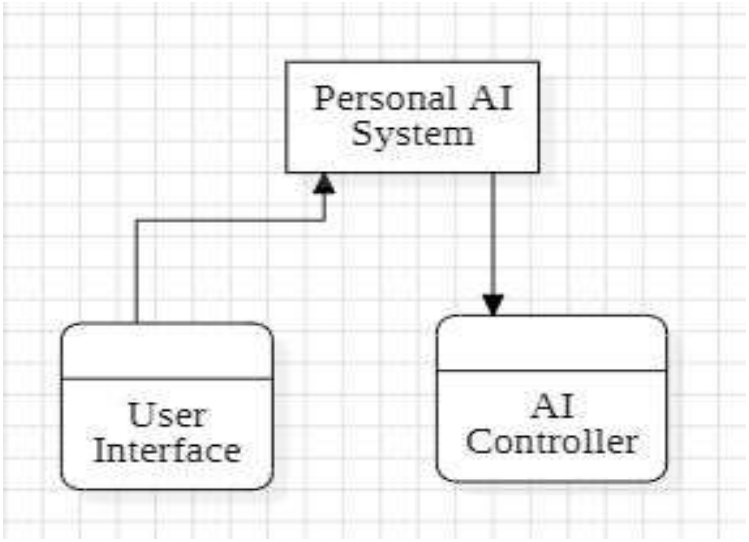

FIGURE 4.3.1 LEVEL 0

Level 1

The next stage is to create the level 1 data flow diagram. This highlights the main functions carried out by the system. As a rule, to describe the system was using between two and seven functions - two being a simple system and seven being a complicated

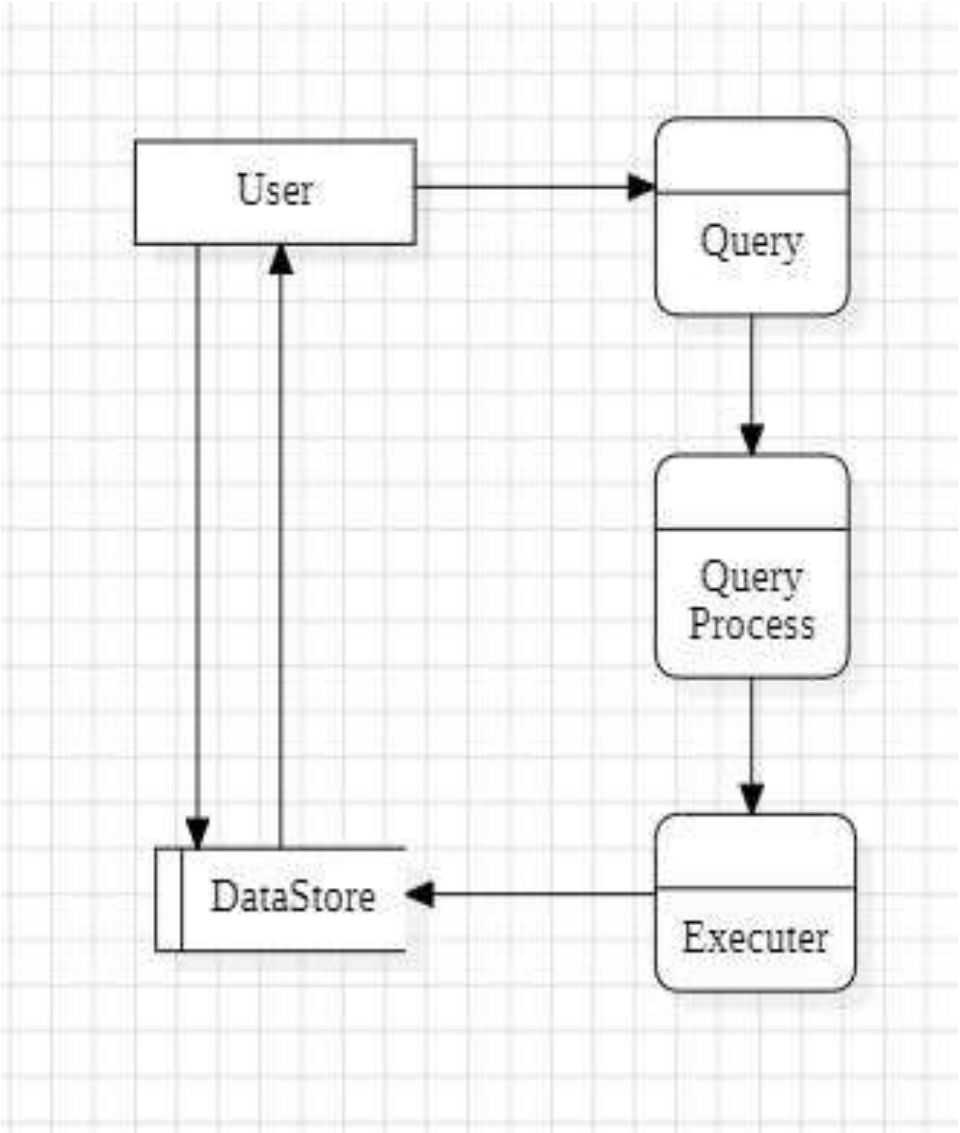

FIGURE 4.3.2 LEVEL 1

Level 2

The next stage is to create the level 2 data flow diagram. This highlights the main functions carried out by the system. As a rule, to describe the system was using between two and seven functions - two being a simple system and seven being a complicated system. This enables us to keep the model manageable on screen or paper.

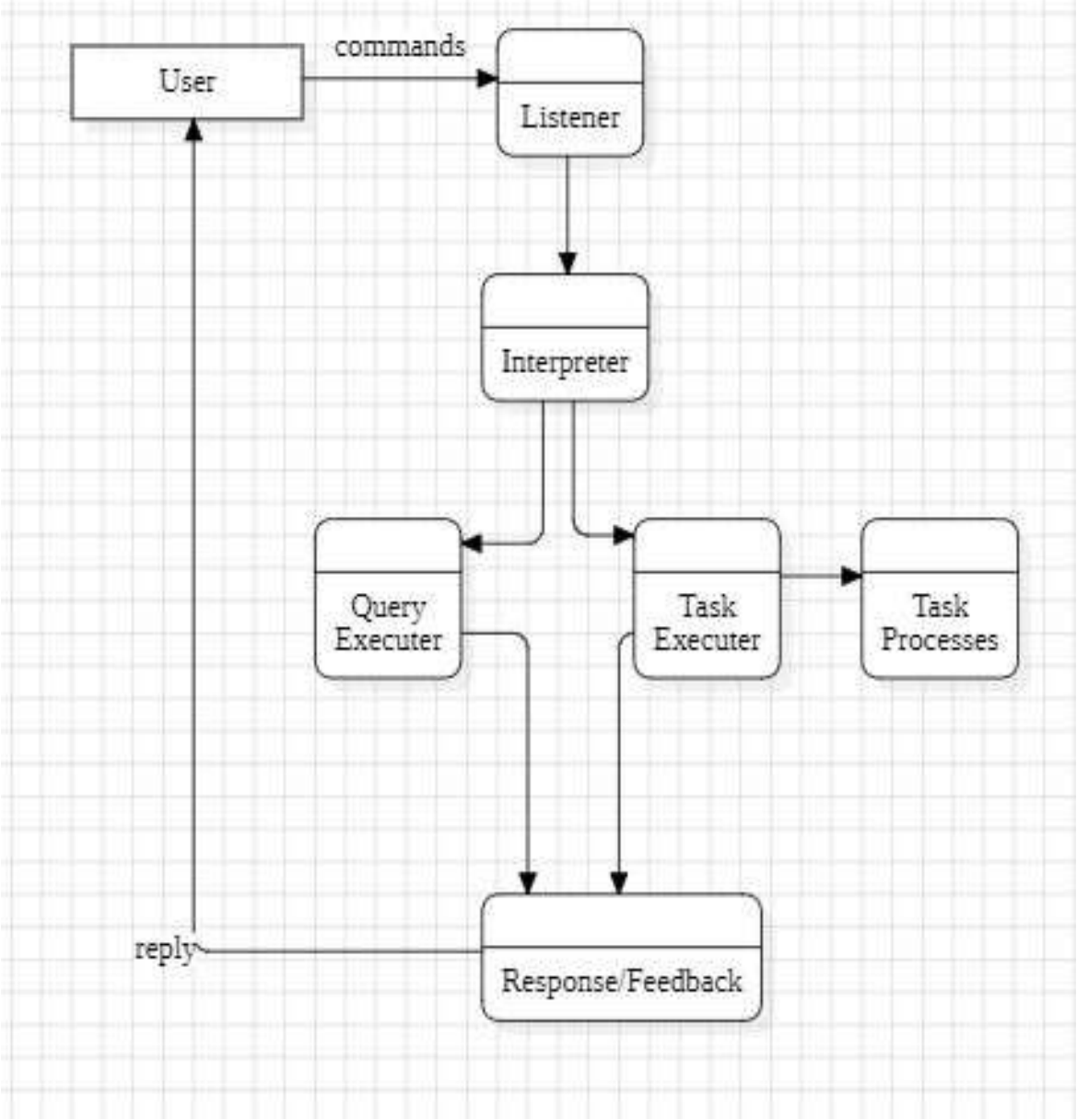

FIGURE 4.3.3 LEVEL 2

CHAPTER 5

SYSTEM REQUIREMENTS

5.1 HARDWARE REQUIREMENTS

Processor	: Dual core processor 2.60 GHZ
RAM	: 8GB
Hard disk	: 240 GB
Keyboard	: Standard keyboard
Monitor	:15 inch color monitor

5.2 SOFTWARE REQUIREMENTS

Operating system	: Windows 10
Front end	: Python
IDE	: PyCharm

CHAPTER 6

SYSTEM IMPLEMNETATION

6.1 MODULES

- Text Preprocessing Module
- Model Training and Evaluation Module
- Dialog Management Module
- User Interface Module
- Data Storage and Retrieval Module

6.2 MODULE DESCRIPTION

6.2.1 Text Preprocessing Module:

This module could include functions for cleaning and preprocessing text data, such as removing punctuation, stopwords, and performing tokenization.

6.2.2 Model Training and Evalution Module:

This module could provide functions for training and evaluating machine learning models for various tasks, such as sentiment analysis, text classification, named entity recognition, etc.

6.2.3 Model Training and Evalution Module:

This module could handle dialog management functionalities, such as understanding user queries, generating responses, and managing conversation flow.

6.2.4 User Interface Module:

This module could include functions for creating a user interface for interacting with the personal AI system, such as a command-line interface (CLI) or a graphical user interface (GUI).

6.2.5 Data Storage and Retrieval Module:

This module could handle storing and retrieving data used by the personal AI system, such as user preferences, training data for machine learning models, or logs of interactions.

CHAPTER 7

ALGORITHM DESCRIPTION

These algorithms enable machines to understand and generate human language, revolutionizing our interactions with technology. Virtual assistants like Siri and Google Assistant use NLP to interpret voice commands and provide responses, making human-computer interaction more natural.

NATURAL LANGUAGE PROCESSING

User Input: Voice command like open youtube, open camera, and open wikipedia etc.

User Output: Youtube will open in browser or Camera open take photo automatically

Steps:

1) User send a command in voice to the AI.
2) AI listening the command.
3) Analyse the command with different perspective.
4) The command analyse in syntax format.
5) Analyse the command sentences or paragraph.
6) After analysing the command and execute the task.
7) output in user interface.

CHAPTER 8

TESTING & MAINTENANCE

7.1 TESTING APPROACHES

After a system has been verified, it needs to be thoroughly tested to ensure that every component of the system is performing in accordance with the specific requirements and that it is operating as it should including when the wrong functions are requested or the wrong data is introduced.

7.1.2 Unit Testing

The first test in the development process is the unit test. The source code is normally divided into modules, which in turn are divided into smaller units called units. Write unit tests for individual components/modules using Python testing frameworks like unittest or pytest. Test edge cases, boundary conditions, and expected behaviors. Ensure proper exception handling and error messages.

7.1.3 End-to-End Testing

Test the entire system's functionality from user input to output. Create test scenarios that mimic real-world usage. Automate end-to-end tests using tools like Selenium for web interfaces or Appium for mobile interfaces.

7.1.4 User Acceptance Testing (UAT)

Involve real users to test the system. Gather feedback on usability, intuitiveness, and functionality. Iterate based on user feedback.

7.1.5 Performance Testing

Test the system's performance under different loads. Measure response times, resource consumption, and scalability. Use tools like Locust or JMeter for load testing.

7.1.6 Security Testing

Perform security audits to identify vulnerabilities. Test for common security issues such as injection attacks, data leaks, and authentication bypasses. Utilize security testing tools like OWASP ZAP or Burp Suite.

CHAPTER 9

CONCLUSION AND FUTURE WORK

CONCLUSION

Developing a personal AI using Python presents a multifaceted journey of learning and innovation. Throughout this project, you've delved into various aspects of artificial intelligence, honing your skills in natural language processing, machine learning, and data manipulation. By customizing its functionalities, you've tailored it to suit your specific needs, whether it's assisting with daily tasks, providing information, or even offering companionship. Along the way, you've encountered challenges that have sharpened your problem-solving abilities and creativity.

Ethical considerations have guided your development process, ensuring privacy, transparency, fairness, and accountability. Engaging with the broader AI community has provided valuable insights and collaboration opportunities, enriching the development experience. Looking ahead, you're poised to explore advanced AI techniques and integrate with emerging technologies, further extending the capabilities of your personal AI and showcasing the transformative potential of artificial intelligence in everyday life.

FUTURE WORK

In future work, advanced natural language understanding, multi-modal interaction support, personalization, integration with external services, emotion recognition, privacy measures, voice assistant integration, and continuous learning. These improvements aim to make the personal AI more versatile, empathetic, and capable of adapting to user needs securely.

CHAPTER 10

APPENDICES

A1.CODING

```
import speech_recognition as sr
import pyttsx3
import datetime
import wikipedia
import webbrowser
import os
import time
import subprocess
from ecapture import ecapture as ec
import wolframalpha
import json
import requests
print('Loading your AI personal assistant - G One')
Engine=pyttsx3.init('sapi5')
voices=engine.getProperty('voices')
engine.setProperty('voice','voices[0].id')
def speak(text):
   engine.say(text)
   engine.runAndWait()
          elif "log off" in statement or "sign out" in statement:
                speak("Ok , your pc will log off in 10 sec make sure you exit
        from all application)
```

```
def wishMe():
    hour=datetime.datetime.now().hour
    if hour>=0 and hour<12:
        speak("Hello,Good Morning")
        print("Hello,Good Morning")
    elif hour>=12 and hour<18:
        speak("Hello,Good Afternoon")
        print("Hello,Good Afternoon")
    else:
        print("Hello,Good Evening")
def takeCommand():
    r=sr.Recognizer()
    with sr.Microphone() as source:
        print("Listening...")
        audio=r.listen(source)
try:
            statement=r.recognize_google(audio,language='en-in')
            print(f"user said:{statement}\n")
 except Exception as e:
            speak("Pardon me, please say that again")
            return "None"
        return statementspeak("Loading your AI personal assistant
G-One")
wishMe()
f __name__=='__main__':
while True:
        speak("Tell me how can I help you now?")
```

```
statement = takeCommand().lower()
        if statement==0:
            continue
if "good bye" in statement or "ok bye" in statement or "stop" in statement:
            speak('your personal assistant G-one is shutting down,Good bye')
            print('your personal assistant G-one is shutting down,Good bye')
            break
 if 'wikipedia' in statement:
            speak('Searching Wikipedia...')
            statement =statement.replace("wikipedia", "")
            results = wikipedia.summary(statement, sentences=3)
            speak("According to Wikipedia")
            print(results)
            speak(results)
elif 'open youtube' in statement:
            webbrowser.open_new_tab("https://www.youtube.com")
            speak("youtube is open now")
            time.sleep(5)
elif 'open google' in statement:
            webbrowser.open_new_tab("https://www.google.com")
            speak("Google chrome is open now")
            time.sleep(5)
elif 'open gmail' in statement:
            webbrowser.open_new_tab("gmail.com")
            speak("Google Mail open now")
            time.sleep(5)
```

```
elif "weather" in statement:
        api_key="8ef61edcf1c576d65d836254e11ea420
base_url="https://api.openweathermap.org/data/2.5/weathr?"speak("whats the city
name")
        city_name=takeCommand()
complete_url=base_url+"appid="+api_key+"&q="+city_name
        response = requests.get(complete_url)
        x=response.json()
        if x["cod"]!="404":
            y=x["main"]
            current_temperature = y["temp"]
            current_humidiy = y["humidity"]
            z = x["weather"]
            weather_description = z[0]["description"]
            speak(" Temperature in kelvin unit is " +str(current_temperature)
+"\n humidity in percentage is"+
elif 'time' in statement:
strTime=datetime.datetime.now().strftime("%H:%M:%S")
speak(f"the time is {strTime}")
    elif 'who are you' in statement or 'what can you do' in statement:
 speak('I am G-one version 1 point O your persoanl assistant. I am
programmed to minor tasks like','opening youtube,google chrome,gmail
and stackoverflow ','in different cities , get top headline news from times
of india and you can ask me computational or geographical questions too!')
```

```
 speak('I am G-one version 1 point O your persoanl assistant. I am
programmed to minor tasks like','opening youtube,google chrome,gmail and
stackoverflow ,predict time,take a photo,search wikipedia,predict weather','in
different cities , get top headline news from times of india and you can ask me
computational or geographical questions too!')
    elif "who made you" in statement or "who created you" in statement or
"who discovered you" in statement:
            speak("I am built by BATCH-16")
            print("I am built by BATCH-16")
     elif "open stackoverflow" in statement:
            webbrowser.open_new_tab("https://stackoverflow.com/login")
            speak("Here is stackoverflow")
             elif 'ask' in statement:
                    speak('I can answer to computational and geographical questions and
             what question do you want to ask now')
                    question=takeCommand()
                    app_id="R2K75H-7ELALHR35X"
                    client = wolframalpha.Client('R2K75H-7ELALHR35X')
                    res = client.query(question)
                    answer = next(res.results).text
                    speak(answer)
                    print(answer)
            time.sleep(3)
             print(answer)
            time.sleep(3)
```

```
    elif 'news' in statement:
        news =
webbrowser.open_new_tab("https://timesofindia.indiatimes.com/home/
headlines")
        speak('Here are some headlines from the Times of India,Happy
reading')
        time.sleep(6)

    elif "camera" in statement or "take a photo" in statement:
        ec.capture(0,"robo camera","img.jpg")

    elif 'search'  in statement:
        statement = statement.replace("search", "")
webbrowser.open_new_tab(statement)
        time.sleep(5)

    elif 'ask' in statement:
        speak('I can answer to computational and geographical
questions and what question do you want to ask now')
        question=takeCommand()
        app_id="R2K75H-7ELALHR35X"
        client = wolframalpha.Client('R2K75H-7ELALHR35X')
        res = client.query(question)
        answer = next(res.results).text
        speak(answer)
```

A2. OUTPUT SCREENSHOT

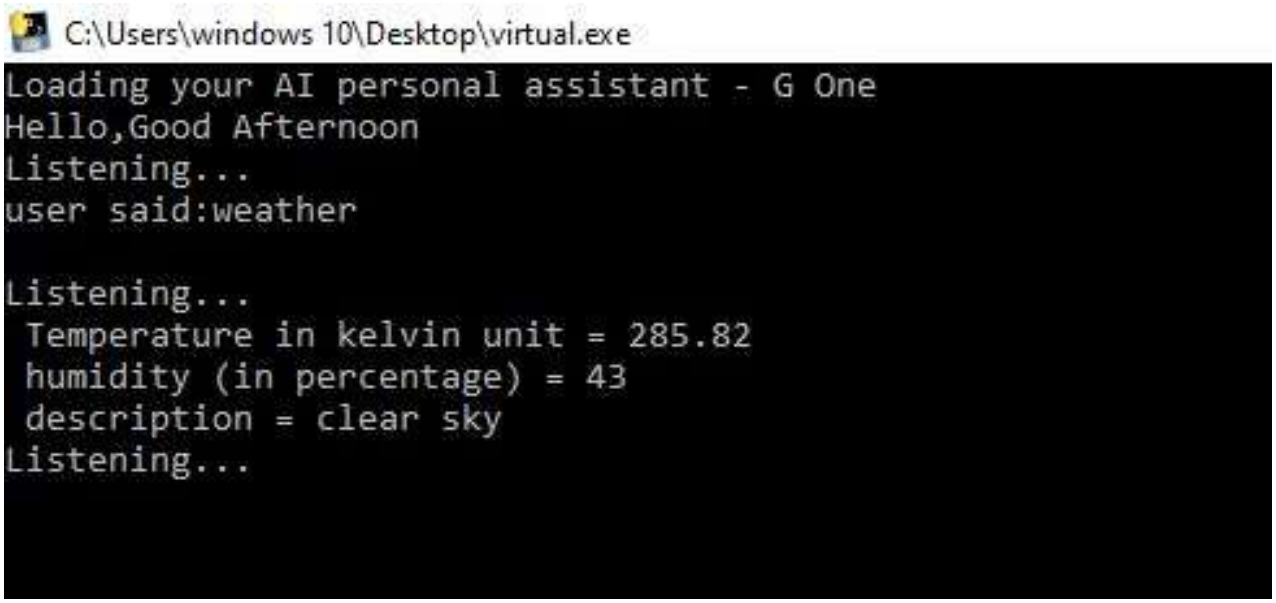

FIG A2.1 CLI Command input

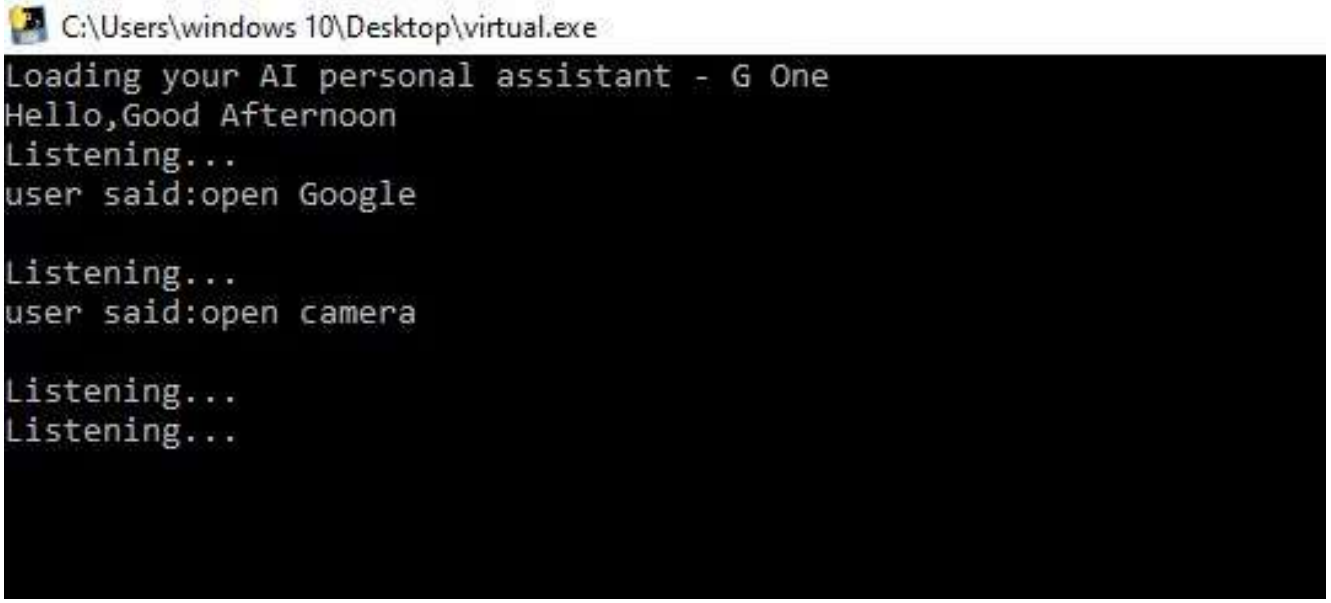

FIG A2.2 CLI Command input

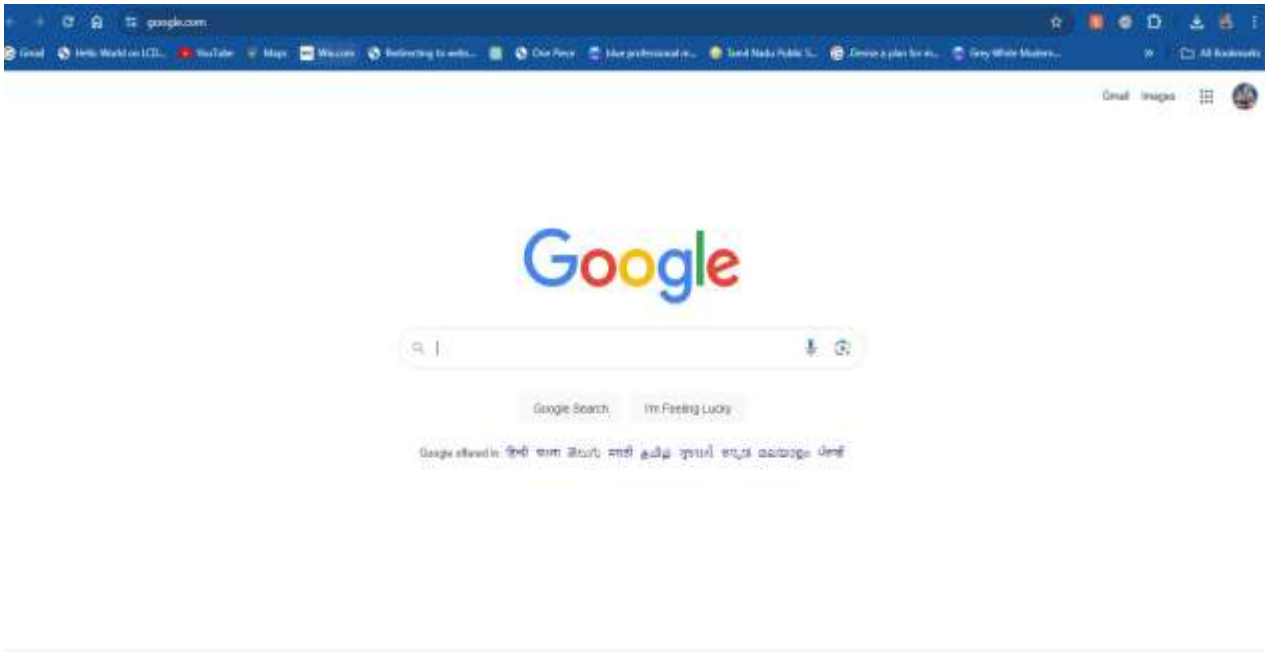

FIG A2.3 Output in browser

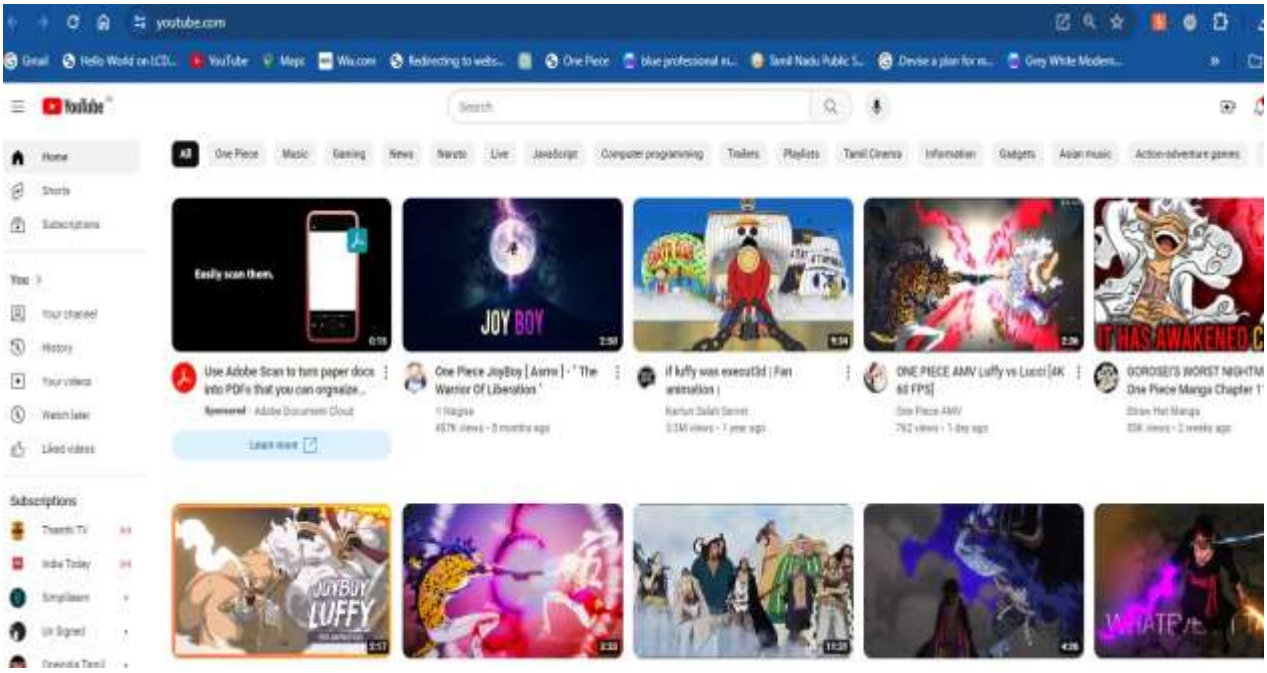

FIG A2.4 Output in browser

REFERENCES

[1] Krishnaraj P,Mohamed Faris J,Rajesh D "Portable Voice Recognition with GUI Automation" 2021.

[2] Rabiner Lawrence, Juang Bing-Hwang"Fundamentals of Speech Recognition Prentice Hall", New Jersey, 1993.

[3] Rani Y, Ms. G. Kaur, H. Rana, . S., and .N., "JARVIS: A Virtual Assistant," Int J Res Appl Sci Eng Technol, vol. 11, no. 2,, Feb.2023

[4] Sai Madhavi D, Sudarshan Reddy R,Shivkumar Made, Siddesh Godinal, Harshit B, 'Jarvis Voice Assistance', June 2022.

[5] Sharma and Dwivedi, "'JARVIS'-AI Voice Assistant," International Journal of Science and Research,2021.

[6] Shrutika Khobragade, "Jarvis, Digital Life Assistant.", Department of Computer, Vishwakarma Vishwakarma Institute of Information Technology, India, 2020.

[7] Sprengholz, Philipp & Betsch, Cornelia "Ok Google: Using virtual assistants for data collection in psychological and behavioral research" Behavior Research Methods, 2021

[8] Vishal Kumar Dhanraj, Lokesh kriplani, Semal Mahajan, "Research Paper on Desktop Voice Assistant" International Journal of Research in Engineering and Science,5 ebruary 2022

[9] Venkatesan V.K. Ramakrishna M.T.; Batyuk A.; Barna A.; Havrysh B. High-Performance Artificial Intelligence Recommendation of Quality Research Papers Using Effective Collaborative Approach. 2021

[10] Yogendra .Narayan Prajapati, U. Sesadri, Mahesh T. R., Shreyanth S., Ashish Oberoi, & Khel Prakash Jayant. (2022). Machine Learning Algorithms in Big Data Analytics for Social Media Data Based Sentimental Analysis. International Journal of Intelligent Systems and Applications in Engineering, 2022