D Y PATIL
INTERNATIONAL
UNIVERSITY
AKURDI PUNE

Leah: A NLP Powered Voice Assistant Streamlining Your Daily Tasks

Project report

Submitted to D Y Patil International University, Akurdi, Pune in partial fulfillment of full-time degree.

BTech Computer Science and Engineering

Artificial Intelligence and Machine Learning

Submitted By:

Raj Kansal 20190802013

Aditya Yadav 20190802060

Sidhant Subedhi 20190802035

Under the Guidance of

Ms. Nikita Pagar

Department of Computer Science and Engineering

D Y Patil International University, Akurdi,Pune, INDIA, 411044

Session 2019-23

Leah: A NLP Powered Voice Assistant Streamlining Your Daily Tasks

Project report

Submitted to D Y Patil International University, Akurdi, Pune in partial fulfilment of full-time degree.

BTech Computer Science and Engineering

Artificial Intelligence and Machine Learning

Submitted By:

Raj Kansal	20190802013
Aditya Yadav	20190802060
Sidhant Subudhi	20190802035

Under the Guidance of

Ms. Nikita Pagar

Department of Computer Science and Engineering

D Y Patil International University, Akurdi,Pune, INDIA, 411044

Session 2019-23

D Y PATIL
INTERNATIONAL
UNIVERSITY
AKURDI PUNE

CERTIFICATE

This is to certify that the project entitled **Leah: A NLP Powered Voice Assistant Streamlining your Daily Tasks** submitted by:

Raj Kansal	20190802013
Aditya Yadav	20190802060
Sidhant Subudhi	20190802035

is the partial fulfillment of the requirements for the award of degree of Bachelor of Technology in Computer Science and Engineering is an authentic work carried out by them under my supervision and guidance.

20/6/23

Ms. Nikita Pagar
(Supervisor)

22.06.23

Director
Dr. Bahubali Shiragapur
School of Computer Science Engineering & Applications
D Y Patil International University, Akurdi
Pune, 411044, Maharashtra, INDIA

D Y PATIL INTERNATIONAL UNIVERSITY
AKURDI, PUNE-44.

DECLARATION

We, hereby declare that the following report which is being presented in the Major Project entitled as **Leah: A NLP Powered Voice Assistant Streamlining your Daily Tasks** is an authentic documentation of our own original work to the best of our knowledge. The following project and its report in part or whole, has not been presented or submitted by us for any purpose in any other institute or organization. Any contribution made to the research by others, with whom we have worked at D Y Patil International University, Akurdi, Pune or elsewhere, is explicitly acknowledged in the report.

Raj Kansal	20190802013	
Aditya Yadav	20190802060	
Sidhant Subudhi	20190802035	Sidhant Subudhi

ACKNOWLEDGEMENT

With due respect, we express our deep sense of gratitude to our respected guide and coordinator Ms. Nikita Pagar, for her valuable help and guidance. We are thankful for the encouragement that she has given us in completing this project successfully.

It is imperative for us to mention the fact that the report of major project could not have been accomplished without the periodic suggestions and advice of our project supervisor mentor Ms. Nikita Pagar.

We are also grateful to our respected Director, Dr. Bahubali Shiragapur and Hon'ble Vice Chancellor, DYPIU, Akurdi, Prof. Prabhat Ranjan for permitting us to utilize all the necessary facilities of the college.

We are also thankful to all the other faculty, staff members and laboratory attendants of our department for their kind cooperation and help. Last but certainly not the least; we would like to express our deep appreciation towards our family members and batch mates for providing support and encouragement.

Raj Kansal (20190802013)
Aditya Yadav (20190802060)
Sidhant Subudhi (20190802035)

Abstract

Leah is a natural language processing (NLP) powered voice assistant designed to streamline daily tasks. With its advanced speech recognition and deep learning capabilities, Leah offers a user-friendly and a multilingual approach to interact with humans. In this project, Leah will be implemented on a Raspberry Pi, allowing users to access the assistant on a low-cost, portable platform. Leah's main objective is to effortlessly fit into users' habits, streamlining work and boosting output. It has a wide range of functions, including the ability to send messages, read news, answer inquiries, provide weather updates, and operate smart home devices. The system will be developed using cutting-edge technologies such as Python+, NLTK, Raspberrypi 3. With its customizable features and ability to handle a wide range of tasks, Leah promises to be a valuable addition to any individual or organization looking to boost productivity and streamline their daily workflows. By leveraging the power of the Raspberry Pi and GoogleTTS, users can have a high-performing voice assistant at their fingertips anytime and anywhere.

Keywords: *Python+, Deep Learning, Speech Recognition, NLTK, Raspberry Pi, GoogleTTS, APIs, Tokenization, GUI, Lemmatization*

TABLE OF CONTENTS

List of Figures

List of Tables

1. Project Plan

Schedule		Date	Project Activity
February	1st week	01/02/2023	Project Topic Selection
	2nd week	06/02/2023	Synopsis Submission
	3rd week	13/02/2023	Presentation on project ideas
	4th week	20/02/2023	Submission of Literature Survey & Feasibility
March	1st week	01/03/2023	Improvising the wake word functionality
	2ndweek	06/03/2023	Study of various research papers
	3rd week	13/03/2023	Study and creation of the Intent Engine
	4th week	20/03/2023	Making different skill modules for our voice assistant
	5th week	27/03/2023	Integration of the modules with the program
April	1st week	03/04/2023	Documentation and creation of report along with diagrams
	2nd week	10/04/2023	Discussion on 2nd stage project report
	3rd week	17/04/2023	Stage 2 Project report presentation, submission and evaluation

	4th week	24/04/2023	Working on the project as per the feedback provided and improving model
May	1st week	01/05/2023	Improving the frequency of the mike and wake word detector
	2nd week	08/05/2023	Starting with the addition of new skills to our voice assistant
	3rd week	15/05/2023	Added skills like time, weather, news, shazam
	4th week	22/05/2023	Started working on blackbook and research paper
	5th week	29/05/2023	Discussion on the correction and proper formats of blackbook and research paper
June	1st week	05/06/2023	Adding more skills and improving the raspberry pi model
	2nd Week	12/06/2023	Blackbook and Presentation preparation under mentor's guidance
	3rd Week	19/06/2023	Blackbook submission and Final presentation

Table 1.1: Project Plan

2. INTRODUCTION

In an era where technology is seamlessly integrated into our daily lives, voice assistants have emerged as a powerful and ubiquitous presence. These virtual entities, powered by artificial intelligence (AI) and natural language processing (NLP), have revolutionized the way we interact with our devices and access information. From smartphones to smart speakers and a myriad of other connected devices, voice assistants have become our digital companions, ready to assist us at any moment.

Voice assistants, such as Amazon's Alexa, Apple's Siri, Google Assistant, and Microsoft's Cortana, have rapidly gained popularity and are now an integral part of many households and businesses. These intelligent systems have the ability to understand and respond to human commands and queries, effectively bridging the gap between humans and machines. They have made our lives more convenient, enabling us to perform various tasks hands-free, from setting reminders and playing music to controlling smart home devices and even making purchases.

The advancements in voice recognition technology have been instrumental in the success of voice assistants. Through the use of sophisticated algorithms and machine learning, these assistants can accurately interpret and understand human speech, adapting to different accents, dialects, and languages. This has significantly improved their usability and made them accessible to a broader user base, transcending geographical and cultural boundaries.

Voice-enabled devices are increasingly being integrated into cars, healthcare systems, and various other sectors, allowing for hands-free interaction and a more intuitive user interface. This integration has created new opportunities for innovation and has paved the way for the concept of the smart home, where voice assistants act as the central hub for controlling and managing connected devices.

However, as voice assistants become more integrated into our lives, concerns surrounding privacy, security, and data protection have also surfaced. The collection and storage of personal data by these assistants raise important questions regarding user privacy and the ethical use of data. Striking the right balance between convenience and privacy will be crucial as we navigate the future of voice assistants.

In this report, we will be building our own voice assistant named "Leah" and try to explore understand its capabilities, and impact , examining its potential applications across various sectors. We will delve into the underlying technologies that power these virtual entities and discuss the challenges and opportunities they present. Furthermore, we will analyze the ethical considerations surrounding the use of voice assistants and propose recommendations to ensure their responsible deployment.

2.1. Problem statement

The problem that the voice assistant project using NLP aims to address is the need for a more efficient and natural way for users to interact with technology. Current methods, such as typing on a keyboard or clicking on a mouse, can be time-consuming and cumbersome, especially for individuals with disabilities or those who prefer a more hands-free approach. Also one more problem that we'll be identifying in this project is that many voice assistants are not multilingual by default, so here we'll be working on an approach to make our voice assistant multilingual. By leveraging NLP techniques to develop a voice assistant, users can communicate with technology in a more intuitive and convenient way, using natural language commands and receiving spoken responses. Therefore, this project seeks to overcome these obstacles to create a robust and user-friendly voice assistant that can improve productivity and enhance the overall user experience.

3. LITERATURE REVIEW

3.1. Literature review

[1]**Personal Voice Assistant Security and Privacy—A Survey — 10.1109/JPROC.2022.3153167**. The paper surveys the state of the art in security and privacy of personal voice assistants (PVAs). PVAs are increasingly used as interfaces to digital environments, and as a result, security and privacy have become a major concern. The paper discusses a number of security and privacy challenges, including:

Voice authentication attacks: These attacks attempt to bypass voice authentication mechanisms by using synthetic or recorded voices. Acoustic Denial of Service (DoS) attacks: These attacks attempt to overwhelm PVAs with noise or other unwanted sounds, preventing them from functioning properly. Hidden voice commands: These attacks attempt to trick PVAs into executing commands without the user's knowledge or consent. Privacy implications of data collection: PVAs collect a large amount of data about users, including their voice recordings, location data, and other personal information. This data could be used to track users, target them with advertising, or even be used for identity theft. The paper also discusses a number of countermeasures that have been proposed to address these challenges. These countermeasures include:

Improved voice authentication mechanisms: These mechanisms can make it more difficult for attackers to bypass voice authentication. Acoustic DoS mitigation techniques: These techniques can help to prevent or mitigate acoustic DoS attacks. Defenses against hidden voice commands: These defenses can help to prevent or detect hidden voice commands. Data minimization and privacy-preserving data collection techniques: These techniques can help to protect user privacy while still collecting the data that PVAs need to function properly. The paper concludes by discussing the future of security and privacy of PVAs. The authors believe that the challenges in this area are significant, but that they are also solvable. They call for continued research in this area to develop new and improved security and privacy mechanisms for PVAs.

In conlusion, the paper suggests- Security and privacy are major concerns for personal voice assistants (PVAs). A number of security and privacy challenges exist, including voice authentication attacks, acoustic DoS attacks, hidden voice commands, and privacy implications of data collection. A number of countermeasures have been proposed to address these challenges. The authors believe that the challenges in this area are significant, but that they are also solvable. They call for continued research in this area to develop new and improved security and privacy mechanisms for PVAs.

[2]**"Why Do People Use Artificial Intelligence (AI)-Enabled Voice Assistants?" — 10.1109/TEM.2021.3117884** This paper examines the various consumption values associated with the use of voice assistants. The authors conducted a mixed-methods study that included interviews with experts and consumers, as well as a survey of active users of voice assistants. The study found that five consumption values—social identity, convenience, personification, perceived usefulness, and perceived playfulness—were all positively associated with the use of voice assistants. The authors also found that trust and frequency of use significantly moderated the association between usefulness and usage of voice assistants. The findings of this study can be used by technology providers and marketers to develop various voice-enabled applications and services that enhance the consumer experience and consumer engagement.

In conclusion the paper suggests- Social identity and personification are two of the strongest predictors of the use of voice assistants. Perceived usefulness and perceived playfulness are also important predictors of the use of voice assistants. Trust and frequency of use moderate the association between usefulness and usage of voice assistants. These findings suggest that technology providers and marketers can increase the use of voice assistants by designing them to appeal to consumers' social identity, personification, perceived usefulness, and perceived playfulness. They can also increase the use of voice assistants by building trust with consumers and encouraging them to use voice assistants more frequently.

[3]**Deep Spoken Keyword Spotting: An Overview — 10.1109/ACCESS.2021.3139508** Key Word Spotting is a speech recognition task that aims to detect the presence of pre-defined keywords in an audio stream. Deep KWS systems have become increasingly popular in recent years due to their superior performance compared to traditional KWS systems.

The paper discussed about the ,Deep KWS systems that have achieved state-of-the-art performance on a variety of KWS tasks. Deep KWS systems are more robust to noise and other environmental factors than traditional KWS systems. Deep KWS systems can be used for a variety of applications, such as voice assistants, hands-free control of devices, and security systems. The paper provides a comprehensive overview of deep spoken keyword spotting. It is a valuable resource for researchers and practitioners who are interested in this technology.

[4]**Monkey Says, Monkey Does: Security and Privacy on Voice Assistants — 10.1109/ACCESS.2017.2747626** The paper discusses the security and privacy risks of voice assistants. Voice assistants are always listening for a wake word, and they can be activated by accident. This means that a malicious actor could potentially eavesdrop on private conversations or even take control of the device.

The paper discusses a number of attacks that can be launched against voice assistants. These attacks include:

Wake word spoofing: This attack involves playing a recording of the wake word to the device, which can cause it to activate. Command injection: This attack involves sending a malicious command to the device, which can be used to take control of it. Data exfiltration: This attack involves stealing data from the device, such as the user's voice recordings or personal information. The paper concludes by discussing a number of security measures that can be taken to protect voice assistants. These measures include:

Using a strong wake word: This will make it more difficult for an attacker to spoof the wake word. Keeping the device up to date: Software updates often include security patches that can help to protect the device from attacks. Using a privacy mode: This mode will prevent the device from listening for a wake word when it is not in use. The paper provides a valuable overview of the security and privacy risks of voice assistants. It is important for users to be aware of these risks and to take steps to protect their devices.

The authors conducted a security analysis of four popular voice assistants: Amazon Alexa, Google Home, Apple Siri, and Microsoft Cortana. They found that all four assistants were vulnerable to wake word spoofing attacks. They also found that all four assistants were vulnerable to command injection attacks. The authors recommend that users take the following steps to protect their voice assistants: Use a strong wake word. Keep the device up to date. Use a privacy mode. Be careful about what information you share with the device. The paper concludes by stating that voice assistants are a powerful new technology, but they also pose a number of security and privacy risks. Users should be aware of these risks and take steps to protect their devices.

[5]**Implementation of Google Assistant and Amazon Alexa on Raspberry Pi**- The paper ”Implementation of Google Assistant on Rasberry Pi” investigates how to implement the voice Google Assistant on a Raspberry Pi microcomputer. The authors first present the Voice Kit of Google, which can be attached to a Raspberry Pi to create a voice Google assistant. They then go on to describe the necessary steps to implement the voice Google assistant on a system that contains only a microphone and a speaker, instead of the Voice Kit. This allows for a more flexible and easy-to-use voice Google assistant. The paper then analyzes the working modes of the newly created device, and finally presents and evaluates a speech recognition system that works in Romanian language.

The paper concludes by discussing the advantages and disadvantages of the proposed implementation, as well as some potential future work. Overall, the paper provides a detailed and well-researched overview of how to implement the voice Google Assistant on a Raspberry Pi.

3.2. Drawbacks of existing system

While voice assistants utilizing Natural Language Processing (NLP) have gained popularity and offer numerous benefits, they also have some drawbacks. Here are a few drawbacks of voice assistants using NLP:

1. **Contextual understanding:** Voice assistants often struggle with understanding context and maintaining continuity in conversations. They may fail to comprehend follow-up questions or references to previous queries, leading to fragmented interactions and user frustration.

2. **Ambiguity resolution:** Voice assistants sometimes struggle with disambiguating ambiguous queries or requests. When faced with multiple interpretations, they may provide generic or unrelated responses, resulting in a lack of accuracy and relevance.

3. **Proactive assistance:** Current voice assistants primarily operate in a reactive mode, responding to explicit user requests. However, they often lack the capability to anticipate user needs or provide proactive assistance without specific prompts. This limits their ability to provide timely and helpful information.

4. **Natural language generation:** While voice assistants have improved in understanding natural language, generating natural and coherent responses remains a challenge. Responses may sound robotic, lack human-like conversational flow, or fail to provide nuanced and contextual information.

5. **Domain-specific knowledge:** Voice assistants may struggle to provide in-depth and accurate information in specialized or niche domains. They might not have access to up-to-date knowledge or lack the ability to understand and answer complex questions in specific areas of expertise.

6. **Emotional intelligence:** Current voice assistants lack emotional intelligence and empathy. They cannot effectively recognize or respond to user emotions, limiting their ability to provide appropriate support in sensitive situations or personalized experiences.

7. **Multilingual capabilities:** While many voice assistants support multiple languages, their proficiency and accuracy can vary across different languages. Some languages may receive limited support, resulting in a language barrier for non-English speakers.

8. **Privacy and security concerns:** Voice assistants have raised concerns about privacy and data security. Users may worry about their voice recordings being stored, analyzed, or shared without their consent. Ensuring robust privacy measures and transparent data handling practices is crucial.

9. **Integration with third-party services:** While voice assistants can connect with certain third-party applications and services, the integration process can be challenging and limited. Compatibility issues or restrictions on available features may hinder seamless integration and broader functionality.

10. **Personalization limitations:** While some voice assistants offer basic personalization, their ability to adapt to individual preferences and behaviors is often limited. Tailoring experiences based on specific user needs, interests, and contexts is an area that can be further developed.

3.3. Gaps Identified

---

1. **Non-Multilingual** - Many of the voice assistants available today weather it is Cortana, Siri, Google Assistant are not multilingual and does not support many Indian and foreign languages. Since many voice assistants support English as their primary language. This restraints Non-English speaking users from utilizing a voice assistant's full functionality.

2. **Inaccuracy** - Many of the voice assistants are found to be inaccurate and provide inaccurate data many a times as they've limited information the current data, news, events and the new discoveries happening in the field of science.

3. **Difficulty with certain accents or languages** - Voice assistants may have difficulty understanding certain accents or languages that differ from the trained or default language models. This can limit their usability for individuals with non-native accents or speakers of less widely spoken languages.

4. **Can't connect to users emotionally:** Voice assistants primarily focus on providing factual information or performing tasks but lack emotional intelligence. They are not designed to provide emotional support or engage in empathetic conversations.

3.4. Objectives

1. **Enhance user convenience:** The primary objective of a voice assistant is to provide users with a convenient and hands-free way to interact with technology. It should be designed to simplify tasks and make everyday activities easier for users.

2. **Provide personalized assistance:** A voice assistant can aim to understand users' preferences, behaviors, and context to deliver personalized recommendations, reminders, and information. The objective is to create a tailored experience that meets individual needs.

3. **Enable natural language understanding and conversation:** A key goal of a voice assistant is to understand and interpret natural language effectively. It should be capable of processing and responding to user queries, instructions, and requests in a conversational manner, mimicking human-like interaction.

4. **Support a wide range of tasks and services:** The voice assistant should be versatile and able to perform various tasks, such as providing weather updates, answering general knowledge questions, setting reminders, playing music, controlling smart home devices, making reservations, and more. The objective is to be a comprehensive assistant that can handle multiple functions.

5. **Ensure reliability and accuracy:** A voice assistant should strive to provide accurate and reliable information to users. The objective is to minimize errors, misunderstandings, and false responses to ensure a high level of trust and user satisfaction.

6. **Continuously improve and learn:** A voice assistant should have the ability to learn from user interactions and adapt its responses and behavior over time. The objective is to improve its performance, expand its knowledge base, and enhance user experiences through ongoing learning and refinement.

7. **Protect user privacy and data security:** It is crucial to prioritize user privacy and data security when developing a voice assistant. The objective is to implement robust security measures and follow best practices to safeguard user information and maintain confidentiality.

8. **Support integration and interoperability:** The voice assistant should aim to integrate with various devices, platforms, and services to offer a seamless user experience across different technologies. The objective is to ensure compatibility and interoperability with existing systems and enable easy integration with third-party applications and services.

4. PROPOSED METHODOLOGY

4.1. Methodology

For this project we have used a microprocessor i.e Raspberry pi - 3 Model B+. This Raspberry Pi 3 Model B+ is powered by a 1.4GHz 64-bit quad-core ARM Cortex-A53 processor, providing improved performance compared to its predecessors.It is equipped with 512MB LPDDR2 RAM, allowing for smooth multitasking and efficient operation of applications. The board has built-in Wi-Fi 802.11ac and Bluetooth 4.2/BLE, enabling wireless connectivity for internet access, IoT applications, and communication with other devices.

We start by speaking the phrase 'Hey Leah' which acts a wake word engine for the voice assistant. Meanwhile the voice assistant in background keeps looking for the 'Hey Word' phrase despite the noises in the background. and activates only when the phrase is spoken by the user.

Once the user speaks or puts its request to the voice assistant, the voice assistant looks for the 'intent' in the sentence spoken by the user and recognizes what the user has actually spoken or what actually the user's request is.

Once the intent gets recognized it passes that desired request to the skill handler which then chooses a skill based on the request, requested by the user. For example: if a user requests for the current time it passes that intent to the skill handler which then passes the request to the Module/API containing that skill with its respective code, here its 'time.py' module.

After the request gets fulfilled by the Module/API the output gets generated by voice assistant in the form of audio.

4.2. Block diagram

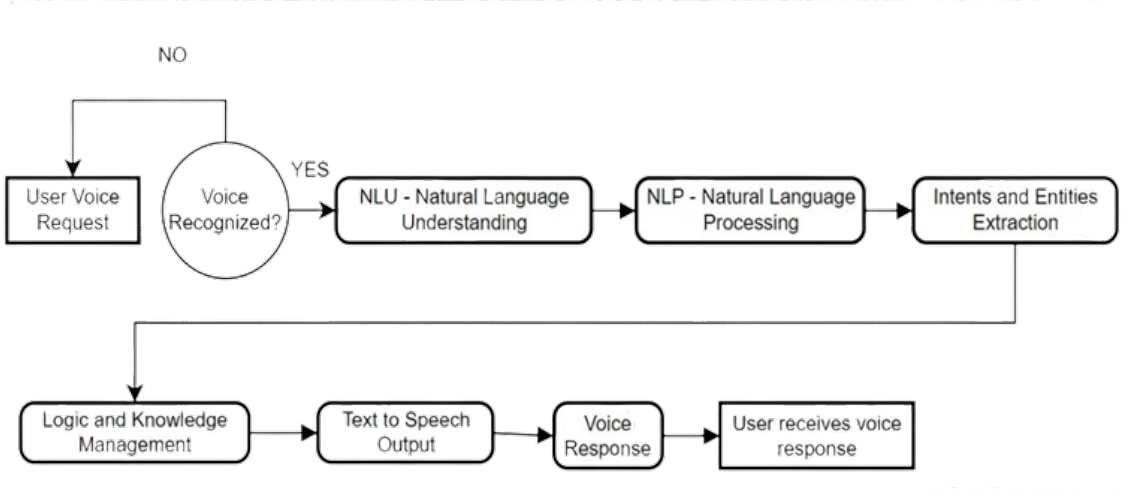

Figure 4.1: Block Diagram

4.3. Flow Diagram

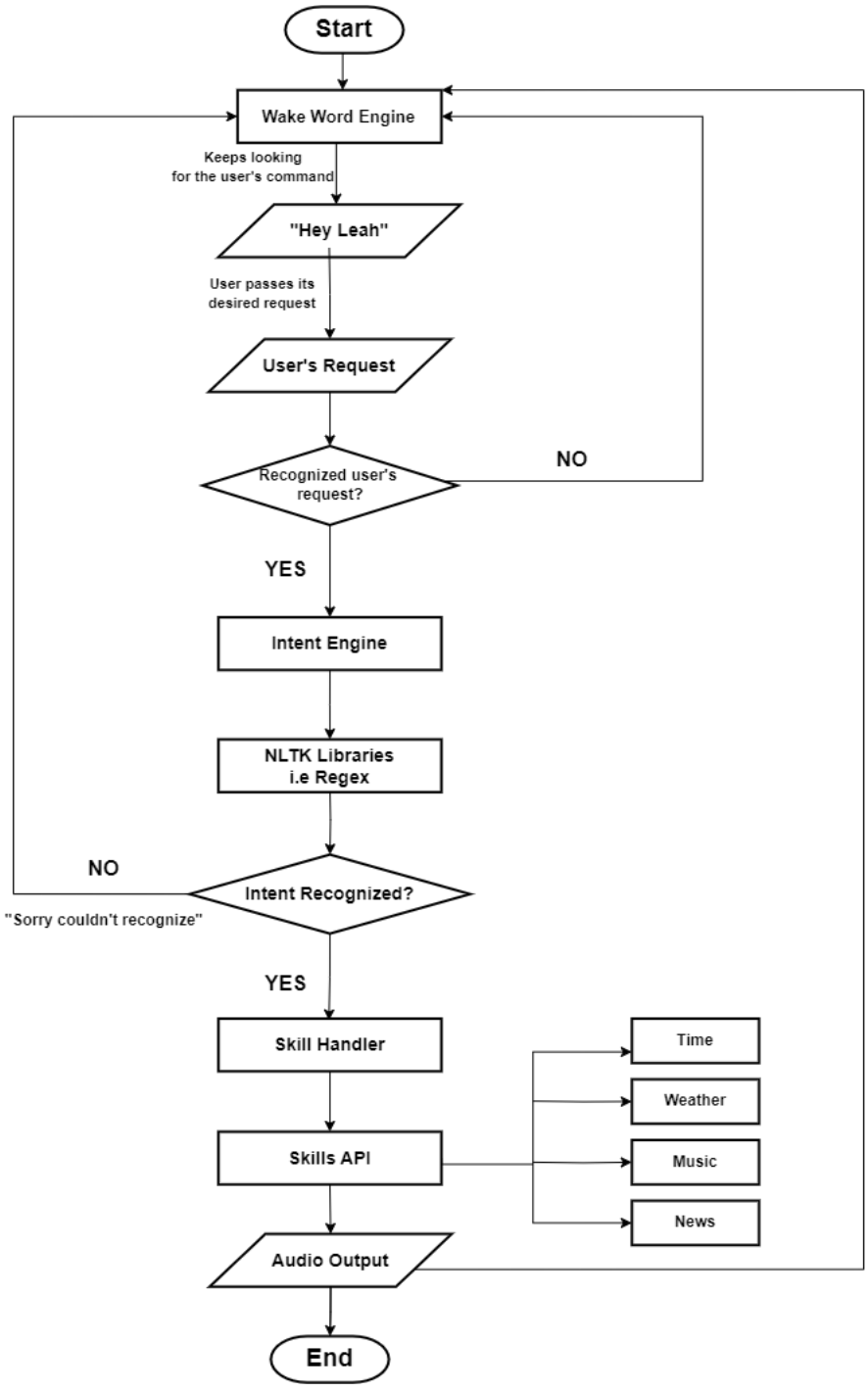

Figure 4.2: Flow Diagram

4.4. Tools Used

Software Tools[6]

- Pycharm – It's a Python Integrated Development Environment (IDE)
- pip - To install python libraries
- Linux OS for raspberry Pi to be installed on its EMMC storage.
- Miniconda package
- Python 3.8.16 working environment
- PyAudio Library
- NLTK library
- VS Code (IDE)
- Google TTS
- VNC Server
- Working Internet Connection

API's Used

- API Ninjas - Facts API [7]
- ThinkSpeak API - for home security monitoring [8]
- Shazam API - for music detection [9]
- NewsonAir API - for playing english/hindi news [10]
- DuckDuckGo API - for single info search [11]
- OpenWeather API - for weather forecast [12]

Hardware Tools

- Raspberry pi - 3 Model B+
- 512 Mb Sandisk SSD
- 3.5mm jack supported
- Speaker - for audio output
- Microphone - for user input
- PIR Sensor and NodeMCU - for Home Security Monitoring
- PC with Linux (raspberry runs on Linux)

Figure 4.3: Image displaying raspberry pi module with connections to - Wifi through LAN, Microphone, Speaker, HDMI, Home Security System through MicroUSB

4.5. Advantages & Disadvantages

Voice assistants have become increasingly popular in recent years and offer a range of advantages and disadvantages.

Advantages:

1. Convenience: Voice assistants are hands-free, allowing users to perform tasks and get information without having to physically interact with a device. This can be especially useful in situations where a user's hands are occupied or if they have limited mobility.

2. Time-saving: With voice assistants, users can perform tasks quickly and efficiently without having to navigate menus or type out commands. This can be especially helpful for tasks like setting reminders, making phone calls, or sending messages.

3. Personalization: Voice assistants can be personalized to a user's preferences, including their preferred music, news sources, and more. This allows for a more tailored and personalized experience.

4. Accessibility: Voice assistants can be a great tool for people with disabilities, such as those with visual impairments, who may find it difficult to interact with devices in other ways.

Disadvantages:

1. Privacy concerns: Voice assistants are always listening and recording, which can raise privacy concerns for some users. There have been instances where voice assistant recordings have been accessed or leaked without the user's knowledge or consent.

2. Accuracy: Voice assistants are not always accurate, and can misinterpret commands or fail to understand accents or dialects.

3. Limited functionality: Voice assistants may not be able to perform all tasks that a user needs, and some may require additional devices or services to function properly.

4. Dependency: Some users may become too reliant on voice assistants, leading to a loss of independence and reliance on technology.

Overall, while voice assistants can provide numerous benefits, users should be aware of the potential drawbacks and take steps to protect their privacy and maintain a healthy relationship with technology.

5. ANALYSIS AND DESIGN

In this project there are three main components, The **Wake Word Engine**, The **Intent Engine**, The **Skill Handler**.

1. **Wake Word Engine:** The Wake Word Engine keeps looking for the phrase ”Hey Leah” in the background and stops after few moments if it doesn't recognize any activity from the user. The Wake Word Engine gets called even when any intent gets failed to be recognized by Leah. Also after , once the user's request gets fulfilled an audio output is generated after which the the wake word engine gets called again - looking for the phrase ”Hey Leah”.

2. **Intent Engine:** The function of Intent Engine is to identify 'What the user has actually spoken/or desired for'. Its aim is to look for certain the keys or words that is present in the English dictionary. So for example if user speaks - 'What is the Time?'. It will extract the word 'time' and pass it over to the Skill Handler, once the Intent gets recognized satisfactorily.

3. **Skill Handler:** The function of a Skill Handler is to pass on the intent received from the intent engine to the specific skill API based on intent derived from the user's request. It then calls that specific skill module as per the user's request. For example: if a user requested for 'time' the skill handler will call the 'time.py' module, after which the output gets generated in the form Audio. Some of the featured skills that can be seen in Leah are:

   (a) News (*Aakashvaani*)

   (b) Dr. Leah

   (c) Daily Quotes

   (d) Time and Date

   (e) Singular Search

   (f) Send Messages on Telegram

   (g) Weather Forecast of any place

   (h) Music/Song Recognition and Playing

   (i) Home Security System using NodeMCU-ESP8266 and PIR Sensor.

   Also we've tried an approach to make Leah multilingual, i.e *hamari leah hindi aur marathi dono bolti hai.*

5.1. UML Diagrams

5.1.1. Use Case Diagram

A use case diagram is a visual representation that illustrates the interactions between actors (users or external systems) and a system, showcasing the various ways the system can be utilized to achieve specific goals. It helps to identify and define the functional requirements of the system by outlining the different use cases, which are specific scenarios or actions performed by the actors. Use case diagrams provide a high-level overview of the system's functionality, highlighting the relationships and dependencies between the actors and the system, aiding in requirements analysis, communication, and design of software systems.

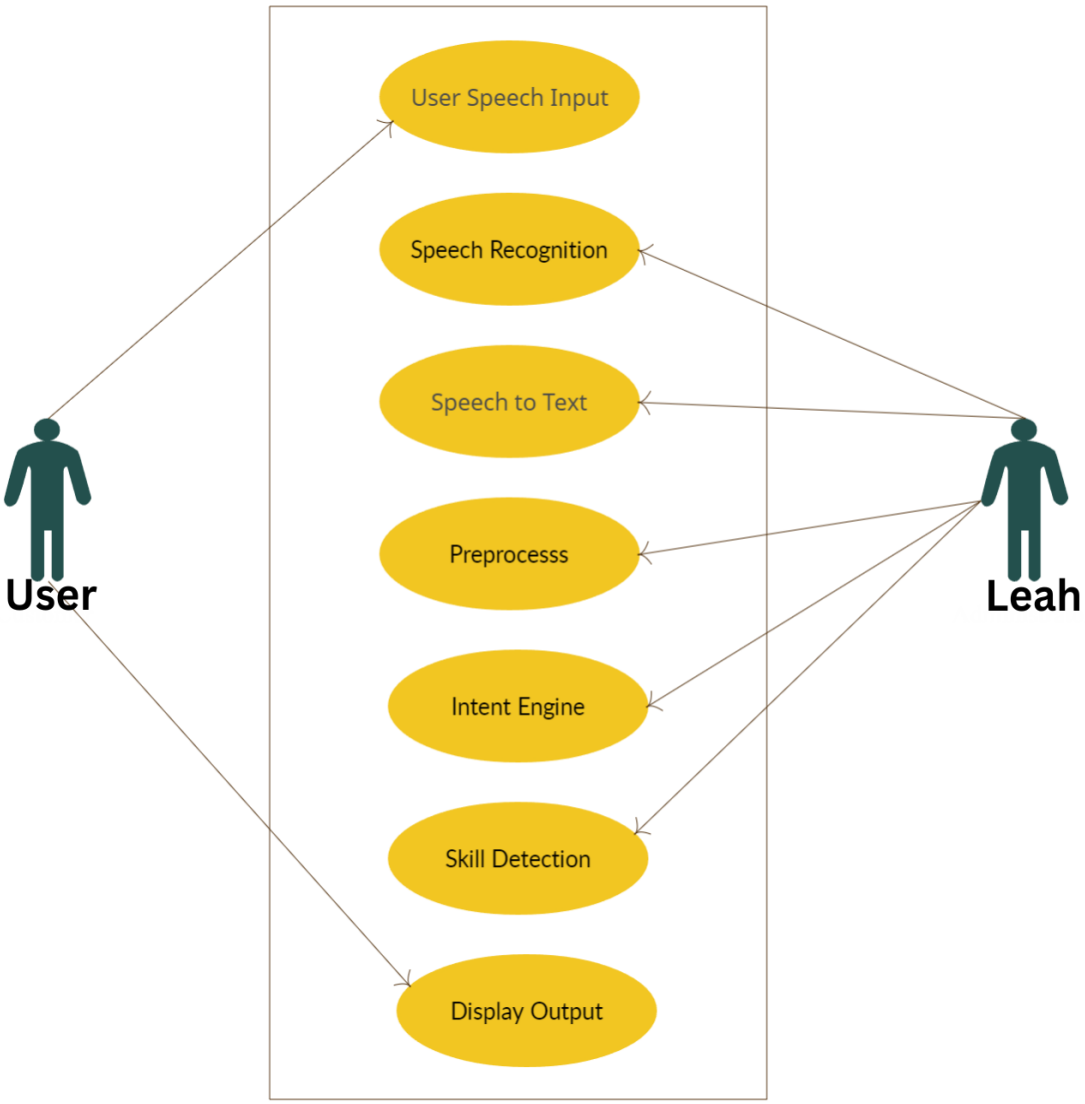

Figure 5.1: Use Case Diagram

5.1.2. Class Diagram

A class diagram is a visual representation of the structure and relationships among classes in an object-oriented system. It illustrates the essential components of the system, such as classes, attributes, operations, and associations. Classes are depicted as rectangles, displaying their name, attributes, and methods. Associations are represented by lines connecting classes, indicating the connections and dependencies between them. The diagram provides a high-level overview of the system's architecture, helping developers understand the interactions between classes and aiding in software design and development. It serves as a blueprint for the system's implementation and facilitates communication among stakeholders involved in the development process.

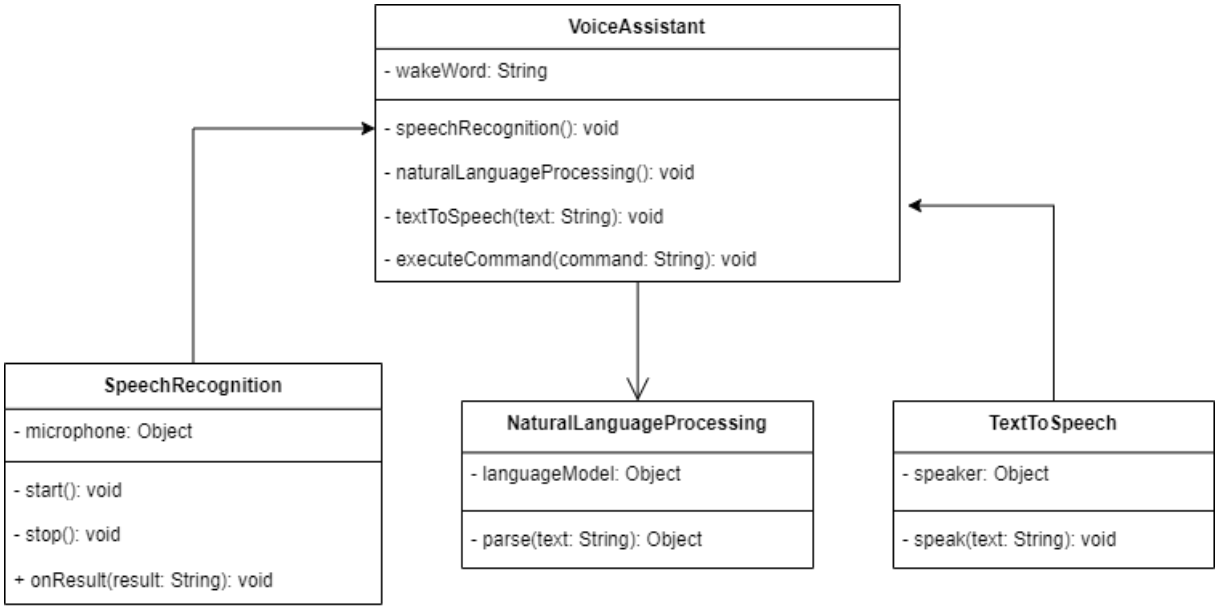

Figure 5.2: Class Diagram

5.1.3. Data flow diagram

A data flow diagram (DFD) is a graphical representation of how data flows through a system, modeling its inputs, outputs, and processes. It is a type of structured analysis and design tool that is used to describe and analyze the flow of data within an organization or system. A DFD consists of a set of interconnected data flow diagrams that show the flow of data through different processes within the system. Each process in the diagram represents a specific activity or operation that transforms the input data into output data. The input and output data are represented as data flows, and they are connected to the processes using arrows.

DFD LEVEL 0

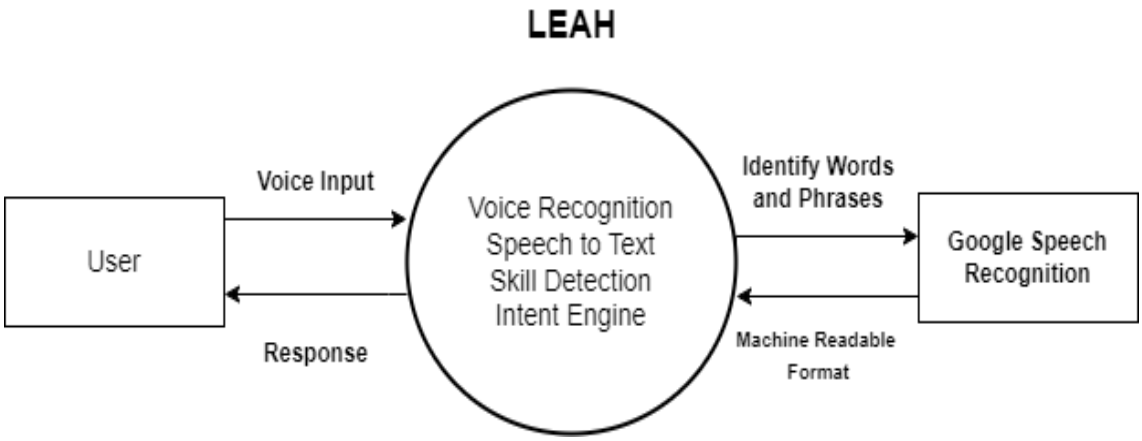

Figure 5.3: DFD Level 0

DFD LEVEL 1

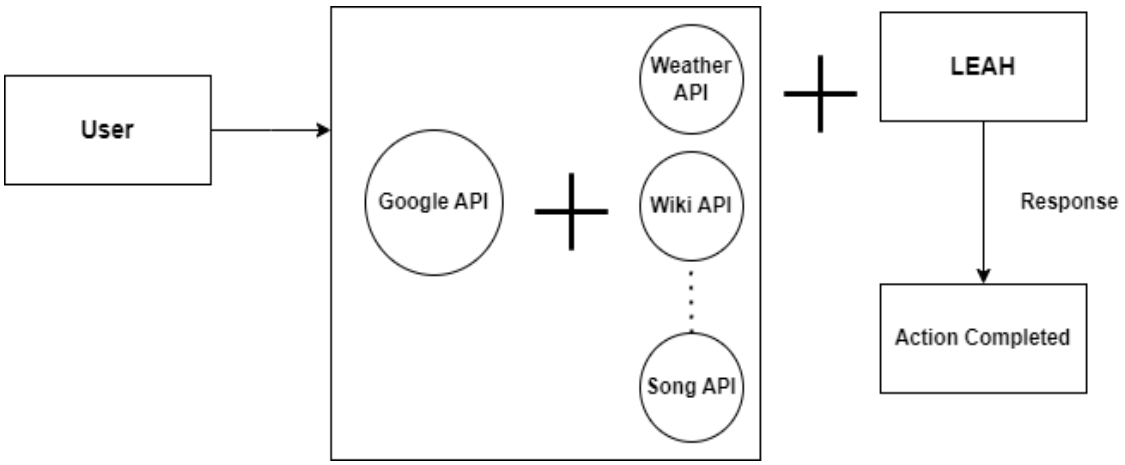

Figure 5.4: DFD Level 1

DFD LEVEL 2

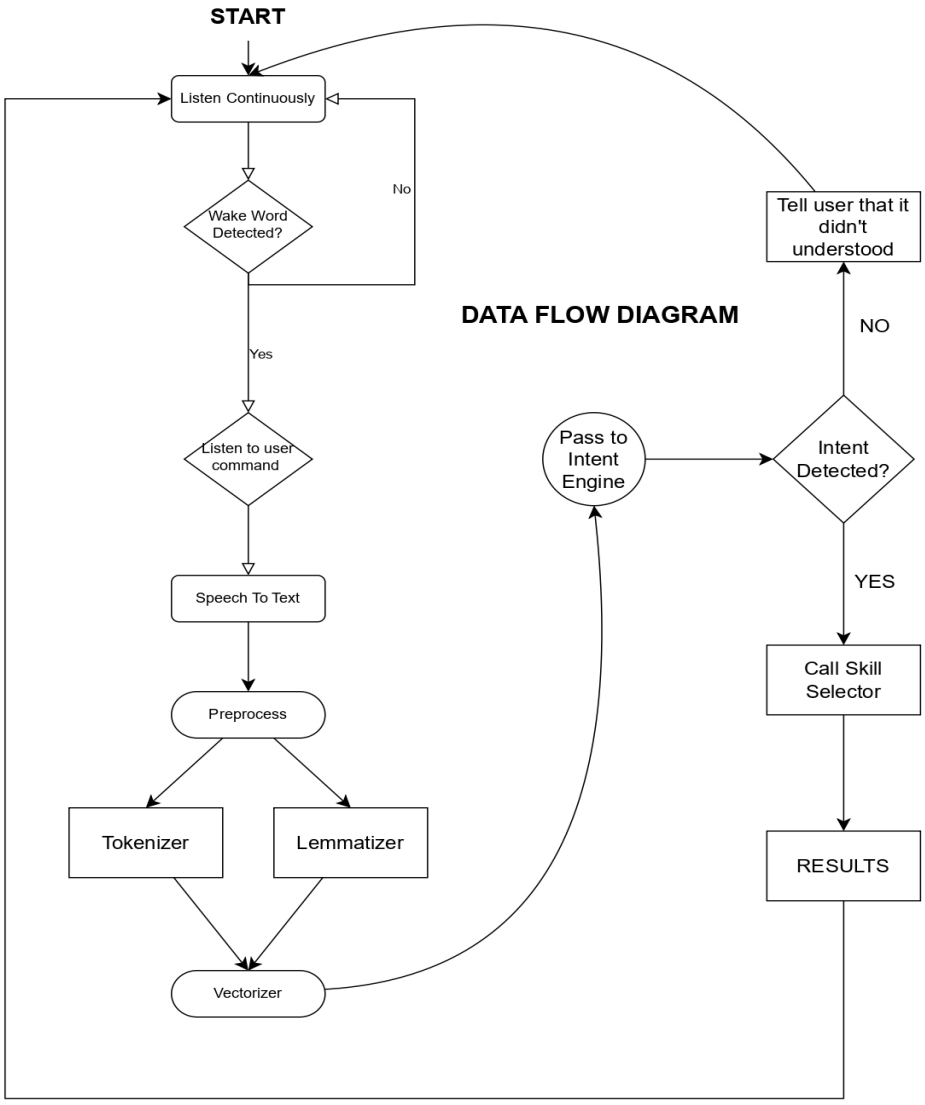

Figure 5.5: DFD Level 2

5.1.4. Component Diagram

A component diagram is a visual representation in UML that depicts the organization and relationships between software components within a system. It shows how components, which can be physical or logical entities, interact through interfaces, ports, and connectors. Components represent modular parts of the system, and connectors illustrate dependencies, associations, or communications between them. Component diagrams help in understanding the system's structure, identifying component dependencies, and facilitating communication and design of software architectures. They are widely used in software engineering to document and communicate the composition and relationships of complex systems.

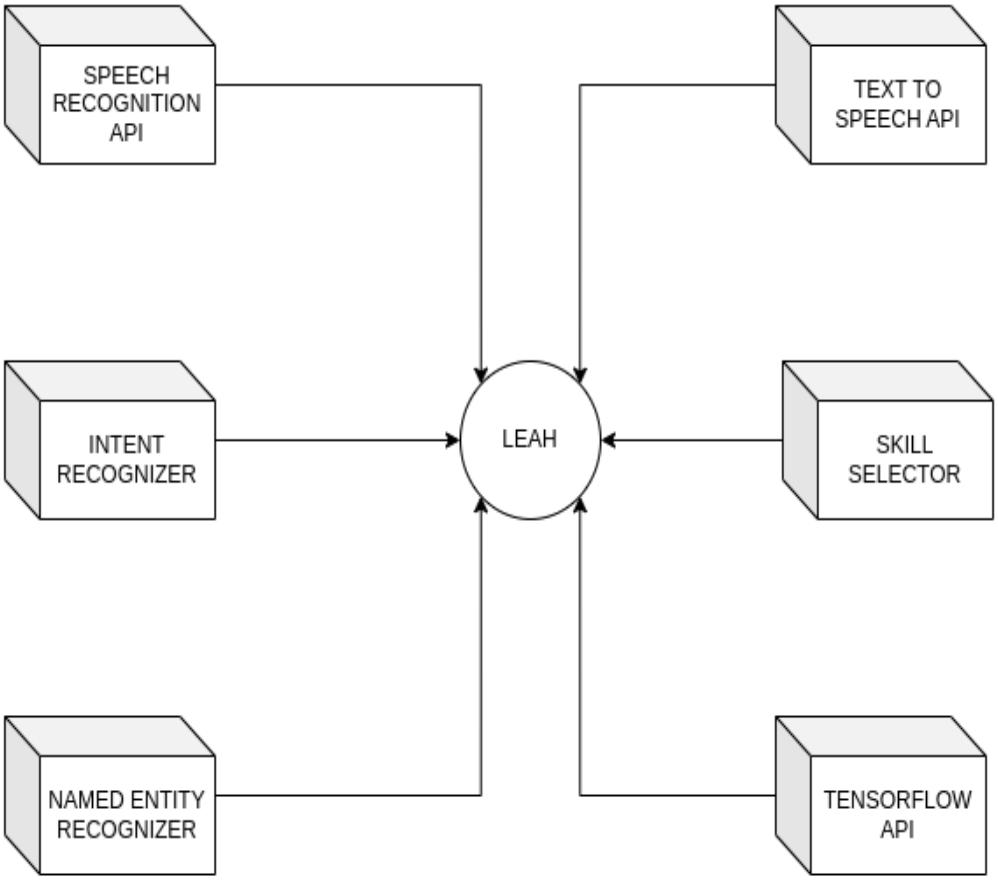

Figure 5.6: Component Diagram

5.1.5. Sequence Diagram

A sequence diagram is a UML (Unified Modeling Language) diagram that illustrates the interactions between objects or components in a system over time. It shows the chronological order of messages exchanged between objects to accomplish a particular functionality or scenario. Objects are represented as vertical lifelines, and the messages are shown as horizontal arrows between the lifelines. The sequence diagram provides a visual depiction of how objects collaborate, including the sequence of method calls, the flow of data, and the order of events. It is commonly used for modeling and understanding the behavior of complex systems, especially in the context of software development.

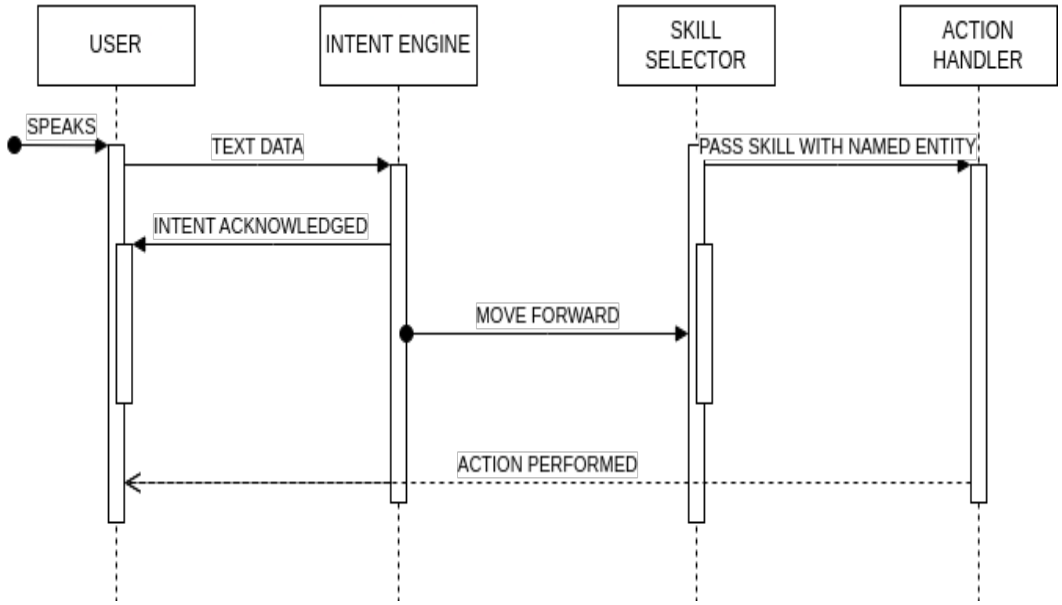

Figure 5.7: Sequence Diagram

5.1.6. Activity Diagram

An activity diagram is a visual representation in UML (Unified Modeling Language) that illustrates the flow of activities or actions within a system or process. It captures the dynamic behavior and logic of a system by depicting the sequence of activities, decision points, concurrency, and the flow of data between different activities. Activity diagrams consist of nodes representing activities, decisions, or start/end points, and arrows representing transitions between these nodes. They are useful for modeling business processes, workflows, software algorithms, or complex behaviors within a system. Activity diagrams provide a clear and intuitive view of how activities are performed and help in analyzing and improving system processes.

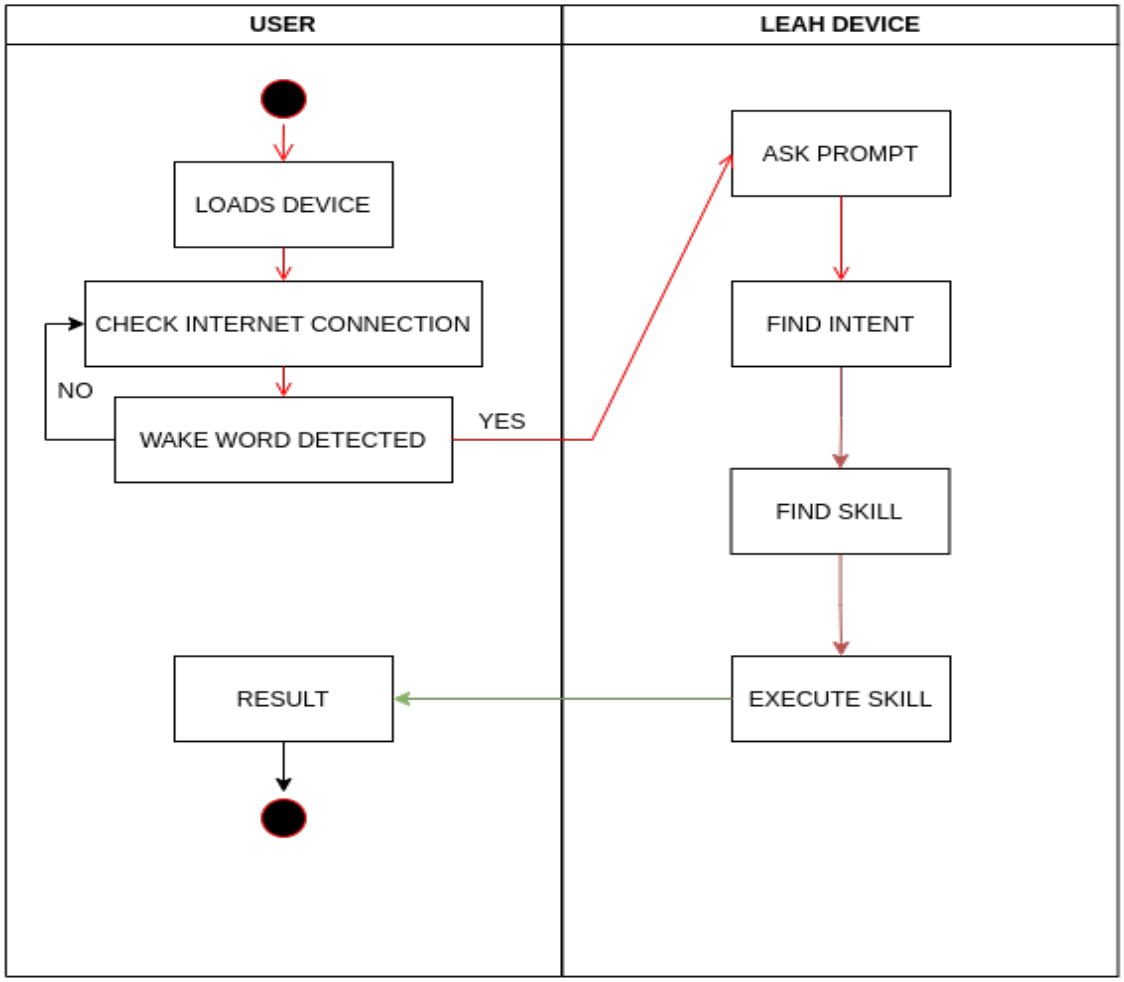

Figure 5.8: Activity Diagram

5.1.7. Deployment Diagram

A deployment diagram is a type of UML (Unified Modeling Language) diagram that depicts the physical arrangement of hardware and software components within a system. It illustrates how software artifacts (such as executables, libraries, and files) are deployed onto hardware nodes (such as servers, computers, or devices) in a distributed system. Deployment diagrams show the relationships between nodes, including the communication paths, dependencies, and associations. They can also depict the allocation of components to specific nodes, indicating where each component is executed or hosted. Deployment diagrams help in understanding the deployment architecture, resource allocation, and the distribution of software across different hardware nodes in a system.

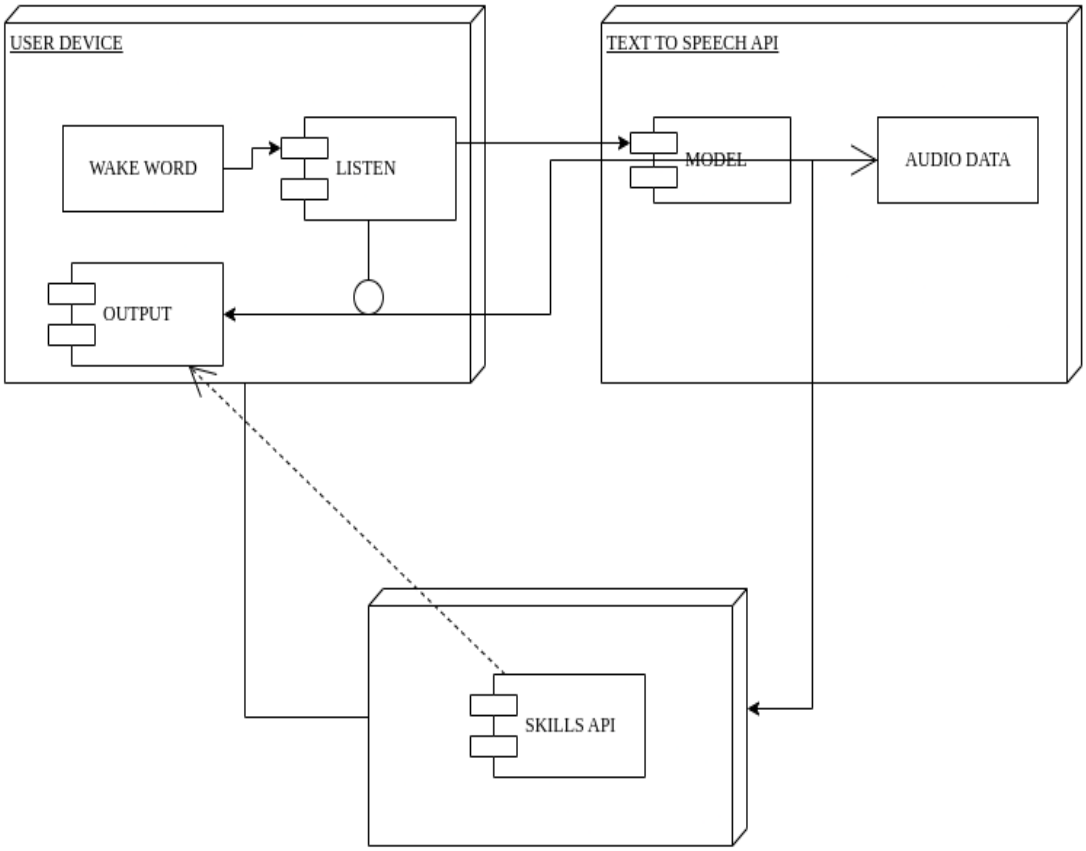

Figure 5.9: Deployment Digram

6. RESULTS AND DISCUSSIONS

Here are the outcomes of our voice assistant.

Leah As a News Anchor and Music Player

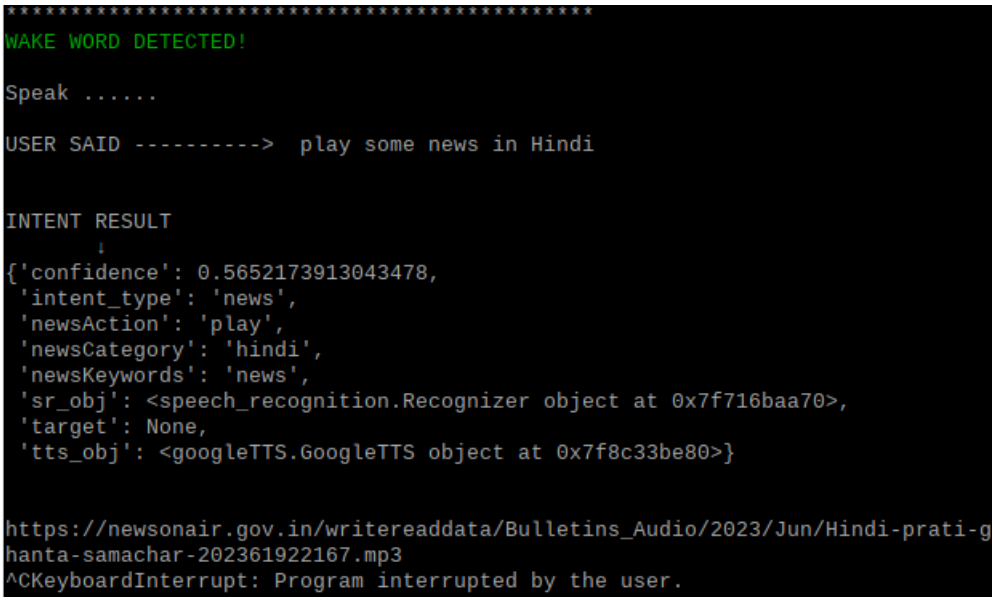

Figure 6.1: Leah Speaking News

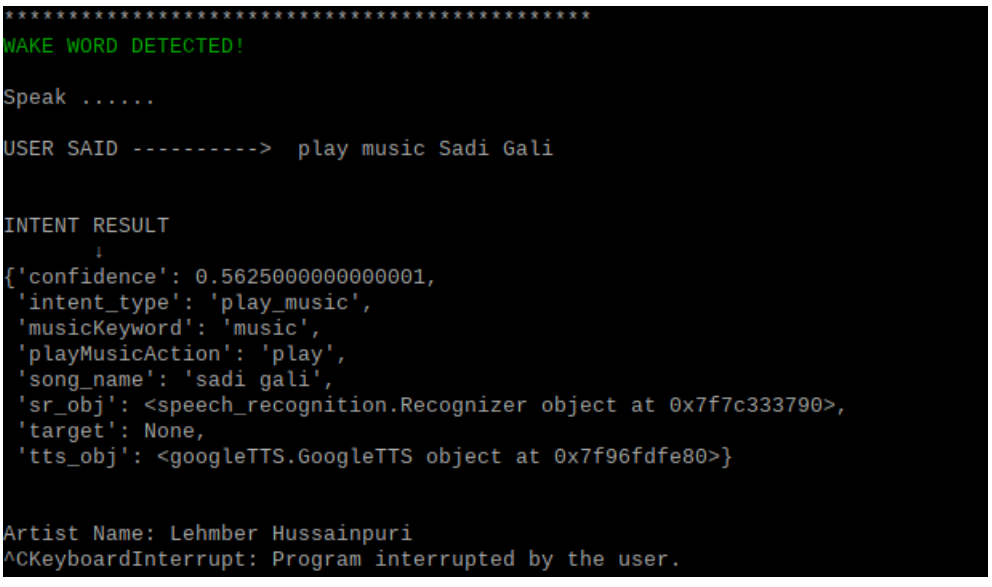

Figure 6.2: Leah Music

Leah As A Home Doctor

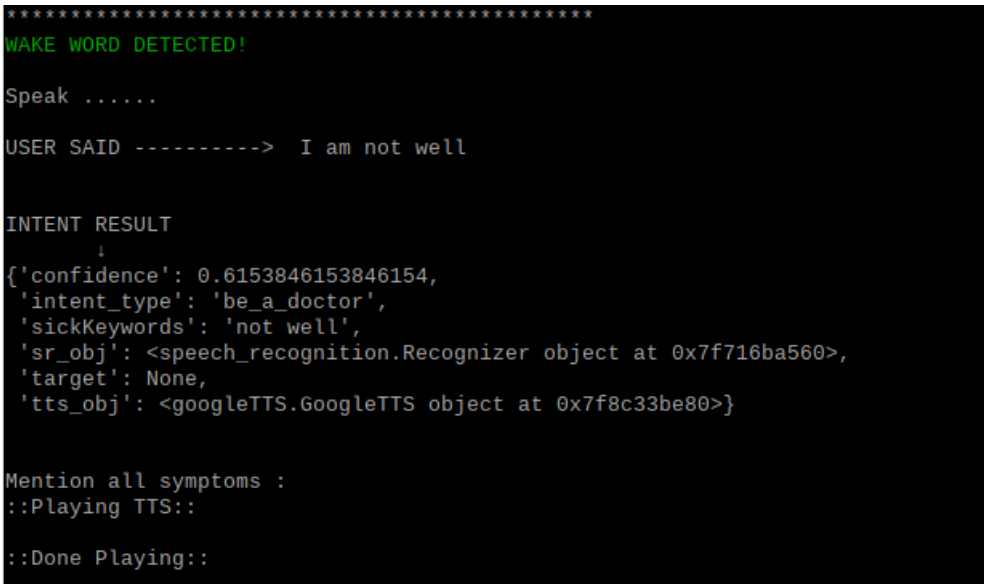

Figure 6.3: Dr. Leah

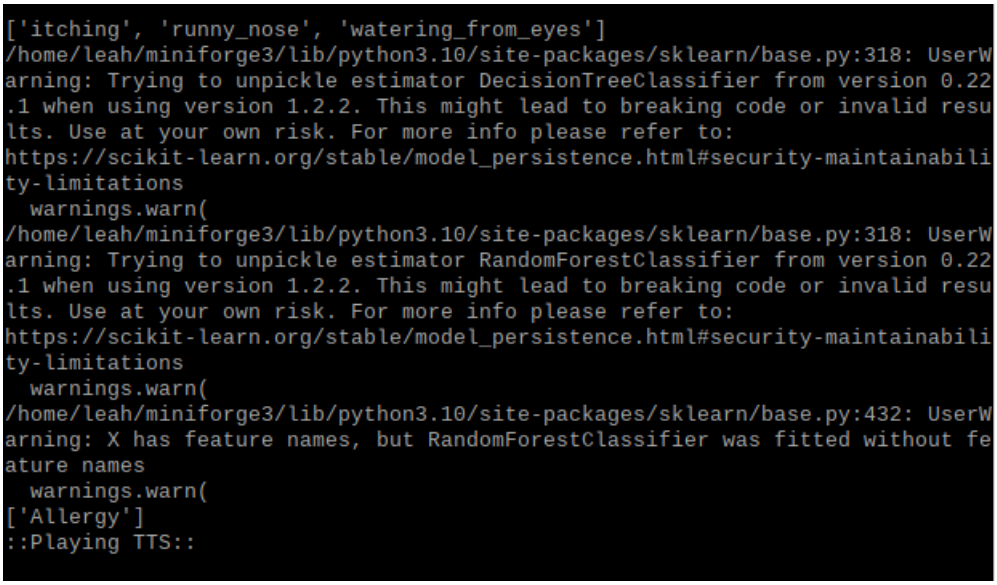

Figure 6.4: Dr. Leah

Leah Sending Daily Quotes and Leah Search

```
**********************************************
WAKE WORD DETECTED!

Speak ......

USER SAID ---------->  start sending me daily motivation


INTENT RESULT
       ↓
{'confidence': 0.6363636363636364,
 'dailyQuotesAction': 'start',
 'dailyQuotesKeywords': 'daily motivation',
 'intent_type': 'daily_quotes',
 'sr_obj': <speech_recognition.Recognizer object at 0x7f716ba980>,
 'target': None,
 'tts_obj': <googleTTS.GoogleTTS object at 0x7f8c33be80>}


Cron job created successfully.
::Playing TTS::

::Done Playing::
```

Figure 6.5: Leah Sending Quotes

```
**********************************************
WAKE WORD DETECTED!

Speak ......

USER SAID ---------->  search for Raspberry Pi


INTENT RESULT
       ↓
{'confidence': 0.3913043478260869,
 'intent_type': 'search_summary',
 'searchKeyword': 'search',
 'search_entity': 'raspberry pi',
 'sr_obj': <speech_recognition.Recognizer object at 0x7f716bae00>,
 'target': None,
 'tts_obj': <googleTTS.GoogleTTS object at 0x7f8c33be80>}


Response ----------> Raspberry Pi is a series of small single-board computer
s developed in the United Kingdom by the Raspberry Pi Foundation in associati
on with Broadcom
```

Figure 6.6: Leah Search

Leah Speaking Time and Weather

```
**********************************************
WAKE WORD DETECTED!

Speak ......

USER SAID ---------->  what is the time


INTENT RESULT
       ↓
{'confidence': 0.19791666666666666,
 'intent_type': 'time',
 'sr_obj': <speech_recognition.Recognizer object at 0x7f716b9ba0>,
 'target': None,
 'timeKeyword': 'time',
 'tts_obj': <googleTTS.GoogleTTS object at 0x7f8c33be80>}


Response ---------->  the time is 10 54 PM


::Playing TTS::
```

Figure 6.7: Leah Time

```
**********************************************
WAKE WORD DETECTED!

Speak ......

USER SAID ---------->  tell me the weather of Valsad


INTENT RESULT
       ↓
{'confidence': 0.25862068965517243,
 'intent_type': 'weather',
 'location': 'valsad',
 'sr_obj': <speech_recognition.Recognizer object at 0x7f716ba230>,
 'target': None,
 'tts_obj': <googleTTS.GoogleTTS object at 0x7f8c33be80>,
 'weatherKeyword': 'weather'}


Response ---------->  In Valsād, it's 29.94 degrees Celsius and broken clouds
, with a high of 29.94 and a low of 29.94.
```

Figure 6.8: Leah Weather

Leah Shazam

```
************************************************
WAKE WORD DETECTED!

Speak ......

USER SAID ---------->  what song is this


INTENT RESULT
       ↓
{'confidence': 0.45098039215686275,
 'intent_type': 'song_recognise',
 'songRecognitionKeywords': 'song',
 'songRecognitionKeywords2': 'what',
 'sr_obj': <speech_recognition.Recognizer object at 0x7f716bae00>,
 'target': None,
 'tts_obj': <googleTTS.GoogleTTS object at 0x7f8c33be80>}


::Playing TTS::

::Done Playing::
```

Figure 6.9: Leah Music Recognition

```
::Playing TTS::

::Done Playing::

Recording 7 seconds of audio...
Recording completed.
Audio saved to '/home/leah/Documents/leah-final/temp/shazam.mp3'.
::Playing TTS::

::Done Playing::

Title :  Kya Mujhe Pyar Hai (From "Woh Lamhe")
Artist :  KK
its Kya Mujhe Pyar Hai (From "Woh Lamhe") by KK
::Playing TTS::

::Done Playing::

Response ---------->  None
```

Figure 6.10: Leah Music Recognition

Leah Home Security System

```
**********************************************
WAKE WORD DETECTED!

Speak ......

USER SAID ---------->  start monitoring my home


INTENT RESULT
       ↓
{'confidence': 0.625,
 'homeSecurityAction': 'start',
 'homeSecurityKeywords': 'monitoring',
 'intent_type': 'home_security',
 'sr_obj': <speech_recognition.Recognizer object at 0x7f716bb070>,
 'target': None,
 'tts_obj': <googleTTS.GoogleTTS object at 0x7f8c33be80>}


Home Monitor:
Waiting 5 Seconds..
Sending Request..
Waiting 5 Seconds..
```

Figure 6.11: Leah Security

```
Home Monitor:
Waiting 5 Seconds..
Sending Request..
Waiting 5 Seconds..
Sending Request..
1 No Detection: 234
Waiting 5 Seconds..
Sending Request..
1 No Detection: 234
Waiting 5 Seconds..
Sending Request..
1 Motion Detected: 235
Telegram Message Delivered Successfully.
{"ok":true,"result":{"message_id":100,"from":{"id":6192792234,"is_bot":true,"
first_name":"leah_telegram","username":"leah_tg_bot"},"chat":{"id":-960255198
,"title":"Leah Test Group","type":"group","all_members_are_administrators":tr
ue},"date":1687196798,"text":"Leah Home Security :\n\nMotion has been detecte
d in your home on 19/6/2023 at 11:16:37 PM\n\nTake Action Immediately!","enti
ties":[{"offset":63,"length":9,"type":"bold"},{"offset":76,"length":11,"type"
:"bold"},{"offset":89,"length":24,"type":"bold"}]}}
Take Action Immediately
```

Figure 6.12: Leah Security

7. CONCLUSION

In conclusion, ”Leah: A NLP Powered Voice Assistant Streamlining Your Daily Tasks” highlights the significance of leveraging Natural Language Processing (NLP) to create a voice assistant that enhances daily productivity.By leveraging the capabilities of the Raspberry Pi, Leah offers a seamless and user-friendly experience for individuals looking to simplify their daily tasks.

The Raspberry Pi’s processing power, combined with its connectivity options, enables Leah to handle natural language processing (NLP) tasks with speed and accuracy. The integration of the Raspberry Pi’s Wi-Fi and Bluetooth capabilities enhances Leah’s ability to connect and interact with various devices and services, expanding its functionality and usability.

Moreover, the Raspberry Pi’s GPIO pins allow for easy integration with external sensors and peripherals, providing additional customization options and extending the capabilities of Leah. This flexibility makes it possible to tailor the voice assistant to meet specific user needs and integrate it into various environments and applications.

Furthermore, the compact form factor of the Raspberry Pi makes it ideal for embedding Leah into different physical devices, such as smart speakers or home automation systems, further enhancing the user experience and convenience.

Through the implementation of advanced NLP techniques, Leah is a powerful tool capable of understanding and responding to user commands with a high level of accuracy and context awareness. This enables a user to seamlessly interact with the voice assistant, simplifying and automating various tasks throughout their day.

Leah’s ability to comprehend natural language allows for fluid communication, empowering users to issue commands, ask questions, and receive personalized responses tailored to their preferences and historical interactions. The integration of a speech recognition model ensures accurate transcription of spoken language, further enhancing the user experience.

By streamlining tasks such as news, time,weather,home security, information retrieval, and more, Leah becomes an invaluable asset in increasing efficiency and productivity. The voice assistant’s intelligent dialogue management system enables smooth multi-turn conversations, adapting to user context and evolving requirements.

The development of Leah involved a meticulous methodology encompassing requirements gathering, data collection and preparation, NLU model training, speech recognition implementation, dialogue management design, response generation, integration of Skills, and

continuous iteration based on user input.

As technology advances, Leah's maintenance and regular updates will ensure it remains at the forefront of voice assistant capabilities. Ongoing evaluation and enhancement of its functionalities will enable Leah to adapt to changing user needs and emerging advancements in NLP and related technologies.

Ultimately, Leah's NLP-powered voice assistant represents a significant breakthrough in daily task management. Its seamless integration into users' lives, combined with its ability to understand and respond to natural language, positions Leah as a valuable tool for streamlining tasks and enhancing overall productivity in a user-friendly and intuitive manner.

References

[1] P. Cheng and U. Roedig, "Personal voice assistant security and privacy—a survey," *Proceedings of the IEEE*, vol. 110, no. 4, pp. 476–507, 2022.

[2] S. Malodia, N. Islam, P. Kaur, and A. Dhir, "Why do people use artificial intelligence (ai)-enabled voice assistants?" *IEEE Transactions on Engineering Management*, pp. 1–15, 2021.

[3] I. López-Espejo, Z.-H. Tan, J. H. L. Hansen, and J. Jensen, "Deep spoken keyword spotting: An overview," *IEEE Access*, vol. 10, pp. 4169–4199, 2022.

[4] E. Alepis and C. Patsakis, "Monkey says, monkey does: Security and privacy on voice assistants," *IEEE Access*, vol. 5, pp. 17 841–17 851, 2017.

[5] S. Arya and S. B. Patel, "Implementation of google assistant & amazon alexa on raspberry pi," *ArXiv*, vol. abs/2006.08220, 2020.

[6] S. S. Aditya Yadav, Raj Kansal, "Leah: A nlp powered voice assistant streamlining your daily tasks," https://github.com/adityay186/leah-raspberrypi, 2023.

[7] "API Ninjas," https://api-ninjas.com/, accessed on Month DD, 2023.

[8] "ThingSpeak," https://thingspeak.com/, accessed on Month DD, 2023.

[9] PyPI, "shazamio," https://pypi.org/project/shazamio/, accessed on Month DD, 2023.

[10] "News on AIR," https://newsonair.gov.in/, accessed on Month DD, 2023.

[11] SerpApi, "DuckDuckGo Search API," https://serpapi.com/duckduckgo-search-api, accessed on Month DD, 2023.

[12] OpenWeatherMap, "OpenWeatherMap API," https://openweathermap.org/api, accessed on Month DD, 2023.