INDUSTRIAL PROJECT REPORT

Internship project report on

Personal AI Assistant Bot

submitted in partial fulfilment of the
Requirements for the award of

Bachelor of Engineering
in
School of Electronics and Communication Engineering

Carried out at

Abhyudyaya Techno Solutions Pvt. Ltd.

Submitted By-

1. PAVAN SHIVALLI USN:01FE22BEC411
2. SARVESH KARAKANNAVAR USN:01FE22BEC412
3. NITIN SAVVASE USN:01FE22BEC413
4. BASAVARAJESHWARI H USN:01FE22BEC416

Under the guidance of

Dr. Tanuja Patil
College Guide
KLE Technological University

Ms. Priya
Industry Guide
Abhyudyaya Techno Solutions Pvt. Ltd.

Submitted To:
School of Electronics and Communication Engineering
KLE TECHNOLOGICAL UNIVERSITY
Hubballi

K.L.E SOCIETY'S
KLE Technological University,
HUBBALLI-580031
2024-2025

SCHOOL OF ELECTRONICS AND COMMUNICATION ENGINEERING

CERTIFICATE

This is to certify that the internship project entitled **"Personal AI Assistant Bot"** is a bonafide work carried out by **"Pavan Shivalli - 01FE22BEC411, Sarvesh Karakannavar - 01FE22BEC412, Nitin Savvase - 01FE22BEC413, Basavarajeshwari H - 01FE22BEC416"** in **Abhyudyaya Techno Solutions Pvt. Ltd.**, in partial fulfillment for the award for Bachelor of Engineering in Electronics and Communication in the School of Electronics and Communication Engineering of KLE Technological University, Hubballi for the academic year 2024-2025.

Dr. Tanuja Patil	**Dr. Suneeta V. Budihal**	**Dr. Basavaraj S. Anami**
Guide	**Head of School**	**Registrar**

External Viva:

Name of Examiners	**Signature with date**
1.	
2.	

DECLARATION

I hereby declare that the Industrial Project Report entitled **"Personal AI Assistant Bot"** is an authentic record of my own work as requirements of Industrial Project during the period from 01/02/2025 to 31/05/2025 for the award of degree of B.E. KLE Technological University, Hubballi under the guidance of **Dr. Tanuja Patil** and Industry guide **Ms. Priya**

- The Project Team

Date :

ABHYUDYAYA
TECHNO SOLUTIONS Pvt Ltd.

GSTIN : 29ABBCA6681J1Z9
3rd Floor Kothari Holdings, Shirur Park Main Road,
Opp. Sukruti College, Vidya Nagar, Hubballi - 580021

Internship Completion Letter

Employee Details

Pavan Shivalli
AIML Intern
Worked from 01/02/2025 to 31/05/2025

To Whom It May Concern,

This is to certify that Pavan Shivalli was employed with our company for a period of 4 months. He was hired on February 1, 2025, and his last working day with us was May 31, 2025. Pavan Shivalli held the position of AIML Intern under the Engineering team.

We sincerely appreciate Pavan Shivalli's contributions during his internship and wish him all the best in his future endeavors.

For any inquiries, clarifications, or verification, please feel free to contact Varsha Patil via email at info@abhyudyayatech.com

Abhyudyaya Techno Solutions Pvt. Ltd,

Varsha Patil
(Human Resource Manager)
Email: info@abhyudyayatech.com
Phone: 733 782 0923

ABHYUDYAYA TECHNO SOLUTIONS PRIVATE LIMITED
HUBLI.

ABHYUDYAYA
TECHNO SOLUTIONS Pvt Ltd.

GSTIN : 29ABBCA6681J1Z9
3rd Floor Kothari Holdings, Shirur Park Main Road,
Opp. Sukruti College, Vidya Nagar, Hubballi - 580021

Internship Completion Letter

Employee Details

Nitin Prakash Savvase
AIML Intern
Worked from 01/02/2025 to 31/05/2025

To Whom It May Concern,

This is to certify that Nitin Prakash Savvase was employed with our company for a period of 4 months. He was hired on February 1, 2025, and his last working day with us was May 31, 2025. Nitin Prakash Savvase held the position of AIML Intern under the Engineering team.

We sincerely appreciate Nitin Prakash Savvase's contributions during his internship and wish him all the best in his future endeavors.

For any inquiries, clarifications, or verification, please feel free to contact Varsha Patil via email at info@abhyudyayatech.com

Abhyudyaya Techno Solutions Pvt. Ltd,

Varsha Patil
(Human Resource Manager)
Email: info@abhyudyayatech.com
Phone: 733 782 0923

ABHYUDYAYA TECHNO SOLUTIONS PRIVATE LIMITED
HUBLI.

Web: http:// abhyudyayatech.com Email : info@abhyudyayatech.com Phone: +91 7337820923

ABHYUDYAYA
TECHNO SOLUTIONS Pvt Ltd.

GSTIN : 29ABBCA6681J1Z9
3rd Floor Kothari Holdings, Shirur Park Main Road,
Opp. Sukruti College, Vidya Nagar, Hubballi - 580021

Internship Completion Letter

Employee Details

Sarvesh Karakannavar
AIML Intern
Worked from 01/02/2025 to 31/05/2025

To Whom It May Concern,

This is to certify that Sarvesh Karakannavar was employed with our company for a period of 4 months. He was hired on February 1, 2025, and his last working day with us was May 31, 2025. Sarvesh Karakannavar held the position of AIML Intern under the Engineering team.

We sincerely appreciate Sarvesh Karakannavar's contributions during his internship and wish him all the best in his future endeavors.

For any inquiries, clarifications, or verification, please feel free to contact Varsha Patil via email at info@abhyudyayatech.com

Abhyudyaya Techno Solutions Pvt. Ltd,

Varsha Patil
(Human Resource Manager)
Email: info@abhyudyayatech.com
Phone: 733 782 0923

ABHYUDYAYA TECHNO SOLUTIONS PRIVATE LIMITED
HUBLI.

Web: http:// abhyudyayatech.com Email : info@abhyudyayatech.com Phone: +91 7337820923

ABHYUDYAYA
TECHNO SOLUTIONS Pvt Ltd.

GSTIN : 29ABBCA6681J1Z9
3rd Floor Kothari Holdings, Shirur Park Main Road,
Opp. Sukruti College, Vidya Nagar, Hubballi - 580021

Internship Completion Letter

Employee Details

Basavarajeshwari Haralakatti
AIML Intern
Worked from 01/02/2025 to 31/05/2025

To Whom It May Concern,

This is to certify that Basavarajeshwari Haralakatti was employed with our company for a period of 4 months. She was hired on February 1, 2025, and his last working day with us was May 31, 2025. Basavarajeshwari Haralakatti held the position of AIML Intern under the Engineering team.

We sincerely appreciate Basavarajeshwari Haralakatti's contributions during his internship and wish him all the best in his future endeavors.

For any inquiries, clarifications, or verification, please feel free to contact Varsha Patil via email at info@abhyudyayatech.com

Abhyudyaya Techno Solutions Pvt. Ltd,

Varsha Patil
(Human Resource Manager)
Email: info@abhyudyayatech.com
Phone: 733 782 0923

ABHYUDYAYA TECHNO SOLUTIONS PRIVATE LIMITED
HUBLI.

Web: http:// abhyudyayatech.com Email : info@abhyudyayatech.com Phone: +91 7337820923

ACKNOWLEDGMENT

The sense of contentment and elation that accompanies the successful completion of the Industrial project report and it would be incomplete without mentioning the names of the people who helped in accomplishing this. I absolutely respect the inspiration, support, and steering of all the people who have been instrumental in making this project fulfillment. I, being a student of KLE Technological University, am extremely grateful to Abhyudyaya Techno Solutions Pvt. Ltd. and the university for the self-assurance bestowed to us and for entrusting our mission entitled **Abhyudyaya Techno Solutions Pvt. Ltd.**

It gives me immense pleasure to express my deepest sense of gratitude and sincere thanks to my highly respected and esteemed guides **Ms. Priya** for their valuable guidance, encouragement, and help in completing this work. Their useful suggestions for this whole work and co-operative behavior is sincerely acknowledged.

I also express my gratitude to **Dr. Tanuja Patil**, University guide, for his constant support and guidance. I also wish to express my gratitude to **Dr. Suneeta V. Budihal**, HOS of School of Electronics and Communication Engineering.

I would like to express my sincere thanks to **Dr. Ashok Shettar**, Vice Chancellor, KLE Technological University, Hubli for his kind-hearted support and for giving us this opportunity to undertake this project. Also, I like to thank Mr. P.G.Tewari, Principal for his whole-hearted support.

- The Project Team

ABSTRACT

The goal of this internship project was to create Tara, a voice-activated, multilingual personal assistant robot that can automate both digital and manual duties. The objective was to create a single, real-time system that included robotic motion control, speech recognition, and multilingual support. Email scheduling, weather and news updates, photo and video capture, document creation and summarization, and five-language support are some of the main features. Using an ESP32 microcontroller for motion and sensor control through the MD10C R3 driver and ultrasonic sensors, Tara is powered by a Jetson Nano for AI computation. REST APIs, UART/Bluetooth connection, and modular Python scripts were used in the system's construction. Tara is able to obey oral instructions such as "come here" or "go forward". This project demonstrates useful AI and embedded system skills, producing a fully functional prototype for assistive technology, smart homes, and education.

Contents

1 Introduction **12**
1.1 Background of the Project 12
1.2 Objectives 13
1.3 Scope of Work 14
1.4 Problem Statement 15

2 Literature Survey **16**

3 Methodology **19**
3.1 Tools and Technologies Used :- 19
3.2 Development Process :- 21
3.3 Tara's Capabilities 22

4 Results and Discussions **26**
4.1 Result 27
4.2 Discussion 27

5 Conclusion **29**

6 References **30**

Chapter 1

Introduction

Tara is a clever personal assistant robot that can be controlled by voice commands to carry out both digital and manual activities. In addition to responding in several languages and automating over 22 tasks including snapping pictures, reading news, writing emails, and summarizing documents, it integrates AI, speech recognition, and embedded devices to comprehend users in real-time.

The system makes use of an ESP32 to operate motors and sensors and a Jetson Nano for AI computation. Whisper for voice input, gTTS for speech output, and REST APIs for online and mobile interaction are all part of its modular architecture. With the use of motor drivers and distance sensors, Tara can move robots in real time and can be controlled by voice, app, or dashboard. Tara is essentially a sophisticated, multilingual, voice-activated personal assistant for daily usage that connects hardware control and software intelligence.

1.1 Background of the Project

Intelligent systems that can help users in both digital and physical surroundings are becoming more and more necessary in today's technologically advanced world. Robots may now be designed to listen to voice commands and carry out a range of automated tasks thanks to developments in artificial intelligence, embedded systems, and natural language processing.

The goal of developing a voice-activated, intelligent robot that could automate repetitive tasks, communicate with users in different languages, and offer support in both virtual and real-world environments gave rise to the concept for the personal assistant robot Tara.

Tara is a fully functional AI-driven personal assistant robot that can perform a range of tasks in real time because to its integration of speech recognition, integrated control, and AI models.

1.2 Objectives

This project's main goal is to create a multilingual personal assistant robot that may:

- Recognizes and processes voice commands in real-time using AI.
- Performs a wide range of digital tasks such as:
    1. Email scheduling
    2. Media capture (photos, videos, audio)
    3. Information retrieval (news, weather, document summaries)
    4. Voice-based reminders
- Responds using text-to-speech across various languages (English, Hindi, Kannada).
- Interacts with physical components for robotic movement based on commands:
    1. "Come here"
    2. "Go back"
    3. "Turn left"
    4. "Stop"
    5. Sensor-based obstacle avoidance
- Offers mobile and web-based control interfaces for enhanced accessibility via APIs.
- Ensures modular, scalable architecture so that new features can be added easily.

1.3 Scope of Work

The scope of this project includes the integration of software and hardware components to develop a fully functional personal assistant robot.

Software Scope :-

- Speech Recognition: Jetson Nano runs Whisper model to transcribe voice commands in multiple languages.
- Text-to-Speech (TTS): Responses are generated using gTTS and played via Pygame.
- Command Routing: Transcribed commands are parsed and routed to the appropriate feature modules.
- Feature Modules: Modular system with 22+ features such as:
    1. Capturing photos and videos
    2. Playing music
    3. Scheduling emails
    4. Reading news and weather
    5. Generating salary slips
    6. Summarizing documents
- API Layer: REST API + Webhook support for mobile and external device communication.
- UI Layer: Mobile app (Flutter or React Native) and web dashboard for real-time interaction.

Hardware Scope :-

- Jetson Nano:
    1. Handles core AI functionalities (speech recognition, TTS, language processing).
    2. Sends control signals to ESP32 for robotic actions.
- ESP32 Microcontroller:

1. Manages motor control and sensor responses.
2. Executes real-world movement based on AI commands from Jetson Nano.

- Motor Driver (MD10C R3):

1. Controls robot movement — forward, backward, left, right, stop.

- Sensors:

1. Ultrasonic distance sensors enable obstacle detection and collision avoidance.

- Communication:

1. Jetson Nano communicates with ESP32 over UART or Bluetooth, providing flexible deployment.

- Robotic Actions Supported

1. "Go forward"
2. "Come here"
3. "Turn left"
4. "Turn right"
5. "Stop"
6. Dynamic obstacle avoidance via sensor input

1.4 Problem Statement

Voice-based solutions that provide natural, hands-free engagement for routine digital chores are becoming more and more popular. In order to ensure seamless and real-time interaction using speech recognition and text-to-speech technologies, this project intends to create an intuitive English-speaking voice assistant that can comprehend voice commands and carry out tasks like taking screenshots, responding to inquiries, and more.

Chapter 2

Literature Survey

A number of recent studies have combined automation, natural language processing, and speech recognition to further the development of voice-based intelligent systems. While Prerna Wadikar et al. improved system accessibility by providing voice automation in Hindi, Yash Mittal et al. presented an affordable Smart Home Automation System using Arduino and voice commands. Chen-Yen Peng used a Raspberry Pi and Google Voice to demonstrate voice-activated smart socket control. Abhishek Singh created a continuous on-device listening architecture in order to overcome the drawback of wake-word dependence. Po-Sheng Chiu created an emotionally intelligent campus assistant with enhanced semantic comprehension, while Youngmoon Jung concentrated on robust speaker verification in loud circumstances using deep learning. Vinayak Iyer developed a voice assistant for the blind and visually handicapped that uses BERT-based summarization to automate web interactions. Together, these pieces show how voice assistants are growing more sophisticated, contextually aware, and approachable. This serves as inspiration for the suggested system, which uses Python, speech recognition, and BERT to build a desktop voice assistant that can do tasks, summarize information, and send out reminders. [1]

Deep neural networks (DNNs) have replaced conventional Gaussian Mixture Models (GMMs) in Hidden Markov Model (HMM) frameworks for acoustic modeling, leading to notable advancements in speech recognition during the last ten years. A collaborative review by four top research groups—University of Toronto, Microsoft, IBM, and Google—that investigated the application of DNNs with multiple hidden layers trained via generative pretraining and discriminative fine-tuning is presented in the publication by Hinton et al. In contrast to GMMs, which have trouble adequately modeling non-linear data manifolds, DNNs greatly increase

recognition accuracy by successfully capturing intricate patterns in expansive acoustic environments. Significant improvements in word error rates (WERs) on benchmark datasets like TIMIT, Switchboard, and Google Voice Search have been made possible by the advent of Deep Belief Networks (DBNs), Restricted Boltzmann Machines (RBMs), and discriminative training methods like Maximum Mutual Information in particular. The basis for hybrid DNN-HMM systems that perform better than traditional models on a range of large vocabulary speech recognition tasks has been established by these developments.[2]

Recent developments in voice recognition have placed a greater emphasis on generalization across tasks and languages, scalability, and resilience. Even when they are accurate in-domain, traditional supervised models have trouble adapting to multilingual environments and distribution adjustments. By using 680,000 hours of weakly supervised internet-scale audio-text pairs in 96 different languages for training, the Whisper model put forth by Radford et al. presents a novel method. In contrast to traditional models that necessitate domain-specific fine-tuning, Whisper performs competitively in zero-shot scenarios without the need for extra oversight. In order to overcome the drawbacks of previous unsupervised techniques, such as Wav2Vec 2.0, this study incorporates a robust decoder into a multitask, multilingual Transformer framework. Scaling weak supervision can be a potent substitute for supervised training in automatic speech recognition, as demonstrated by the model's strong resilience to noise, ability to handle long-form transcription, translation, and even match or exceed human-level accuracy on multiple datasets.[3]

Natural language processing pre-training techniques have advanced dramatically, with models like as BERT and GPT introducing left-to-right generation and masked language modeling, respectively. Building on these, Lewis et al.'s BART model uses a sequence-to-sequence Transformer architecture as a denoising autoencoder framework to combine the advantages of both methods. Token masking, deletion, infilling, sentence permutation, document rotation, and other noising techniques are used to corrupt input text before BART reconstructs the original content. Its adaptable approach enables it to generalize to a variety of downstream tasks. BART produces state-of-the-art results on a variety of benchmarks, such as conversation production (ConvAI2), question answering (SQuAD, ELI5), and summarization (CNN/DM, XSum). It matches strong baselines on classification tasks and performs better, especially on generation tasks. It is a significant

development in pre-trained language models since its architecture enables a more cohesive and reliable model that supports both understanding and creation.[4]

Because of its cost, adaptability, and capacity to foster experiential learning, open-source hardware (OSHW) incorporation into educational settings has drawn a lot of interest. By allowing students to interact with real-world scientific and engineering topics, technologies such as Arduino and Raspberry Pi have been essential in fostering creativity and technological fluency. The usage of OSHW has been studied in the past in a variety of knowledge domains, including electronics, programming, and automation, as well as in a range of educational levels, from K–12 to university. However, the majority of the literature now in publication consists of experiential narratives and solution suggestions, with comparatively little empirical analyses evaluating the efficacy of education. This reveals a significant gap and indicates that more rigorous approaches and solid proof of educational impact are needed in future study. The reviewed work does a thorough systematic mapping assessment of 676 publications, providing insightful information on the development of the subject, current trends, and areas that need more research.[5]

The advantages of expanding model and data sizes to enhance performance across a range of natural language tasks have been highlighted by recent developments in large language models (LLMs). Models like GPT-3, Gopher, and PaLM have historically depended on confidential or proprietary datasets, which has limited their reproducibility and transparency among the scientific community. However, a number of initiatives, such as OPT, GPT-NeoX, and BLOOM, have concentrated on open-source methods, frequently failing to achieve state-of-the-art results. By showing that competitive foundation models (with parameters ranging from 7B to 65B) can be trained using only publically available datasets, the LLaMA models, which are presented in this study, fill this gap. LLaMA models surpass or perform on par with larger proprietary counterparts like GPT-3 and Chinchilla on a variety of benchmarks by utilizing effective architecture and training optimizations and taking inspiration from Chinchilla scaling laws. The viability and importance of transparent, repeatable large-scale language model development are highlighted by this work.[6]

Chapter 3

Methodology

The multilingual AI voice assistant Tara was developed using a human-centered, modular methodology. Across the seven Indian languages—English, Hindi, Kannada, Tamil, Telugu, Marathi, and Gujarati—accuracy, real-time performance, and organic engagement were therefore guaranteed.

For speech-to-text transcription, Tara used OpenAI's Whisper model because of its sophisticated multilingual recognition and noise resilience. It employs Google Text-to-Speech (gTTS) for voice answers, with pyttsx3 acting as a backup offline.

In order to mimic human behavior, Tara goes into sleep mode after a minute of inactivity and can be roused by sayings like "Hello Tara" or "Hi Tara." A wide range of users can utilize the system because it is designed in Python, suited for edge devices like the Jetson Orin Nano 8GB, and can be controlled via voice, text, or a mobile user interface.

3.1 Tools and Technologies Used :-

- ESP32 Microcontroller: The hardware control interface was the ESP32. It was in charge of carrying out robot movements, reading data from ultrasonic sensors, and receiving commands through Bluetooth and UART. Future scalability was also made easier by the microcontroller's integrated Bluetooth and Wi-Fi capabilities.
- Jetson Nano: The central processing unit was this edge-AI computing platform. It managed voice recognition, decision-making logic, and interaction workflows and facilitated the deployment of high-performance AI models. The processing power required for on-device real-time inference was supplied by Jetson Nano.

Figure 3.1: Illustration of NVMe storage on the Jetson Orin Nano.

- MD10C R3 Motor Driver: The robot's two-wheel drive system, which allowed for left, right, forward, and backward movements, was controlled by this motor driver. It converted the ESP32's PWM and direction signals into motion that was controlled by power.
- HC-SR04 Ultrasonic Sensor: Used to prevent collisions and detect obstacles. The ESP32 was able to react to changing surroundings thanks to the sensor data, stopping or rerouting movement as needed.
- OpenAI Whisper: Multilingual voice-to-text transcription was done using Whisper. Tara Bot was able to serve as a multilingual voice assistant due to its capacity to handle a variety of accents and languages.
- gTTS (Google Text-to-Speech): Used for generating spoken responses in different languages. This enhanced Tara's human-like interaction capability.
- Arduino IDE: Utilized for programming the ESP32 microcontroller. It facilitated code deployment and debugging during hardware interaction development.

3.2 Development Process :-

1. Requirement Analysis: The goal of this phase was to determine which functionalities Tara should provide. Cross-platform communication, sensor feedback, physical robot movement, and AI-based speech interaction were important topics.
2. System Architecture Design: There were two logical layers to the system. The ESP32 microcontroller managed real-world robotic movements, while the Jetson Nano managed advanced AI and voice functions. Asynchronous command handling and communication were made possible by the system design.
3. Voice Interaction and AI Integration:
    - Whisper was integrated to provide accurate real-time transcription.
    - A command mapping logic was developed to associate spoken phrases with robot actions.
    - gTTS converted textual outputs into audible feedback.
    - A fallback mechanism was added to handle misunderstood or unsupported commands.
4. Hardware Control and Integration:
    - ESP32 was programmed to interpret commands from Jetson Nano.
    - Motor direction and speed were controlled using PWM logic and MD10C R3 motor driver.
    - Ultrasonic sensor input triggered stop or redirection actions to prevent collisions.
5. Feature Module Development:
    - A total of 22 task-specific modules were created. These included photo capture, video recording, alarm setup, email scheduling, salary computation, document summarization, and news/weather updates.
    - Each module operated independently and was triggered via a unified command routing logic.
6. Testing and Iteration:

Figure 3.2: Launching 6.2 Jetpack on Jetson Orin Nano.

- Software modules were tested in isolation and then in combination.
- Multilingual voice command testing ensured functional speech interaction across five languages.

3.3 Tara's Capabilities

Figure 3.3: Testing Features

- **speech_recognition.py**
  - Uses OpenAI Whisper to convert real-time speech to text.
  - Handles noise filtering, audio recording, and transcribing voice commands in multiple languages.
  - Central to understanding user input and enabling natural interaction.
- **tts_output.py**
  - Converts text responses into human-like speech using gTTS (Google Text-to-Speech).
  - Plays audio via gTTS for fast response.
- **language_selector.py**
  - Prompts the user to choose a language (English, Hindi, Kannada, etc.).
  - Remembers and sets the interaction language globally.
  - Ensures Tara speaks and understands the user in their preferred language.
- **command_router.py**
  - Main logic that matches spoken commands to feature modules.
  - Imports all available features and routes commands accordingly.
  - Uses fuzzy matching and keyword identification to handle variation in phrasing.
- **feedback.py**
  - Handles unclear or invalid user input.
  - Provides helpful suggestions, confirmation prompts, or fallback responses.
  - Improves Tara's interactivity and user guidance.
- **snap.py**
  - Takes pictures using OpenCV.
  - Can also display the latest photo.

  - Triggered with commands like "Take a snap" or "Show the last photo".

- **video.py**
  - Records video using OpenCV and webcam.
  - Can also play or show the latest video clip.
  - Integrates with `cv2` for media capture.

- **audio.py**
  - Records audio or plays audio clips.
  - Useful for tasks like voice memos.

- **weather.py**
  - Gets real-time weather and 5-day forecast using OpenWeatherMap API.
  - Supports fuzzy city name matching via `city.list.json`.
  - Uses voice interaction to confirm city names and deliver audio reports.

- **news.py**
  - Fetches top headlines from news APIs like NewsAPI.
  - Can summarize news and speak it out loud.

- **email.py**
  - Handles sending and reading emails via voice.
  - Integrates Gmail login, Google People API for contact lookup, and email scheduling.

- **navigation.py**
  - (Optional placeholder for GPS-based or map-related services).
  - Can be integrated with Google Maps or similar services for direction-related queries.

- **task_automation.py**
  - Automates tasks like setting reminders, alarms, and scheduling meetings.

  - Responds to voice prompts to define the task content and time.

- **salary_generator.py**
  - Calculates monthly/yearly salary from spoken input (e.g., "I earn 30 thousand rupees").
  - Supports Hindi, Kannada, and English.
  - Parses salary phrases including lakhs, crores, and paisa.

- **document_summary.py**
  - Summarizes the content of uploaded or dictated documents.
  - Uses NLP to extract and simplify important information.
  - Ideal for reading long reports, legal documents, etc.

Chapter 4

Results and Discussions

Figure 4.1: Interactive Service Robot Integrated with Tara AI Assistant

When it came to responding promptly in English and managing voice commands, the Tara assistant functioned admirably. A seamless user experience was facilitated by its comprehension of natural speech, accuracy in task execution, and human-like response. Under various test conditions, essential capabilities like wake word detection, photo/video management, and screenshot capture operated dependably. High transcription and speech quality were guaranteed with the usage of gTTS and Whisper. All things considered, the assistant achieved its design objectives and demonstrated itself to be a productive, amiable, and interactive voice-based

system for routine digital chores.

4.1 Result

Several voice-driven tasks in English were successfully tested and deployed using the Tara assistant. Smooth interaction flow and dependable real-time performance were displayed by the system. When Tara first started up, it gave users voice-based instructions and greeted them contextually according to the time of day. Whisper was used to accurately identify voice instructions, and Google Text-to-Speech was used to provide a nice, human-sounding response.

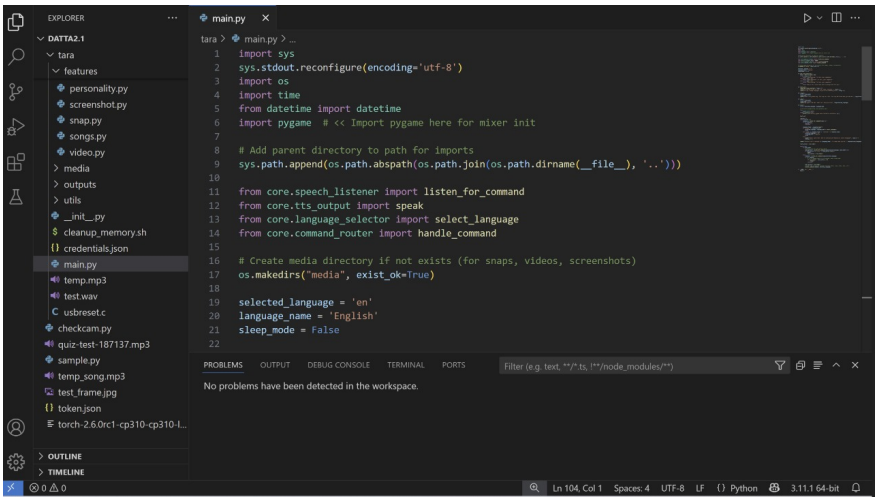

Figure 4.2: Tara Assistant's Boot Sequence

4.2 Discussion

A voice-driven interface can greatly enhance the way consumers interact with digital systems, as demonstrated by the creation and testing of the English-language Tara assistant. High-accuracy transcriptions were produced by integrating OpenAI's Whisper model for speech recognition, which made it appropriate even in rather loud settings. Likewise, Google Text-to-Speech made sure Tara's answers were understandable, cordial, and human-like, which helped to facilitate a seamless dialogue.

The assistant's reactivity is among its most noteworthy features. By lowering idle resource utilization and guaranteeing Tara was always prepared to react promptly when called upon, the wake word detection and sleep

mode implementation improved system performance. Fuzzy matching allowed the assistant to comprehend different command phrasings, which increased the system's adaptability and user-friendliness.

Chapter 5

Conclusion

Voice-driven systems have the ability to improve and streamline consumer contact with technology, as the Tara assistant effectively illustrates. Tara offers English-speaking customers a responsive and amiable experience by combining real-time speech detection with Whisper and natural voice answers with gTTS. Essential features like command recognition, wake word detection, and modular feature execution were successfully implemented. The architecture is made for easy scalability, thus even though this version only supports English, it is ready for future multilingual and multimodal improvements. All things considered, Tara represents a major advancement in the development of intelligent, user-focused voice assistants.

Chapter 6

References

1. Rose Thomas, Surya V S, Tincy A Mathew, Tinu Thomas.VOICE BASED INTELLIGENT VIRTUAL ASSISTANT FOR WINDOWS USING PYTHON. ICCIDT - 2023 Conference Proceedings.

2. Geoffrey Hinton, Li Deng, Dong Yu, George E. Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent Vanhoucke, Patrick Nguyen, Tara N. Sainath, and Brian Kingsbury. Deep Neural Networks for Acoustic Modeling in Speech Recognition. 15 October 2012

3. Alec Radford Jong Wook Kim Tao Xu Greg Brockman Christine McLeavey Ilya Sutskever. Robust Speech Recognition via Large-Scale Weak Supervision. 1OpenAI, San Francisco, CA 94110, USA.

4. Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Ves Stoyanov, Luke Zettlemoyer. BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension. 29 Oct 2019.

5. RUBEN HERADIO, JESUS CHACON JACOBO SAENZ, LUIS DE LA TORRE, HECTOR VARGAS, DANIEL GALAN ,ANDSEBASTIAN DORMIDO. Open-Source Hardware in Education: A Systematic Mapping Study. November 16, 2018.

6. Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet Marie-Anne Lachaux, Timothee Lacroix, Baptiste Rozière, Naman Goyal Eric Hambro, Faisal Azhar, Aurelien Rodriguez, Armand Joulin Edouard Grave, Guillaume Lample. LLaMA:OpenandEfficient Foundation Language Models. 27 Feb 2023.

re1

ORIGINALITY REPORT

8%	8%	3%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.coursehero.com Internet Source	5%
2	myhq.in Internet Source	1%
3	www.kscst.iisc.ernet.in Internet Source	1%
4	www.slideshare.net Internet Source	<1%
5	dspace.daffodilvarsity.edu.bd:8080 Internet Source	<1%
6	jnnce.ac.in Internet Source	<1%
7	department-of-electronics-and-communication-engineering.newhorizoncollegeofengineering.in Internet Source	<1%
8	www.pressnews.biz Internet Source	<1%

Exclude quotes	On	Exclude matches	< 5 words
Exclude bibliography	On		