iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** IV **Month of publication:** April 2025

DOI: https://doi.org/10.22214/ijraset.2025.69634

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

Building an Offline Virtual Voice Assistant using AI and NLP

Mrs. P. Jhansi Lakshmi[1], D. Sowmya[2], A. Alekhya[3], N. Kalyani[4], P. Swarna Latha[5]
[1]*MTech, Computer Science & Engineering, Bapatla Women's Engineering College, Bapatla, AP, India*
[2, 3, 4, 5]*BTech, Computer Science & Engineering, Bapatla Women's Engineering College, Bapatla, AP, India*

***Abstract:** As human-computer interaction continues to evolve, voice assistants like Siri, Alexa, and Google Assistant have become widely used for simplifying everyday tasks. However, these platforms often depend on a stable internet connection and cloud-based infrastructure, which can lead to privacy concerns, slower response times, and reduced functionality in offline or low-connectivity environments. To overcome these limitations, this project presents the development of an Offline Virtual Voice Assistant. By utilizing Artificial Intelligence (AI) and Natural Language Processing (NLP), the assistant can understand and respond to voice commands without relying on internet access. It is capable of executing various offline tasks, such as setting alarms, accessing locally stored data, managing device settings, and holding basic voice-based interactions. The system employs efficient speech recognition and text-to-speech engines that are tailored for use on low-power edge devices.*

I. INTRODUCTION

Voice-based interactions have become a central aspect of modern technology, with virtual assistants like Siri, Alexa, and Google Assistant playing a vital role in simplifying daily tasks through voice commands. These platforms enable users to perform actions such as setting reminders, playing music, and controlling smart devices effortlessly. However, a significant limitation is their dependence on cloud services and a stable internet connection, which can lead to concerns regarding privacy, slower response times, and limited functionality in offline or low-connectivity environments. To overcome these challenges, this project introduces the development of an Offline Virtual Voice Assistant designed to function independently of the internet. By leveraging Artificial Intelligence (AI) and Natural Language Processing (NLP), the assistant processes and responds to voice commands locally. It is capable of performing a range of offline tasks, including setting alarms, retrieving stored data, and adjusting device settings. The system incorporates lightweight speech recognition and text-to-speech technologies optimized for edge devices, ensuring a fast, secure, and reliable user experience, even without internet access.

II. EXISTING SYSTEM

While most mainstream voice assistants (like Siri, Alexa, and Google Assistant) require cloud connectivity, several open-source and commercial tools are available for building offline-capable voice assistants. These tools focus on the local processing of voice commands using lightweight models suitable for edge devices.

III. PROPOSED SYSTEM

The proposed system is designed to function entirely offline, enabling users to interact with their devices through voice commands without relying on cloud-based services or internet connectivity. It emphasizes data privacy, low latency, and efficient performance on edge devices such as mobile phones or embedded systems.

IV. IMPLEMENTATION

To implement an offline virtual voice assistant, we combine several key components to ensure functionality without relying on the internet. The system begins with wake word detection, where the assistant listens for a predefined trigger (e.g., "Hey Assistant") using a lightweight engine like Porcupine. Once the wake word is detected, the assistant captures the user's speech and converts it to text using offline speech-to-text models such as Vosk or Whisper. The transcribed text is then processed to determine the user's intent through Natural Language Understanding (NLU) using libraries like Rasa NLU or a custom model. Based on the recognized intent, the system executes predefined commands, such as setting alarms or adjusting device settings. The assistant then responds with text-to-speech (TTS) output using offline solutions like Coqui TTS or Mimic3, ensuring all actions and responses occur without internet connectivity.

This setup ensures data privacy, low latency, and reliable performance on edge devices such as Raspberry Pi or smartphones, making it an efficient, self-contained system. The proposed system is designed to function entirely offline, enabling users to interact with their devices through voice commands without relying on cloud-based services or internet connectivity. It emphasizes data privacy, low latency, and efficient performance on edge devices such as mobile phones or embedded systems.

V. MODULES

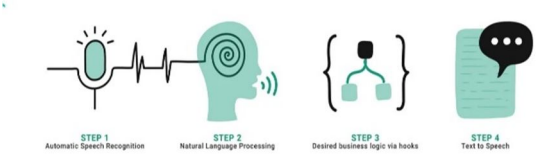
1) Speech Recognition - Converts voice to text
2) NLP Module - Understands user intent
3) Intent & Response - Classifies intent and generates response
4) Text-to-Speech (TTS) - Converts text to speech
5) Integration - Connects to third-party services

VI. ALGORITHMS

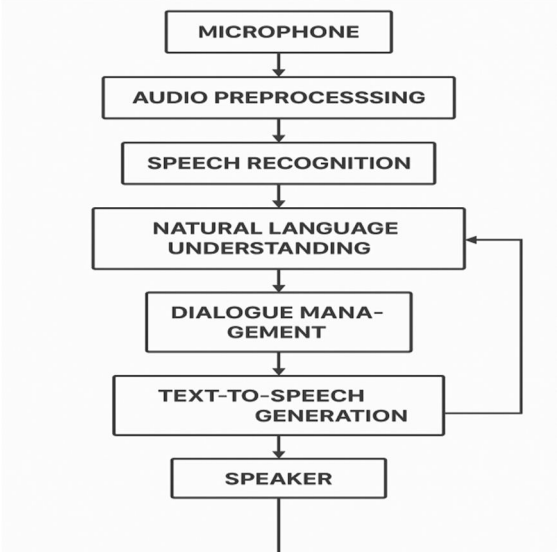
1) Regex pattern Extraction: Virtual assistants use regex to pick out specific details from messy or unstructured text, making it easier to understand and carry out what the user wants.
2) Keyword Matching: In NLP, keyword matching helps recognize certain words or phrases in what the user says, so the system can respond or take the right action.
3) Threading Alarm: Threading helps voice assistants handle tasks like alarms in the background while still responding to your commands.

VII. SYSTEM ARCHITECTURE

The diagram illustrates the four key steps of the voice-assistant workflow. It begins with Automatic Speech Recognition (ASR), wherein spoken words are converted into text. Natural Language Processing (NLP) interprets the meaning of the text. In Step 3, the assistant processes the request using a predefined logic or hooks to determine the appropriate action. Finally, Text-to-speech (TTS) converts the response back into spoken words, enabling the assistant to reply vocally. This seamless flow enables an interactive and intelligent voice communication.



VIII. DATAFLOW DIAGRAM



IX. RESULT AND ANALYSIS

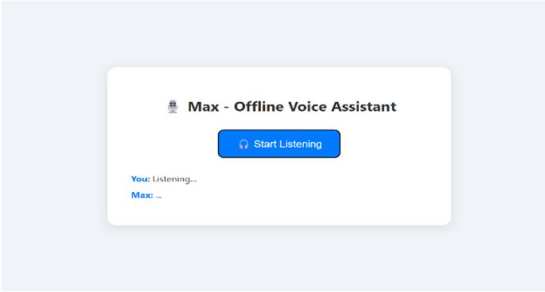


Figure 1: Voice as input

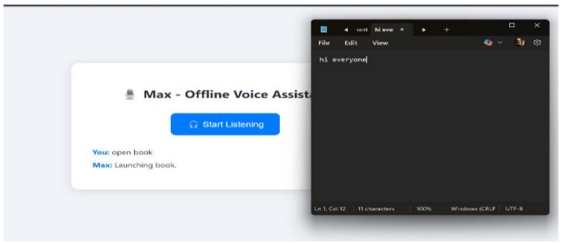


Fig2: Opening Book Feature

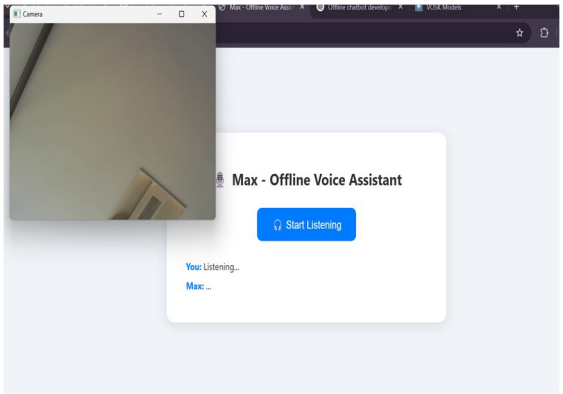


Fig 3: Camera Preview

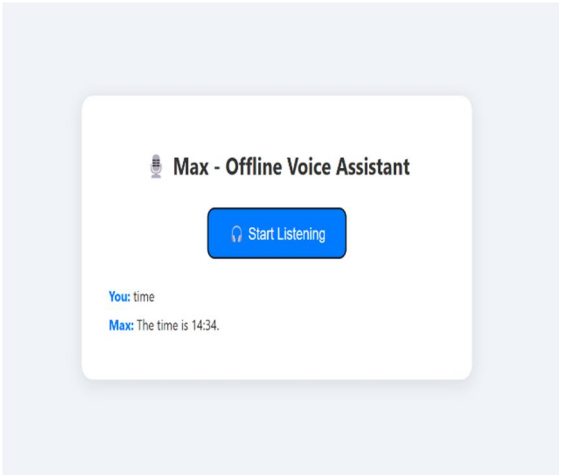


Fig 4: Time Response

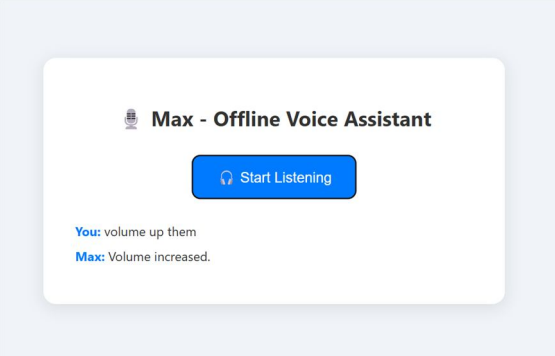


Fig 5: Volume Control Response

X. CONCLUSION

An offline virtual voice assistant provides fast, secure, and reliable speech recognition without needing an internet connection, ensuring user privacy and functionality in any environment. It's ideal for applications where internet access is limited or data security is crucial.

XI. FUTURE SCOPE

Ensure data privacy by processing voice commands locally without cloud access. Enable reliable performance in remote or low-connectivity environments. Power smart devices and on-device processing.

REFERENCES

[1] SELMA: A Speech-Enabled Language Model for Virtual Assistant Interactions (2025)
 https://arxiv.org/abs/2501.19377
[2] Jaco: An Offline Running Privacy-aware Voice Assistant (2022)
 https://arxiv.org/abs/2209.07775
[3] Distilling an End-to-End Voice Assistant Without Instruction Training Data (2024)
 https://arxiv.org/abs/2410.02678
[4] Privacy Preserving Personal Assistant with On-Device Diarization and Spoken Dialogue System (2024)
 https://arxiv.org/abs/2401.01146
[5] Enhancing Voice Assistant Systems through Advanced AI and NLP Techniques (2025)
 https://jricst.com/index.php/JRICST/article/view/19
[6] Jarvis-Virtual Voice Assistant (2024)
 https://qtanalytics.in/publications/index.php/JoCSVL/article/view/219
[7] An Efficient Virtual Voice Assistant for Physically Challenged People (2023)
 https://www.ijert.org/an-efficient-virtual-voice-assistant-for-physically-challenged-people

AUTHOR'S PROFILES

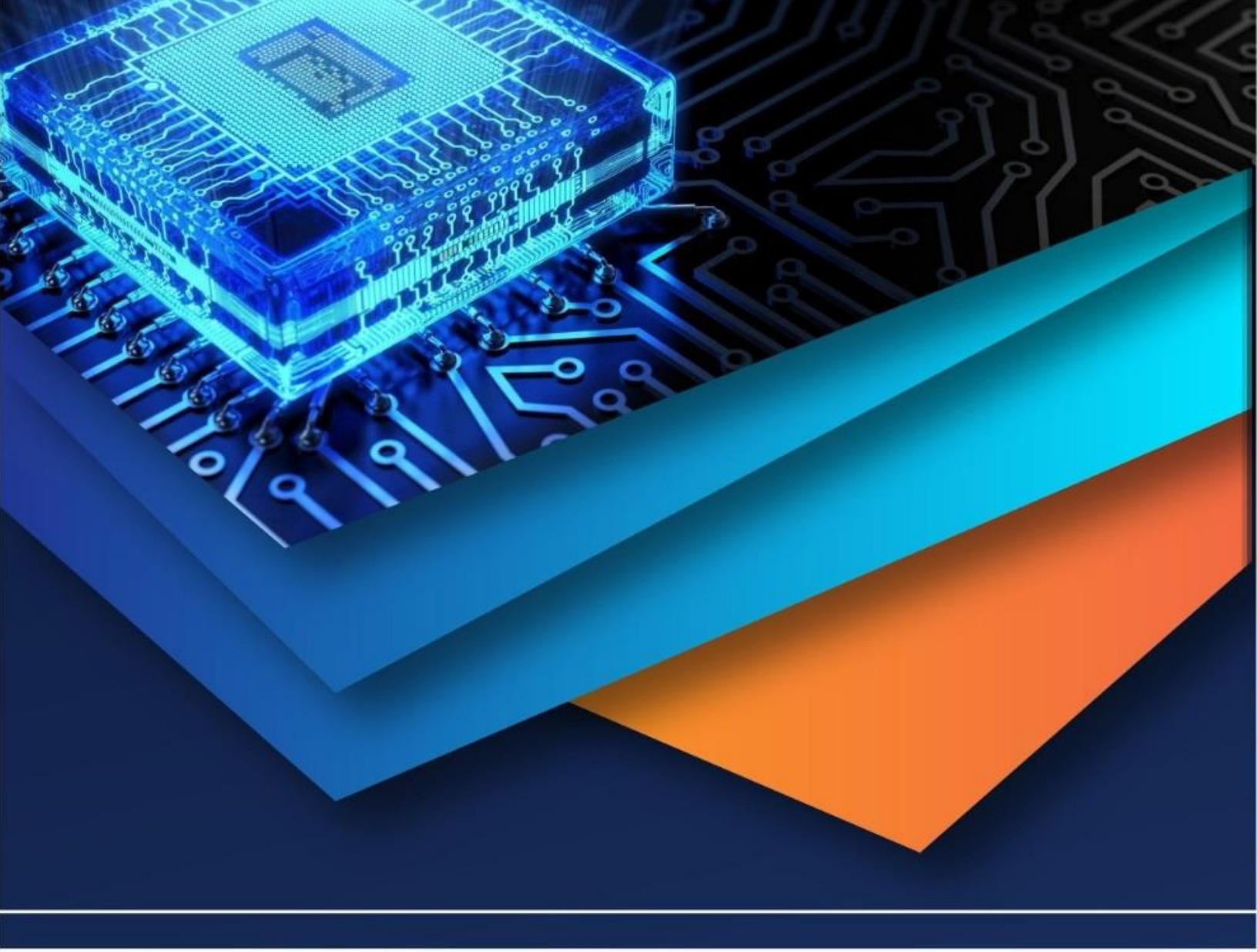
Mrs. P. Jhansi Lakshmi is working as an Assistant Professor in the Department of CSE, BWEC, Bapatla.

D. Sowmya B. Tech with specialization in Artificial Intelligence and Machine Learning at Bapatla Women's Engineering College, Bapatla.

A.Alekhya B.Tech with specialization in Artificial Intelligence and Machine Learning at Bapatla Women's Engineering College, Bapatla.

P. Swarna Latha B. Tech with specialization of Artificial Intelligence and Machine Learning in Bapatla Women's Engineering College, Bapatla.

P. Swarna Latha B. Tech with specialization of Artificial Intelligence and Machine Learning in Bapatla Women's Engineering College, Bapatla.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY