iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IV **Month of publication:** April 2026

DOI: https://doi.org/10.22214/ijraset.2026.81511

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 14 Issue IV Apr 2026- Available at www.ijraset.com

Voice-Enabled Smart Email Assistance for the Visually Impaired

Mrs. P. Saraswathi[1], V. Ramyasri[2], P. Indu[3], K. Sunethra[4]
[1]*Assistant Professor, Department of Computer Science & Engineering Bapatla Women's Engineering College Bapatla, AP, INDIA*
[2,3,4]*B. Tech Computer Science & Engineering, Bapatla Women's Engineering College Bapatla, AP, INDIA*

Abstract: ***This project focuses on building an easy-to-use system that helps visually impaired users manage their emails using voice commands. The system uses Artificial Intelligence techniques along with Speech Recognition and Natural Language Processing (NLP) to enable users to read, send, and respond to emails without relying on a screen. The system processes voice input, converts it into text, and performs actions such as retrieving emails and generating responses.***
To make the system more user-friendly, a web-based application is developed that allows users to interact with emails through simple voice commands. The system can read out email content using text-to-speech and also includes features like email summarization and priority detection to help users quickly understand important messages.
Keywords: Voice-Based Email System, Speech Recognition, Natural Language Processing (NLP), Email Automation, Text-to-Speech (TTS), Email Summarization, Accessibility for Visually Impaired

I. INTRODUCTION

Email communication plays an important role in our daily lives, helping people share information quickly and efficiently. However, managing emails can be difficult for visually impaired users, as most email systems depend heavily on visual interaction. Tasks such as reading messages, composing replies, and navigating through inboxes require constant screen usage, which creates challenges and limits accessibility. To overcome these issues, this project focuses on developing a voice-enabled system that allows users to manage emails using simple voice commands.

With the rapid growth of digital communication, the number of emails people receive every day has increased significantly. Managing these emails manually can be time-consuming and may lead to missing important information. For visually impaired individuals, this problem becomes even more serious, as they often rely on screen readers or external assistance. While these tools provide basic support, they may not always be efficient or user-friendly, especially when handling large volumes of emails.

To address these challenges, modern technologies such as Artificial Intelligence and Natural Language Processing (NLP) can be used to create smarter solutions. Speech recognition allows the system to understand voice commands, while text-to-speech technology helps convert email content into audio. These technologies make it possible to design systems that can interact with users in a natural and convenient way, reducing the need for visual interfaces.

In this project, we have developed a system that allows users to access, read, and respond to emails through voice interaction. The system retrieves emails using standard protocols and processes them using NLP techniques. It can summarize long emails, detect important messages, and provide responses based on user commands. This helps users quickly understand the content of emails without reading them manually. In this project, we have built a deep learning model that is trained to classify different types of blood cells. The dataset used contains images of various types of blood cells, including neutrophils, eosinophils, lymphocytes, monocytes, and more. Each image is labelled based on the type of cell it represents. We used data augmentation techniques to increase the diversity of images during training, which helps the model learn better and generalize well to new, unseen images.

The system architecture is designed to ensure smooth and accurate performance. It includes modules for speech recognition, email retrieval, text processing, and voice output. Each module works together to provide a seamless user experience. The system is capable of handling real-time voice commands and delivering responses with good accuracy.

To make the system more accessible, a simple and user-friendly interface is developed. Users can interact with the system without requiring any technical knowledge. The interface allows users to perform actions such as reading emails, replying, and managing messages easily through voice commands. This reduces effort and improves overall efficiency.

Another useful feature of this project is the ability to summarize email content. This helps users save time by focusing only on important information. Additionally, the system can identify priority emails, allowing users to respond to urgent messages quickly. These features improve productivity and make email management more efficient.

The system is designed to be flexible and scalable, making it suitable for use in different environments such as personal use, education, and business communication. It can also be extended in the future with additional features like better language support and improved voice recognition accuracy.

This approach of combining Artificial Intelligence with a simple voice-based interface makes the system both effective and easy to use. It does not require users to have any technical knowledge of email systems or programming. Users can simply give voice commands, and the system handles tasks such as reading, summarizing, and responding to emails automatically. In the future, this system can be further improved and integrated into mobile applications or smart devices to provide wider accessibility and better support for visually impaired users.

Overall, this project aims to create an efficient, reliable, and user-friendly solution for voice-based email management. It demonstrates how Artificial Intelligence can be used to improve accessibility and support visually impaired users in their daily communication. By reducing manual effort and simplifying interaction, the system provides a practical approach to modern email handling.

II. LITERATURE SURVEY

Over the past decade, the use of Artificial Intelligence in email management and communication systems has grown significantly. Researchers have focused on developing intelligent systems that can automate email processing using technologies such as speech recognition and natural language processing (NLP). These approaches have proven to be faster and more efficient compared to traditional manual methods, which often depend on user effort and constant interaction with screens.

In earlier systems, email management relied on rule-based filtering and basic machine learning techniques for tasks such as spam detection and classification. While these systems provided some level of automation, they were limited in understanding the context and meaning of email content. Users still had to manually read and respond to emails, which made the process time-consuming, especially when handling large volumes of messages. With the advancement of deep learning and NLP, more intelligent solutions have been developed. Techniques such as text summarization and intent recognition allow systems to process email content more effectively. Modern models are capable of generating meaningful summaries and identifying important information within emails. These improvements have made it easier for users to quickly understand email content without reading it entirely. Speech recognition technology has also played a major role in improving accessibility. It enables systems to convert spoken commands into text, allowing users to interact with applications using voice. This is particularly useful for visually impaired users, as it reduces the need for screen-based interaction. Combined with text-to-speech technology, systems can both understand user input and provide audio output, creating a complete voice-based interaction environment. Recent research has focused on integrating multiple technologies to build smarter email systems. Features such as email summarization, priority detection, and multilingual support have been introduced to improve user experience. These systems are designed to handle real-time processing and provide faster responses, making email management more efficient and user-friendly. In addition, the use of cloud-based APIs and scalable architectures has made it easier to develop and deploy such systems. These technologies allow applications to handle large amounts of data and support multiple users without performance issues. As a result, modern email systems are becoming more reliable and adaptable to different use cases. Despite these advancements, challenges still remain. Issues such as background noise in voice input, variations in user accents, and maintaining accuracy in summarization can affect system performance. Researchers continue to work on improving these areas to make voice-based systems more robust and reliable. In summary, the literature shows a clear shift from traditional email handling methods to intelligent AI-based systems. The integration of speech recognition, NLP, and automation has significantly improved the efficiency and accessibility of email communication. These developments provide a strong foundation for building voice-enabled email assistance systems, especially for visually impaired users. Speech recognition and Natural Language Processing (NLP) have significantly transformed how users interact with digital systems by enabling more natural and intuitive communication. In recent years, these technologies have been widely applied in applications such as virtual assistants, automated customer support, and accessibility tools. Researchers have shown that NLP-based systems can effectively understand user intent, process textual information, and generate meaningful responses. Techniques such as text summarization and intent recognition allow systems to extract important information from large volumes of text, making them highly useful for email management. Several comparative studies have highlighted the differences between traditional machine learning methods and modern deep learning approaches in text processing tasks. For instance, methods like Support Vector Machines (SVM) and Random Forest perform well when working with structured data and predefined features. However, they often struggle with unstructured text data such as emails. In contrast, deep learning models are capable of understanding context and handling variations in language more effectively, resulting in better performance in tasks like email classification and summarization. Despite these advantages, deep learning models generally require more computational resources and larger datasets.

Recent research has also focused on hybrid approaches that combine multiple techniques to improve system performance. For example, integrating speech recognition with NLP-based processing allows systems to accurately convert voice input into meaningful actions. These combined approaches help in improving both accuracy and efficiency, especially in real-time applications such as voice-based email systems. Such models provide a balanced solution by leveraging the strengths of different technologies, making them suitable for practical and scalable implementations.

III. METHODOLOGY

The methodology of this project involves designing and developing a voice-enabled system that allows users to manage emails using speech-based interaction. The system is built by integrating speech recognition, natural language processing (NLP), and text-to-speech technologies. The overall process includes collecting user input through voice, processing it into meaningful commands, retrieving emails, and providing responses in audio format. The system is designed to ensure real-time performance, accuracy, and ease of use.

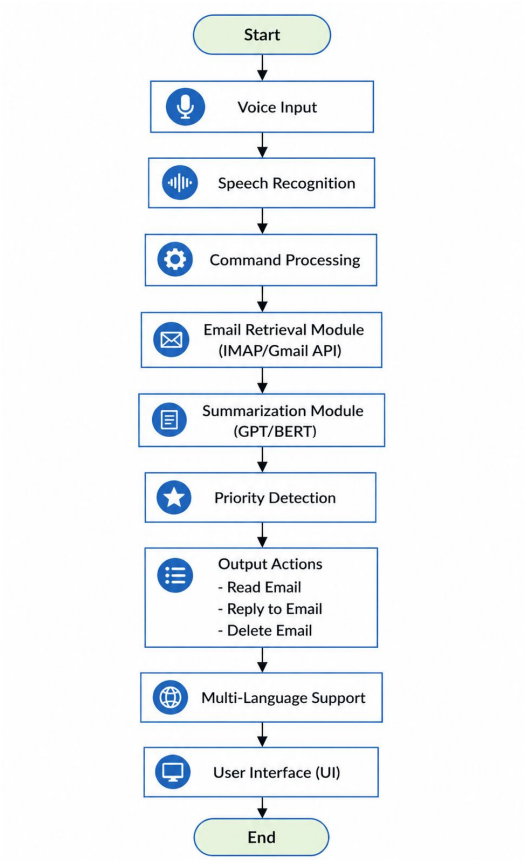
A. Data Collection

The data collection phase is important for building the Voice-Enabled Smart Email Assistance System. The system mainly uses two types of data: voice input from users and email data from email servers.

Voice data is collected in real time through a microphone when the user gives commands such as reading or replying to emails. This input is then sent to the speech recognition module for processing.

Email data is retrieved using IMAP or Gmail APIs. The system collects details like sender, subject, and email content, which are used for tasks such as summarization and priority detection.

Diagram 1: System Architecture



After retrieving emails, the system processes the content through several modules to provide meaningful output to the user. The Summarization Module analyzes the email text and generates a short and clear summary using NLP techniques. This helps users quickly understand the main idea of the email without reading the full content.

Next, the Priority Detection Module evaluates the importance of the email by analyzing keywords, subject, and sender information. Based on this analysis, emails are classified as priority or normal, allowing users to focus on important messages.
The system then moves to the Output Actions Module, where it performs operations based on user commands. These actions include reading emails, replying to messages, or deleting emails. The results are processed and prepared for output.
The Multi-Language Support Module ensures that the content can be translated into different languages if required, improving accessibility for users.
Finally, the processed output is delivered through the User Interface, where the results are displayed on the screen and also converted into speech using text-to-speech technology, enabling hands-free interaction.

B. Pre - processing

Pre-processing is an important step in the Voice-Enabled Smart Email Assistance System, as it ensures that the user's voice input and email data are properly prepared for accurate processing. Based on the system workflow shown in the activity diagram, preprocessing begins with capturing the user's voice input and converting it into a usable format.
The first step involves voice input acquisition, where the system records the user's speech through a microphone. This raw audio input may contain background noise or variations in speech. Therefore, basic noise reduction and normalization techniques are applied to improve clarity and ensure better recognition accuracy.
Once the audio is cleaned, it is passed to the speech recognition module, where the voice input is converted into text. This conversion is essential for further processing, as the system works with textual data to understand user commands.
After conversion, the text is processed in the command processing stage, where Natural Language Processing (NLP) techniques are applied. This includes tokenization, removal of unnecessary words, and identification of key phrases. The system then maps these processed inputs to specific actions such as reading emails, replying, or deleting messages.
In parallel, the system also prepares email data for processing. Retrieved emails are cleaned by removing unnecessary characters, formatting issues, and irrelevant content. This ensures that the data is structured and suitable for further analysis in the summarization and priority detection modules.
Additionally, the processed data is organized in a structured format to support real-time execution. This improves system performance and reduces delay during user interaction. The preprocessing phase also handles variations in user input, such as different accents or speaking styles, to make the system more robust.
Overall, the preprocessing stage ensures that both voice input and email content are clean, structured, and ready for further processing. This step plays a key role in improving accuracy, reducing errors, and enabling smooth operation of the entire system.

C. Email Processing and Interaction

Email processing is the task of analysing and handling email data based on user commands. It involves using known inputs, such as voice commands and email content, to perform actions like reading, replying, or managing emails. In this project, a combination of speech recognition and Natural Language Processing (NLP) techniques is used to process emails efficiently.

1) Voice-Based Email Interaction System

In this project, an intelligent system is developed to enable users to manage emails using voice commands. The system is designed to perform tasks such as reading emails, summarizing content, replying to messages, and deleting emails without requiring visual interaction.
The system processes user input through a speech recognition module, which converts voice commands into text. These commands are then analyzed using NLP techniques to understand user intent. The email data is retrieved using IMAP or Gmail APIs, ensuring real-time access to inbox messages.
The processing module applies techniques such as text cleaning and summarization to extract important information from emails. Long emails are reduced into short and meaningful summaries, allowing users to quickly understand the content. Additionally, a priority detection mechanism is used to identify important emails based on keywords and context.
The system architecture is designed to ensure smooth and accurate performance. It includes modules for speech recognition, command processing, email retrieval, summarization, and text-to-speech output. These modules work together to provide a seamless user experience.
The system is implemented using efficient technologies to ensure reliable performance. Voice commands are processed in real time, and responses are delivered through audio output. The system can handle multiple commands and provides accurate results under normal conditions.

To enhance usability, a simple and user-friendly interface is developed. Users can interact with the system easily without requiring technical knowledge. The interface allows users to perform all email-related operations through voice interaction.
This approach combines efficient data processing, intelligent NLP techniques, and a user-friendly interface to create a practical and accessible email management system.
The implementation of the system demonstrates reliable performance and ease of use. The system successfully processes voice commands, retrieves emails, and provides accurate responses. It also supports features such as summarization and priority detection, improving overall efficiency.

2) Use Case Diagram

The Use Case Diagram shows how the user interacts with the Voice-Enabled Smart Email Assistance System. It highlights the main functionalities provided by the system and how they are accessed through voice commands.

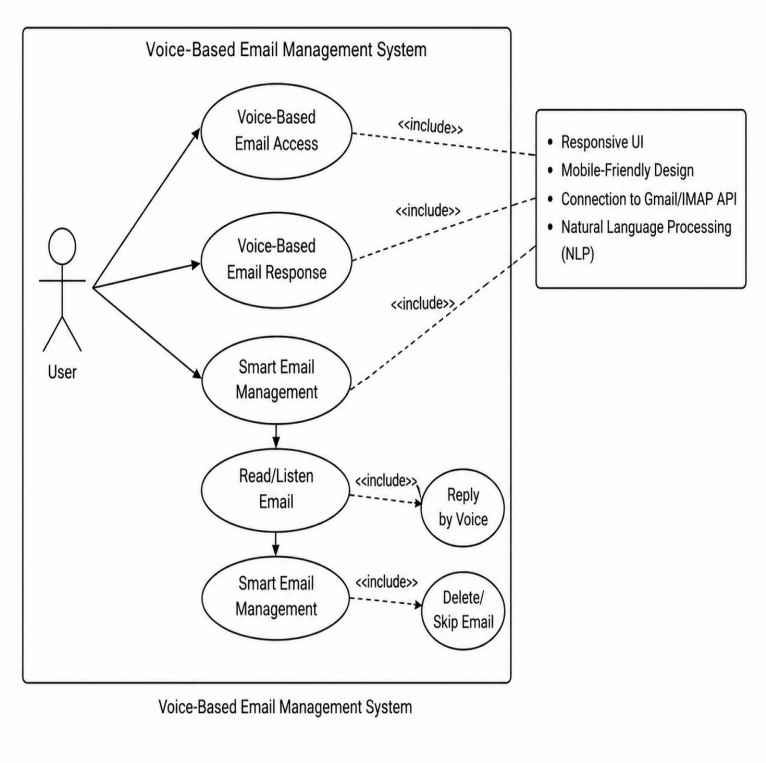


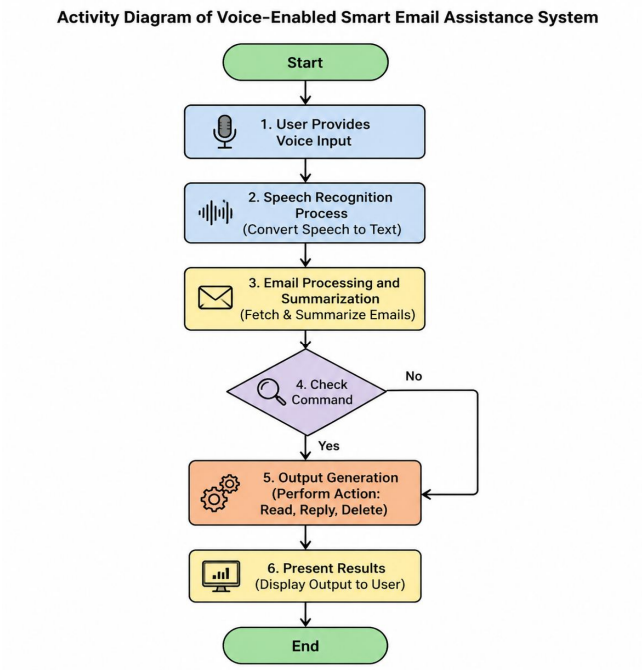
Fig 2.3.2: The Use Case Diagram of Voice-Based Email System

The user is the primary actor and can perform actions such as voice-based email access, replying to emails, and managing emails. The system allows users to listen to emails without opening the inbox, making it highly useful for visually impaired individuals.
The voice-based email response feature enables users to reply to emails using speech, providing a hands-free experience. The smart email management functionality allows users to read, delete, or skip emails easily. Additionally, the system includes features like Natural Language Processing (NLP), Gmail API integration, and responsive UI, which support smooth and efficient operation.
Overall, the diagram demonstrates how the system simplifies email interaction by converting voice commands into meaningful actions, improving accessibility and user experience.

3) Activity Diagram

The Activity Diagram illustrates the workflow of the Voice-Enabled Smart Email Assistance System, showing the sequence of operations from capturing user input to generating the final output. The main steps involved in the process are as follows:

- User Provides Voice Input: The user gives a voice command, which is captured by the system through a microphone.
- Speech Recognition Process: The captured voice input is converted into text using speech recognition technology.
- Email Processing and Summarization: The system retrieves relevant emails and processes them. It generates concise summaries to help the user quickly understand the content.



- Command Verification: The system checks the user's command to determine the appropriate action to be performed.
- Output Generation: Based on the command, the system performs actions such as reading emails, replying, or deleting messages.
- Result Presentation: The processed results are presented to the user through the interface in an accessible format.

IV. IMPLEMENTAION

The implementation of the Voice-Enabled Smart Email Assistance System involves several key steps to ensure smooth functionality and proper integration of components.

1) System Development Environment: The system is developed using Python as the primary programming language for backend processing. A web framework such as Flask is used to build the application and handle communication between frontend and backend. The frontend interface is designed using HTML, CSS, and JavaScript to provide a simple and user-friendly experience.

2) Speech Recognition Module: The speech recognition functionality is implemented using libraries such as Speech Recognition or APIs. This module captures voice input from the user and converts it into text format. Basic audio preprocessing techniques are applied to improve recognition accuracy.

3) Text Processing Module: Once the speech is converted into text, it is processed using Natural Language Processing (NLP) techniques. The system cleans and analyses the text to identify user commands such as reading emails, replying, or deleting messages.

4) Email Retrieval Module: The system retrieves emails using IMAP or Gmail APIs. It fetches email details such as sender, subject, and content, enabling real-time access to user inbox data.

5) Summarization Module: The retrieved email content is processed using NLP techniques to generate concise summaries. This helps users quickly understand long emails without reading the entire content.

6) Text-to-Speech Module: The processed output is converted into speech using text-to-speech (TTS) technology. This allows users to listen to emails and system responses, improving accessibility.

7) Web Framework & User Interface: A web-based interface is developed to allow users to interact with the system easily. Users can give voice commands and receive responses without requiring technical knowledge.

8) Backend Processing & API Integration: The backend handles speech recognition, email retrieval, processing, and output generation. APIs are integrated to ensure efficient communication between system components.

9) Testing and Execution: The system is tested with different voice inputs and commands. It demonstrates real-time processing with minimal delay and reliable performance.

V. RESULT AND ANALYSIS

The proposed system, Voice-Enabled Smart Email Assistance for the Visually Impaired, was successfully implemented and tested using real-time voice inputs. The system demonstrated effective performance in speech recognition, email processing, and voice-based interaction.

A. Signup and Login Screens

The signup screen allows new users to create an account easily by entering basic details, while the login page provides secure access to existing users. These screens ensure authentication and personalized usage of the system. The simple and clean design improves user experience and makes the system accessible even for non-technical users.

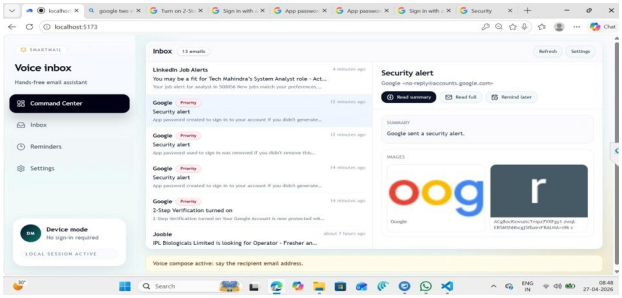


Fig. 4 Main Dashboard

B. Settings

This screen shows the mailbox configuration and settings interface for connecting the email account. It allows users to set preferences, enable voice features, and establish secure IMAP synchronization for email access.

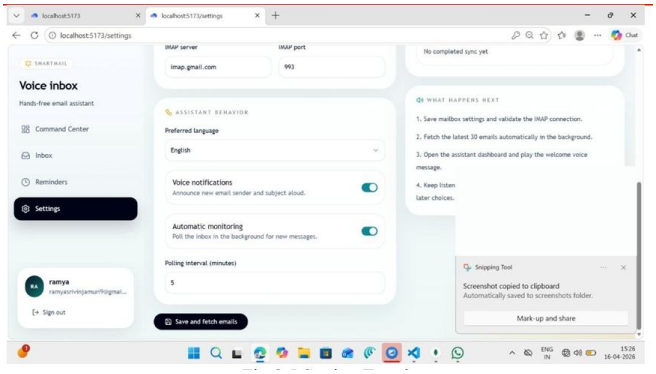


Fig. 5 Settings

C. Remainders

This screen displays the reminder and email management dashboard with saved emails and alerts. It allows users to view summaries, read emails, delay actions, and manage reminders efficiently.

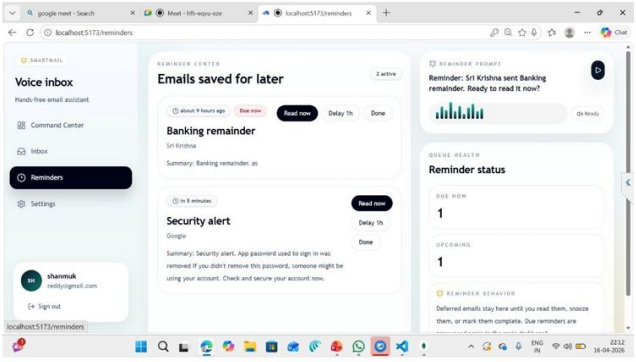
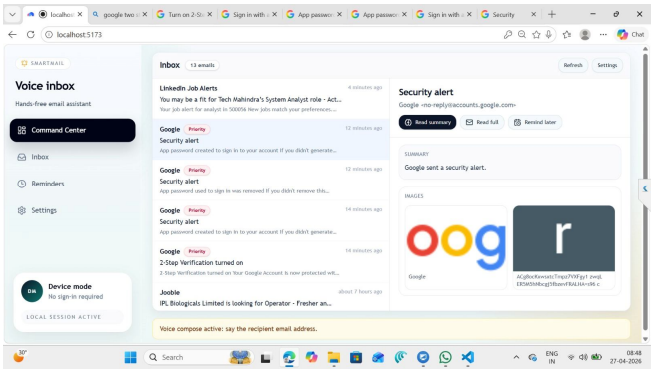


Fig. 6 Remainders

Inbox Dashboard

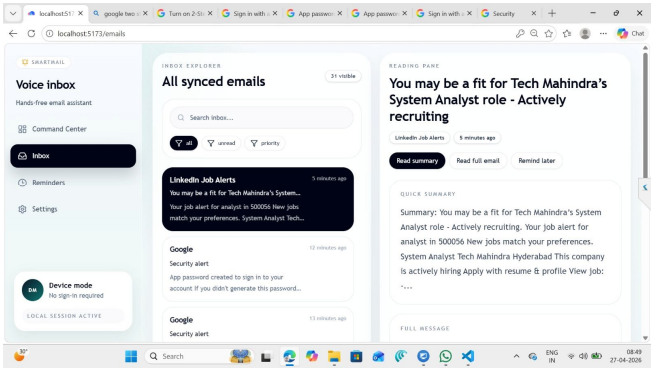
The figure shows the main interface of the Voice-Enabled Smart Email Assistance System. The inbox displays emails with sender details, subject, and priority tags. The selected email is shown on the right with options like reading summary, full content, or setting reminders.

The system also provides a voice command feature, allowing users to manage emails hands-free. The simple and user-friendly design improves accessibility and ease of use.



D. Output Screen

The figure shows the output screen of the Voice-Enabled Smart Email Assistance System. It displays the selected email along with options like reading summary, full message, and setting reminders. A quick summary of the email is provided to help users understand the content easily. The interface supports voice-based interaction, allowing users to perform actions efficiently. The clean layout ensures better readability and accessibility.



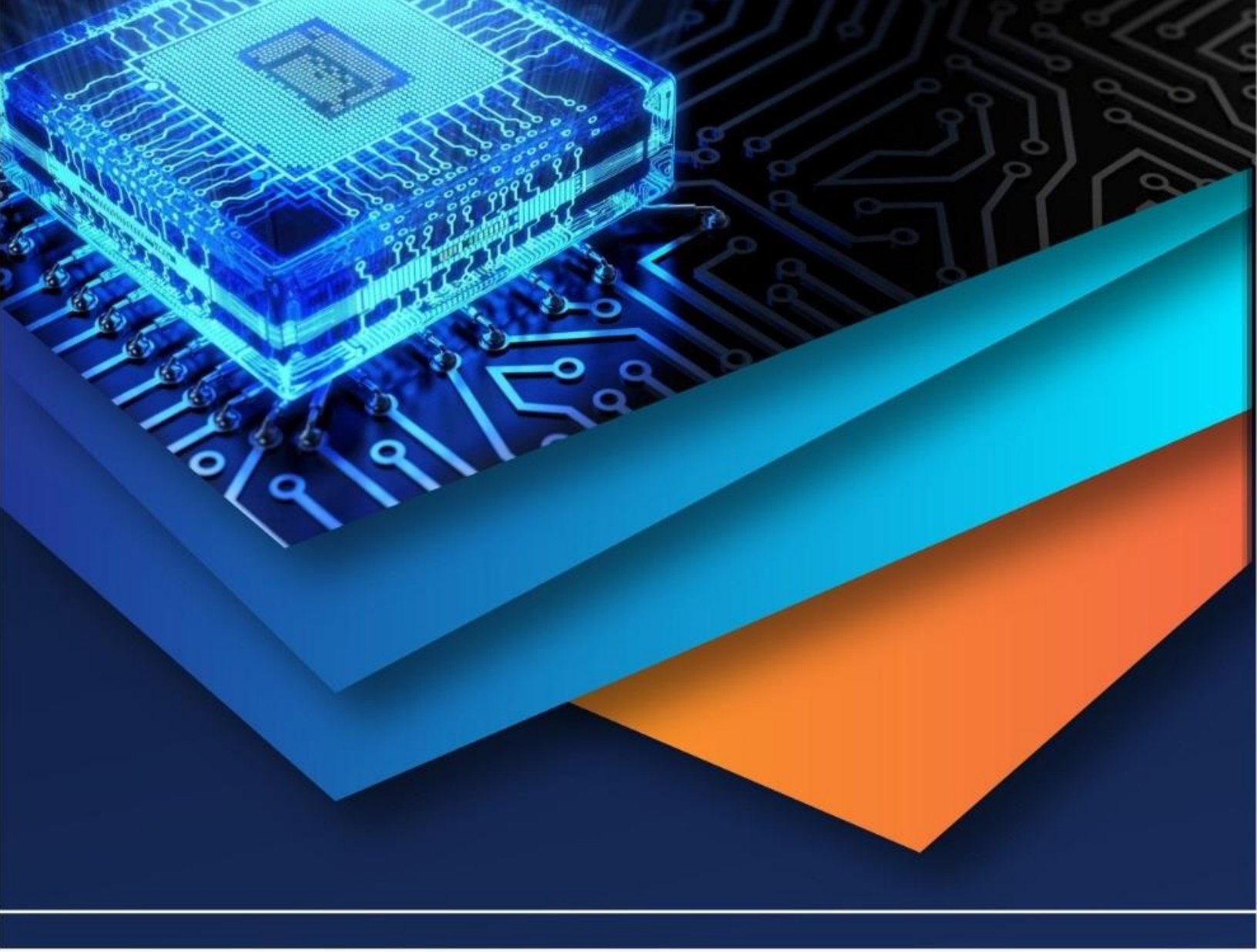
VI. CONCLUSION

The Voice-Enabled Smart Email Assistance System improves email accessibility and communication by using speech recognition and NLP technologies. It allows users to read, manage, and respond to emails through voice commands, reducing the need for manual interaction. The system provides features such as email retrieval, summarization, and voice-based responses, making it useful for visually impaired users and general users alike. It helps improve productivity and simplifies email handling.

Overall, the project demonstrates the effectiveness of voice-based systems in enhancing accessibility and efficient communication.

REFERENCES

[1] D. Jurafsky and J. H. Martin, Speech and Language Processing, 3rd ed., Pearson, 2020.
[2] J. Devlin et al., "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in NAACL-HLT, 2019.
[3] OpenAI, "Whisper: Robust Speech Recognition via Large-Scale Weak Supervision," 2022.
[4] Google, "Gmail API Documentation," Available: https://developers.google.com/gmail/api
[5] Python Software Foundation, "SpeechRecognition Library Documentation," Available: https://pypi.org/project/SpeechRecognition/
[6] Flask Documentation, "Flask Web Framework," Available: https://flask.palletsprojects.com/
[7] T. Brown et al., "Language Models are Few-Shot Learners," in Advances in Neural Information Processing Systems (NeurIPS), 2020.
[8] Microsoft, "Azure Cognitive Services – Speech-to-Text and Text-to-Speech," Available: https://azure.microsoft.com/en-in/services/cognitive-services/



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY