WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

VISUALLY IMPAIRED PEOPLE E-MAIL COMMUNICATION USING NATURAL LANGUAGE PROCESSING

[1] Ms. S. SWETHA,[2] Mrs. R. Gayathri

[1]Student, Master of Computer Applications, Vivekanandha Institute of Information and Management Studies, Tiruchengode, Tamil Nadu, India.
[2] Department of MCA, Vivekanandha Institute of Information and Management Studies, Tiruchengode, TamilNadu, India.(Assistant Professor)

ABSTRACT:

Due to their heavy reliance on visual components, digital communication technologies like email can be difficult for those with visual impairments to access and use. By providing an accessible and user-friendly email system, this project offers a way to assist people in overcoming these challenges. By converting written email content into spoken words using natural language processing techniques, the suggested solution allows visually impaired users to hear their communications rather than read them. It also incorporates speech-to-text technology, enabling users to send emails by dictating them with their voice. Additionally, while users create emails, the system can provide real-time changes and suggestions to improve the precision and readability of their communications. Machine learning enables the system to adjust in response to human interactions, thereby enhancing its suggestions for improved communication. In order to provide a comprehensive email solution for those with visual impairments, this project entails developing a prototype that combines text-to-speech, voice recognition, and natural language processing technologies. Large font, highly contrasted color schemes, and an intuitive navigation structure will all be incorporated into the system's development with accessibility as a primary goal to guarantee a flawless user experience. The ultimate goal of this project is to give visually impaired people an easy-to-use and accessible email platform, which will greatly improve their communication skills. The system makes communication more effective and autonomous by removing barriers caused by traditional visual interfaces through the use of natural language processing.

KEYWORDS: Voice E-mail, Voice System, AI, Visually Impaired AI Tool, E-mail Impaired System, Voice Recognition.

I. INTRODUCTION

Email communication is vital in today's digital age but can be challenging for visually impaired individuals due to its reliance on visual elements like images and colors. Natural Language Processing (NLP), a branch of artificial intelligence (AI), can convert text-based emails into speech, improving accessibility. The proposed solution is an inclusive email system that integrates NLP, speech recognition, and text-to-speech technologies. It will feature large fonts, high contrast, and simple navigation. Machine learning will provide real-time suggestions and improve based on user interactions, making email composition easier. Machine learning, which allows computers to learn from data without explicit programming, enhances the system's capabilities by identifying patterns and improving predictions. It has applications in fields like image recognition, fraud detection, and recommendation systems, revolutionizing industries like healthcare and finance.

OBJECTIVES

The main objective of the project titled "Visually Impaired People E-mail Communication Using Natural Language Processing" is to design and implement a voice-based email system that enables visually impaired individuals to efficiently access and manage their email communication without relying on visual interfaces. By leveraging Natural Language Processing (NLP) techniques, the system aims to convert voice commands into actionable inputs and convert email content into speech, allowing users to compose, send, read, and organize emails through natural voice interaction. The system seeks to provide an intuitive and user-friendly interface that supports multiple languages, enhancing accessibility and ease of use. Additionally, the project focuses on ensuring the security and privacy of users' data during email transmission. Ultimately, the goal is to empower visually impaired users with greater independence and improve their ability to communicate digitally by minimizing their reliance on traditional assistive technologies.

III. LITERATURE SURVEY

Recent research in the fields of speech recognition and natural language processing (NLP) has aimed to address key challenges such as multilingual code-switching, low-resource language processing, and speech recognition error prediction. Lee et al. (2022) proposed a technique for improving Chinese-English mixed speech recognition by augmenting code-switched text data using neural networks [1]. They trained a BERT-BiLSTM- CRF model on the SEAME corpus to identify code-switching points and generate synthetic sentences that preserve the linguistic characteristics of the

original dataset. Their experimental results demonstrated superior performance in mixed-language speech recognition compared to existing methods. Focusing on Indian language processing, Pandya and Kalani (2022) conducted a survey on the preprocessing phase of text sequence generation in Gujarati, an Indo-Aryan language derived from Sanskrit [2].

The paper highlights various challenges in text processing, including word sense disambiguation, part-of-speech tagging, named entity recognition, and ambiguity resolution. These preprocessing tasks are crucial for enabling effective NLP applications in low-resource Indian languages, especially for identifying syntactic elements like nouns, verbs, adjectives, and conjunctions, while also handling stop words and foreign word detection. In the area of automatic speech recognition (ASR) error simulation, Serai, Stiff, and Fosler-Lussier (2020) presented an end- to-end sequence-to-sequence model to predict ASR errors directly from clean text input [3]. Unlike traditional models relying on phonetic confusion and lexicon-based mappings, their convolutional model directly predicts errorful outputs, improving recall of recognition errors on the Switchboard dataset and demonstrating robustness across unseen domains like the Fisher corpus and a Virtual Patient ASR task. Expanding on this work, Serai, Wang, and Fosler-Lussier (2020) introduced enhancements suitable for modern neural acoustic models [4].

They proposed a sampling-based method to better simulate posterior probabilities and replaced static confusion matrices with context-aware sequence-to-sequence models. Their evaluation showed improved predictive accuracy, especially in a 100-best prediction scenario, and comparable performance to traditional confusion matrix approaches. Addressing another crucial area in speech technology, Kizito et al. (2020) developed a rule- based text normalization module for the Luganda language to support text-to-speech (TTS) systems [5]. Their system uses handcrafted grammar rules and regular expressions to detect and classify non-standard words (NSWs) based on semiotic and noun class features. Once detected, the NSWs are verbalized and converted into a normalized output. Testing on seven datasets yielded average detection and normalization accuracies of 82% and 77.7%, respectively, demonstrating the effectiveness of their approach for under-resourced language processing.

EXISTING SYSTEM

Neural end-to-end text-to-speech (TTS) systems have significantly outperformed traditional statistical methods, offering improved speech naturalness and intelligibility. However, these systems often suffer from exposure bias—a performance degradation caused by the mismatch between training (where ground-truth data is used) and inference (where predictions are based on previous outputs), particularly in autoregressive models like Tacotron2. This issue becomes more pronounced when the system is tested on out-of-domain data.

To address this challenge, a novel decoding knowledge transfer approach is proposed through a multi-teacher knowledge distillation (MT-KD) framework for the Tacotron2 TTS model. The approach involves pre-training two teacher models—one using teacher forcing and the other using scheduled sampling—and transferring their knowledge to a student model that operates in a free-running decoding mode. This strategy allows the student model to learn the advantages of both training techniques, thereby narrowing the gap between training and inference phases.

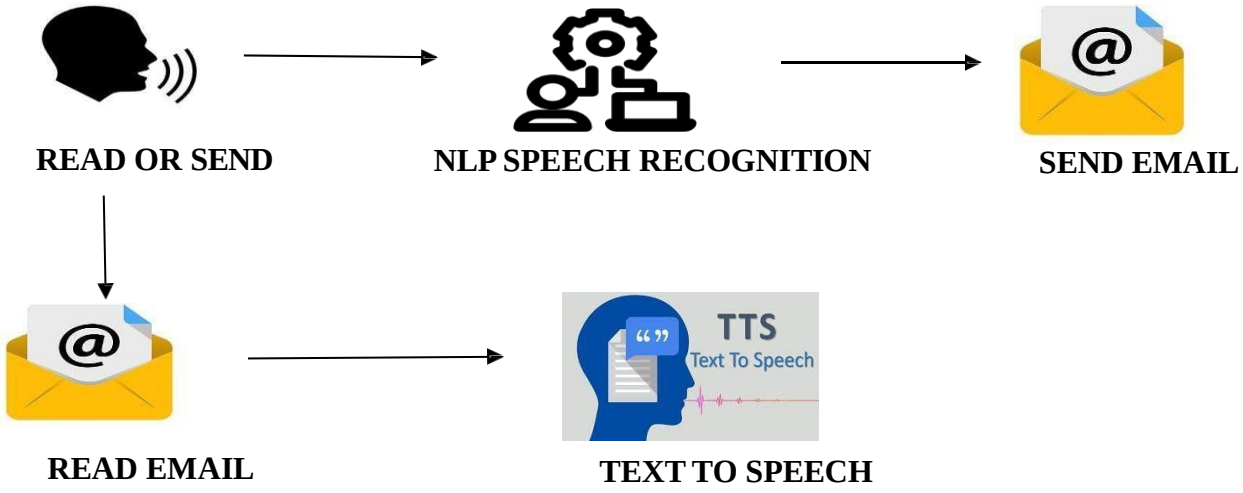
Experimental results on both Chinese and English datasets demonstrate that the MT-KD framework consistently delivers superior performance in terms of naturalness, robustness, and expressiveness across both in-domain and out-of-domain test scenarios. Furthermore, the study shows that this knowledge distillation method is more effective than adversarial learning and data augmentation in mitigating exposure bias, making it a promising solution for enhancing neural TTS systems.

PROPOSED SYSTEM

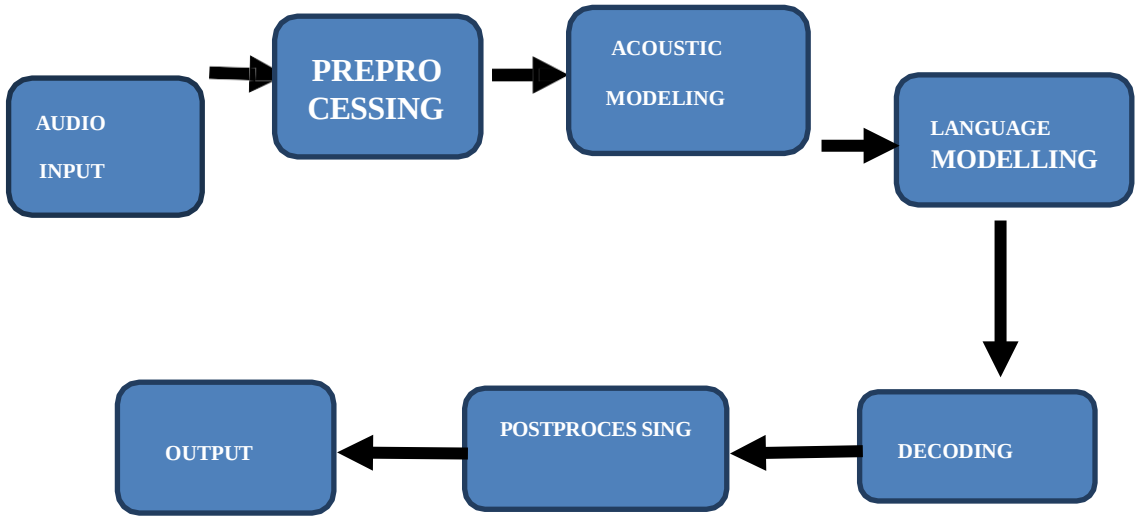
The proposed system is a voice-based email platform that enables users to read and send emails using speech recognition and text-to-speech technologies. Developed in Python, the program utilizes several libraries such as ‘smtplib’, ‘speech_recognition’, ‘pyttsx3’, ‘email’, and ‘imaplib’.The program begins by greeting the user and asking whether they would like to read their latest email or send one. If the user opts to read the latest email, the program prompts them to enter their email credentials and logs into their account using ‘imaplib’. It then retrieves the most recent email from the inbox and extracts the subject, sender, and body using the email library.

Finally, the text-to-speech technology reads out the subject, sender, and body of the email to the user. If the user chooses to send an email, the program prompts for the recipient's name, subject, and body, using speech recognition to capture the information. The email is then sent using the ‘smtplib’ library. The program leverages the pyttsx3 library for text-to-speech conversion and the ‘speech_recognition’ library for capturing speech input. The email library handles email messages, while ‘imaplib’ is used for logging into the user's email account and fetching the latest email. The ‘smtplib’ library facilitates sending emails. This project highlights how speech recognition and text-to-speech technologies can be applied to build a voice-based email system that is accessible to individuals with visual impairments or other disabilities.

PROPOSED BLOCK DIAGRAM



NLP ARCHITECTURE BLOCK DIAGRAM



METHODOLOGY

OVERVIEW OF THE PROJECT

The project titled "Voice-Based Email System for Visually Impaired Individuals Using NLP" aims to enhance email accessibility for people with visual impairments by leveraging natural language processing and speech technologies. It integrates speech recognition to allow users to dictate emails and text-to-speech (TTS) to read incoming emails aloud, enabling hands-free interaction. The system uses Python libraries like speech_recognition, pyttsx3, smtplib, email, and imaplib to perform these functions. Machine learning enhances personalization and improves user experience by adapting to individual communication patterns. Designed with accessibility in mind, the interface includes high-contrast visuals and large fonts, although its primary interaction is voice-based. The system removes the dependency on screen-based input/output, thus making digital communication more inclusive. Ultimately, it empowers visually impaired users to manage emails independently and efficiently, promoting digital equality.

MODULES

- Audio input
- Preprocessing

- Acoustic modeling
- Language modelling
- Postprocessing

8.3MODULE DESCRIPTION

AUDIO INPUT:

Captures the user's voice using a microphone and converts the analog signal to digital. Preprocessing techniques like noise reduction and filtering are applied. Segments the audio into phonemes or words for further recognition.

PREPROCESSING:

Enhances raw audio quality through noise reduction, filtering, normalization, and resampling. Removes unwanted background sounds and adjusts volume and sampling rate. Improves accuracy and reliability for downstream modules like acoustic modeling.

ACOUSTIC MODELING:

Extracts features such as MFCCs or LPC from preprocessed audio data.Maps features to phonetic units using HMMs or deep neural networks. Acts as a core component in converting speech signals to text form.

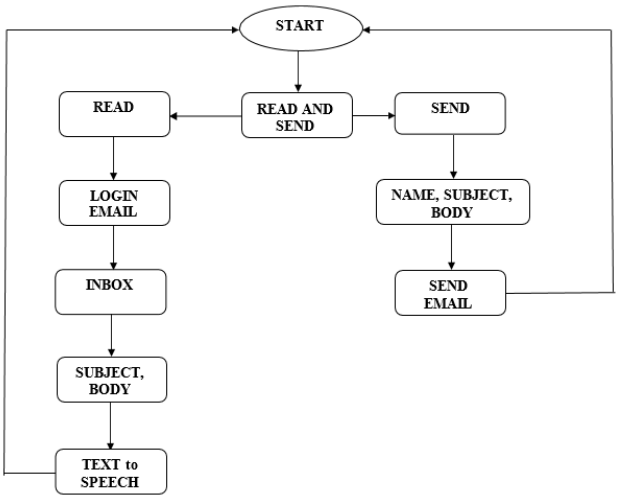
LANGUAGE MODELLING:

Refines text output by predicting the most probable word sequences. Uses statistical (n-gram) or neural models (RNNs, Transformers) for better context understanding. Improves coherence and grammar in the transcription.

POSTPROCESSING:

Adds punctuation and formatting to improve readability of the text.Corrects grammatical errors and fixes spacing or indentation issues.Uses rule-based or machine learning techniques to polish the final output.

SYSTEM ARCHTECTURE



EXPERIMENTAL RESULT

Figures shows the results:

1.Email can be read:

1. Read latest email
2. Send email
You said: read
Reading latest email...
Subject: Follow-up on Your Recovery
From: Dhana Priya <dhanapriya988@gmail.com>
Body: Dear Swetha,

I hope you are in a better condition now. Kindly let me know whether you have recovered and are able to resume work, or if further rest is required.

Wishing you a smooth recovery.

Best regards,
Priya.

2. Email can be send:

1. Read latest email
2. Send email
You said: send email
To whom do you want to send this mail?
programing is listening.....
You said: python
Speak the subject of the email
programing is listening.....
You said: project
Speak the message of the email
programing is listening.....
You said: message I hope this message finds you well I am writing to inform you regarding the project kindly ma ke time to attend a brief meeting schedule on tomorrow in the office to discuss the project current progress and ne xt steps your presence and input would be highly valuable
Your email has been send!!

CONCLUSION

In conclusion, the development of an email communication system using Natural Language Processing (NLP) for visually impaired individuals marks a significant step toward digital inclusivity and accessibility. This system allows users to interact with email services through voice commands, eliminating the need for traditional visual interfaces. By leveraging advanced Python libraries such as speech_recognition, pyttsx3, smtplib, imaplib, and email, the system enables seamless two-way communication—allowing users to compose and send emails as well as read incoming messages using speech. The integration of speech recognition for input and text-to-speech for output creates a hands-free and intuitive experience tailored to the unique needs of visually impaired users. The interface is designed to be simple, providing clear auditory prompts and feedback throughout the process, ensuring ease of use for users with little to no technical background. When the system is launched, it greets the user and offers choices such as reading the latest email or composing a new one. Upon selection, the user is guided through the required steps using natural voice interactions. For reading emails, the system logs into the user's inbox, retrieves the most recent email, and reads it aloud, including sender details, subject, and content. For sending emails, the system captures the recipient's address, subject, and message content through voice input and sends the email securely. This functionality empowers users to manage their email communication independently and efficiently.

FUTURE ENHANCEMENT

This voice-based email system can be enhanced by integrating multilingual support, allowing users to communicate in their preferred languages. Incorporating advanced natural language understanding would enable more conversational and context-aware interactions. Voice-based biometric authentication could be added toimprove security and ease of login. The system can be upgraded to include features like spam detection, email categorization, and smart reply suggestions using AI. Compatibility with mobile devices, smart assistants, and wearables would increase accessibility and convenience. Offline support and cloud synchronization could allowusers to draft and manage emails without constant internet access. Personalized voice feedback and adaptive learning based on user behavior could enhance usability. Regular usability testing with visually impaired users would ensure the system meets real-world needs. Integration with Braille displays and haptic feedback devices can also be explored for users with severe impairments. These future improvements aim to create a more intelligent, secure, and user-friendly email communication tool tailored specifically for visually impaired individuals.

REFERENCE:

1. Y. Wang et al., "Tacotron: Towards end-to-end speech synthesis", *Proc. INTERSPEECH*, pp. 4006-4010, 2017.
2. J. Shen et al., "Natural TTS synthesis by conditioning wavenet on MEL spectrogram predictions", *Proc. IEEE ICASSP Int. Conf. Acoust. Speech Signal Process.*, pp. 4779-4783, 2018.
3. K. Tokuda, Y. Nankaku, T. Toda, H. Zen, J. Yamagishi and K. Oura, "Speech synthesis based on hidden Markov models", *Proc. IEEE*, vol. 101, no. 5, pp. 1234-1252, May 2013.
4. H. Zen, A. W. Senior and M. Schuster, "Statistical parametric speech synthesis using deep neural networks", *Proc. IEEE ICASSP Int. Conf. Acoust. Speech Signal Process.*, pp. 7962-7966, 2013.
5. Z. Wu, O. Watts and S. King, "Merlin: An open source neural network speech synthesis system", *Proc. ISCA Speech Synth. Workshop*, pp. 202-207, 2016.

WEBSITE REFERRENCE:

- https://www.w3.org/WAI/
- https://pypi.org/project/SpeechRecognition/
- https://pyttsx3.readthedocs.io/
- https://www.nltk.org/