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.2025.69510

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

Understanding Emotional Tone using NLP and MLAlgorithms

Mrs.M. Anitha[1], G. Pavani[2], V. Gunasai[3], Sd.Jaibunnisa[4], A. Ananya[5]
[1]*M.Tech, Assistant Professor in Dept of CSE,* [2, 3, 4, 5]*B.Tech, Computer science & Engineering, Bapatla Women's Engineering College, Bapatla, AP, INDI*

Abstract: ***Emotional intelligence is pivotal in improving the interaction between humans and machines, especially for building AI systems that are user-friendly. This research proposes a new framework that enables Õinteractive, emotion-sensitive communication between users and digital systems. It achieves this by integrating speech emotion detection with the generation of empathetic responses. Emotions like happiness, sadness, and anger are recognized using both acoustic signals (e.g., MFCC features and Librosa) and textual inputs. A deep learning setup utilizing CNN and LSTM networks ensures precise emotion recognition. Furthermore, the system uses advanced NLP models such as BERT and GPT to either generate or retrieve motivational quotes aligned with the detected emotional state. These responses are vocalized using text-to-speech tools to support natural verbal exchanges. The final implementation is an interactive web app that delivers emotionally intelligent assistance in real time. Test results indicate that the system performs effectively in recognizing emotions and delivering contextually appropriate feedback, showcasing its utility in domains like mental wellness and AI-driven personalization.***

I. INTRODUCTION

The growing integration of technology into everyday life, it has become increasingly important for machines to recognize and appropriately respond to human emotions. Modern human-computer interaction (HCI) is evolving from mere functional exchanges to emotionally aware systems that can detect, interpret, and react in a human-like way. This study presents the development of a web-based application that can identify emotions from speech in real time and respond with relevant, voice-generated feedback. The system combines methods from speech processing, deep learning, and natural language processing (NLP) to offer more personalized and empathetic interactions. Emotion-sensitive technologies like this have promising applications in digital assistants, emotional wellness support, and intelligent educational tools.

II. LITERARURE SURVEY

Recent strides in artificial intelligence have enabled the creation of systems that can interpret human emotions and respond accordingly. Various approaches to speech emotion recognition (SER) have been studied, with a focus on extracting key audio features like Mel-Frequency Cepstral Coefficients (MFCCs) and spectral patterns using tools such as Librosa. These extracted features are often fed into advanced neural networkssuch as Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) modelsto effectively classify emotions like happiness, sadness, and anger.

III. EXISTING SYSTEM

Current systems for recognizing emotions through speech often utilize Deep Neural Networks (DNNs) to classify emotional states. These models are typically trained on datasets like the Berlin Emotional Speech Database and are limited to recognizing specific emotions such as anger, neutrality, and sadness. Audio inputs undergo preprocessing steps such as waveform normalization, segmentation, and silence removal using methods like Voice Activity Detection (VAD).

IV. PROPOSED SYSTEM

The proposed solution integrates both acoustic and textual features from spoken input to accurately identify emotions in real time. It employs sophisticated preprocessing techniques and deep learning models such as CNN and LSTM to analyse audio signals. In parallel, transformer models like BERT and GPT are used to interpret the linguistic content. A user-friendly web application serves as the interaction medium, where responses are delivered through text-to-speech technology. To further enhance user experience, the system also recommends motivational quotes tailored to the detected emotion. This makes it especially valuable for use cases like emotional wellness support, intelligent virtual assistants, and personalized digital experiences.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

V. IMPLEMENTATION

The system begins by capturing the user's voice through audio recording tools and then transcribes the spoken input into text using speech recognition techniques. It analyses both the acoustic characteristics and the textual content of the speech to accurately identify the user's emotional state. Once an emotion is recognized, the system generates contextually appropriate and emotionally aware responses using advanced natural language processing techniques and transformer-based models. These responses are then converted into speech using text-to-speech technology, allowing for a more natural and human-like interaction.

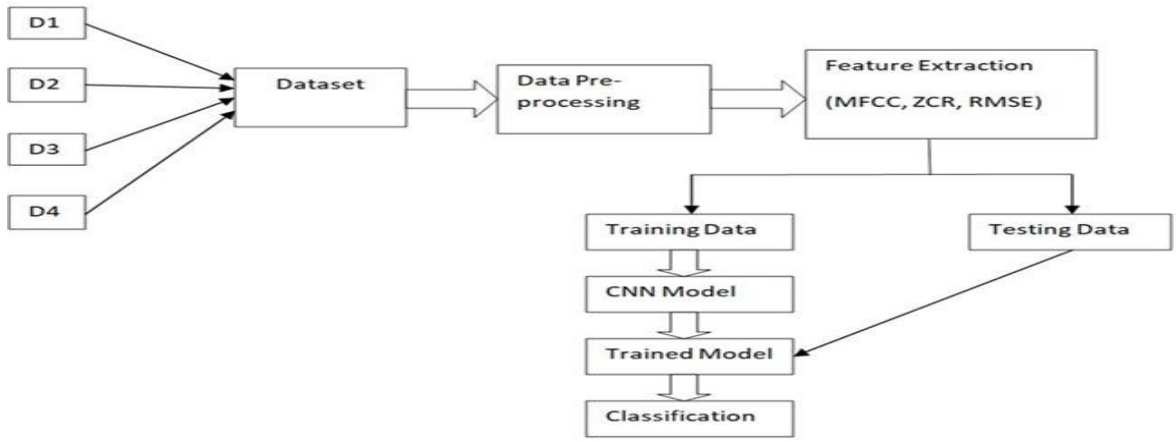
VI. MODULES

1) *Speech Input:* Captures the user's voice through audio recording and converts spoken words into text using speech recognition techniques.
2) *Emotion Recognition:* Analyses both the sound characteristics and the textual content of speech to identify the user's emotional state.
3) *Response Generation:* Creates personalized and emotionally appropriate replies using advanced natural language processing techniques and transformer-based models.
4) *Voice Response:* Uses text-to-speech technology to deliver the system's replies in a natural, human-like voice.
5) *Quote Recommendation:* Selects and provides motivational or supportive quotes that align with the user's identified emotion to enhance engagement and emotional support.

VII. ALGORITHMS

1) *Voice input:* The system utilizes audio tools such as PyAudio and the speech_recognition library to record and interpret spoken input. Speech is then converted into text using APIs like Google Speech-to-Text or OpenAI's Whisper, enabling further processing.
2) *Audio Feature Extraction:* Key sound characteristics—such as Mel-Frequency Cepstral Coefficients (MFCCs), chroma features, and spectrogramsare extracted using audio analysis libraries like Librosa. These features help in detecting emotional nuances in speech.
3) *Emotion Classification:* To determine the speaker's emotional state, the system combines deep learning models like LSTM and CNN with traditional machine learning algorithms such as Support Vector Machines (SVM) and Random Forest.
4) *Response Generation (Text):* Emotionally relevant responses are created using state-of-the-art transformer models, including GPT-2, GPT-3, T5, and BART. These are further enhanced with rule-based NLP approaches and fine-tuned using pre-trained models from Hugging Face to ensure accurate and meaningful replies.
5) *Motivational Quote Module:* The system intelligently selects motivational or supportive quotes from a database annotated with emotional categories. In some cases, transformer models may refine or generate quotes based on the user's emotional state for improved resonance.
6) *Voice Output:* To deliver responses naturally, the text output is converted into speech using text-to-speech engines such as pyttsx3, Google TTS, Microsoft Edge TTS, or Google Cloud TTS.

VIII. SYSTEM ARCHITECTURE



©IJRASET: All Rights are Reserved

IX. RESULT AND ANALYSIS

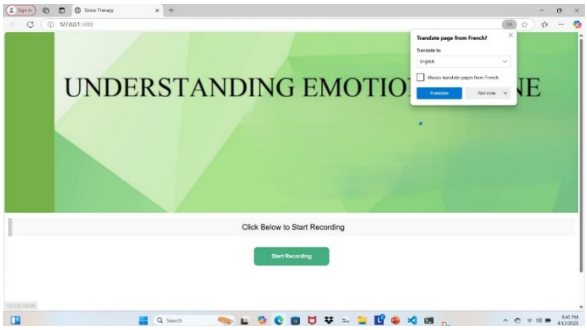


Fig1: Select a Language

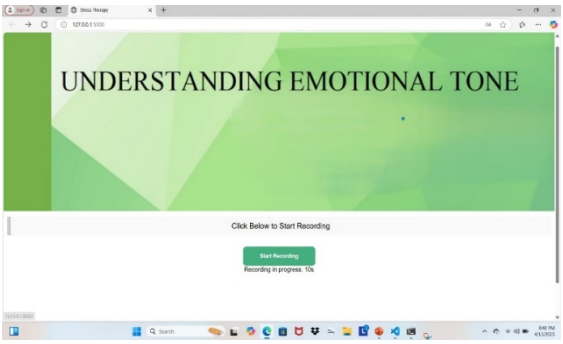


Fig2: Start Recording

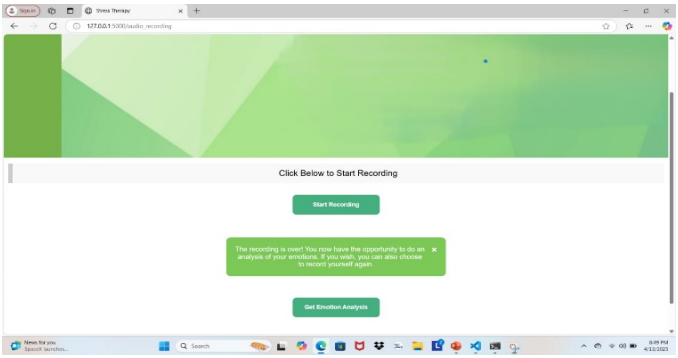


Fig3: Get Analysis

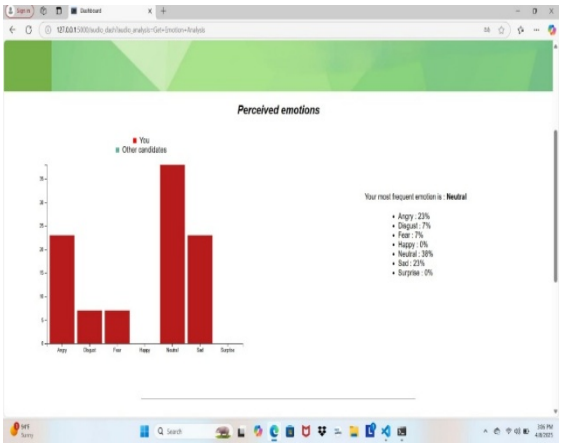


Fig4: NeutralDominant

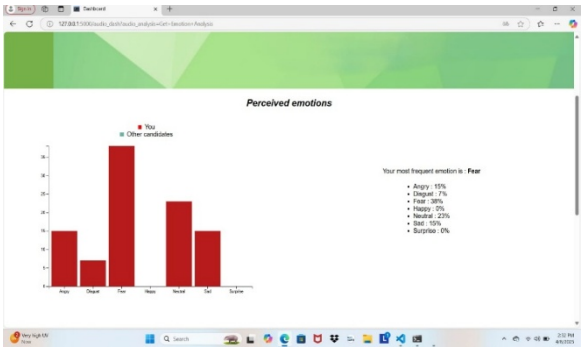


Fig 5: FearDominant

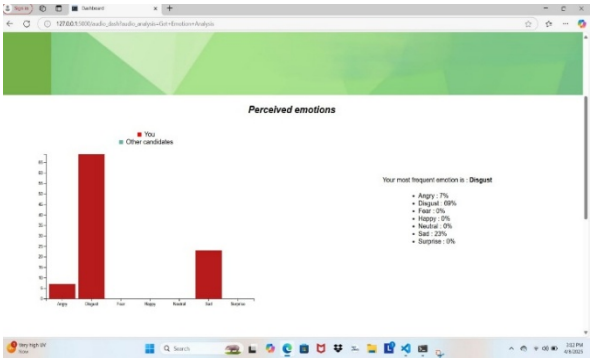


Fig 6: DisgustDominant

X. CONCLUSION

This emotion-aware framework facilitates more natural and empathetic communication by leveraging natural language processing and machine learning to recognize and interpret emotional cues. Through real-time, voice-driven interaction and the delivery of motivational responses, the system significantly improves user engagement. This initiative marks a notable advancement in building emotionally intelligent AI systems with broad potential across sectors like mental wellness support and virtual assistance.

XI. FUTURE SCOPE

Future developments may include multilingual support and adaptation to various regional accents, making the system more inclusive. Enhancing its emotional sensitivity could lead to more nuanced and context-aware responses. Additionally, incorporating offline capabilities would allow the system to function without internet connectivity, thus improving user privacy and extending accessibility to remote or low-connectivity regions.

REFERENCE PAPERS

[1] Ahmed, M. R., Islam, S., & Islam, A. K. M. M. (n.d.). Improved Speech Emotion Recognition Using an Ensemble of CNN-LSTM-GRU Models with Data Augmentation.
[2] Hamsa, S., Shahin, I., & Iraqi, Y. (2020). Speech Emotion Classification Using Wavelet Packet Transform and Random Forest. IEEE Access, 8, 96994–97006.
[3] Murugan, H. (n.d.). Exploring Speech Emotion Recognition Using Convolutional Neural Networks (CNNs). SRM Institute of Science and Technology.

Author's Profiles



Mrs.M. Anithaworking as Assistant Professor in Department of CSE, BWEC, Bapatla.



G. PavaniB.Tech with specialization of Artificial Intelligence and Machine Learning in Bapatla Women's Engineering College, Bapatla.



V. GunasaiB.Tech with specialiazation of Artificial Intelligence and Machine Learning in Bapatla Women's Engineering College, Bapatla.

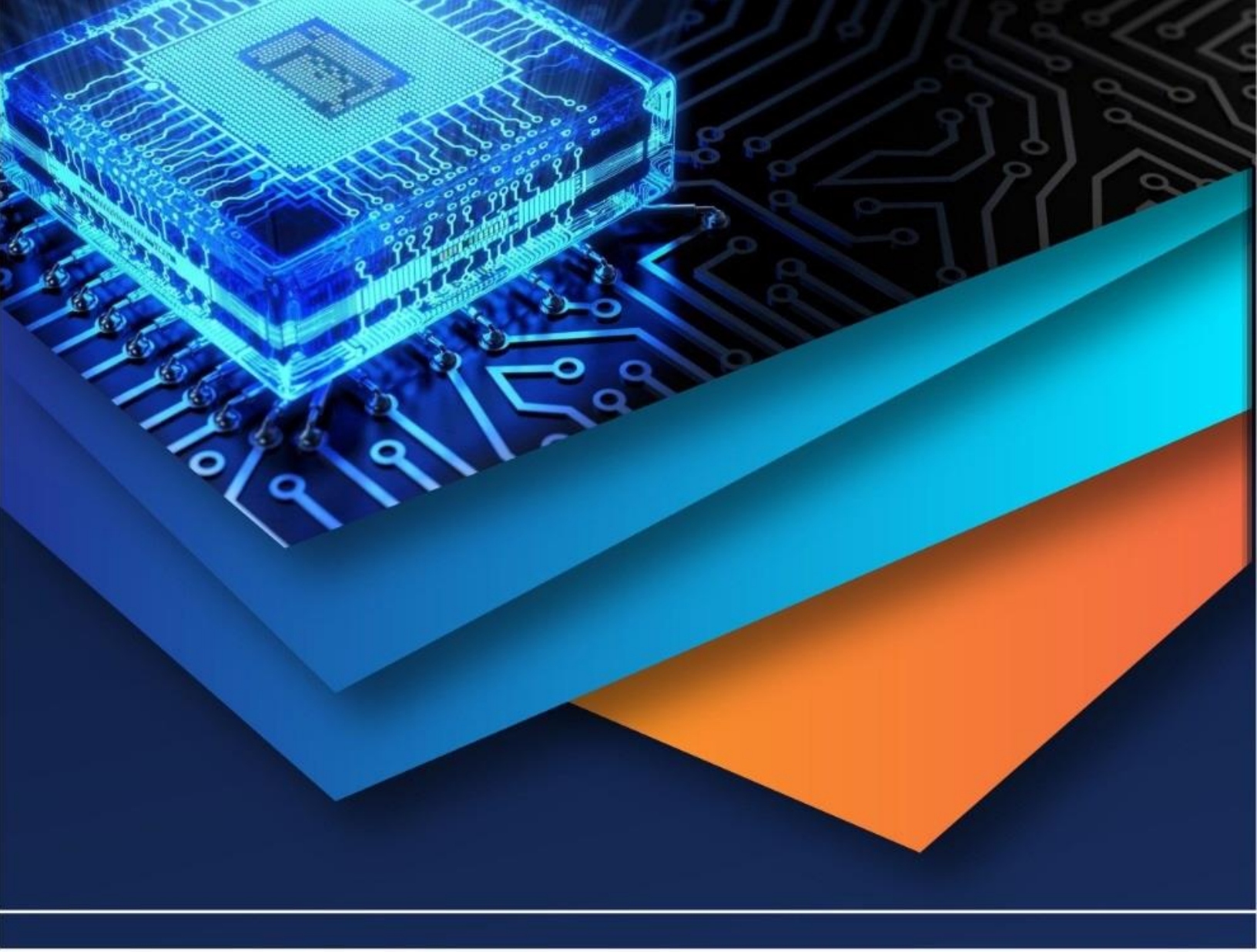


Sd.Jaibunnisa B.Tech with specialization of Artificial Intelligence and Machine Learning in BapatlaWomen's Engineering College,Bapatla.



A. AnanyaB.Tech with specialization of Artificial Intelligence and Machine Learning in Bapatla Women's Engineering College, Bapatla.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)