Voice-Based Emotion Recognition Using Machine Learning Classifiers

Trupti Kalokhe and Mrs.Nanda Kulkarni
Department of Computer Engineering,
Siddhant College of Engineering, Sudumbre, Pune, India
trupti.bhase18@gmail.com

Abstract: *Emotion recognition from speech has emerged as a pivotal research area in artificial intelligence, enabling more empathetic and context-aware human-computer interactions. This paper presents the implementation of a machine learning-based system for real-time emotion recognition using voice data. The proposed system utilizes Mel Frequency Cepstral Coefficients (MFCC) and Linear Predictive Coding (LPC) as primary feature extraction techniques to capture critical emotional cues embedded in speech, such as pitch, tone, and spectral characteristics. These features are fed into various supervised learning models, including Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Decision Tree (DT), Random Forest (RF), and Gradient Boosting (GB). The models were trained and evaluated using benchmark emotional speech datasets, ensuring robustness through preprocessing, data normalization, and cross-validation. Performance metrics such as accuracy, precision, recall, and F1-score were computed to assess each model's efficacy. Experimental results demonstrate that ensemble-based models, particularly Random Forest and Gradient Boosting, outperform others, with Random Forest achieving the highest accuracy of 95.5%. Even with overlapping speech patterns and mild background noise, the system can distinguish emotions. This implementation showcases the feasibility of deploying lightweight, real-time speech emotion recognition systems in practical scenarios such as customer service, mental health monitoring, and intelligent virtual assistants. The study highlights the significance of choosing suitable feature extraction techniques and classifiers while addressing challenges such as emotional overlap and generalizability across diverse datasets. The proposed framework is a foundation for future work involving deep learning, multimodal integration, and deployment on edge devices..*

Keywords: Speech Emotion Recognition, MFCC, LPC, Machine Learning, Random Forest, Gradient Boosting, Voice Data, Real-Time Emotion Detection, Human-Computer Interaction, Feature Extraction