Comparison Study of Machine Learning Classifiers for Speech Emotion Recognition

Authors : Aishwarya Ravindra Fursule, Shruti Kshirsagar

Abstract : In the intersection of artificial intelligence and human-centered computing, this paper delves into speech emotion recognition (SER). It presents a comparative analysis of machine learning models such as K-Nearest Neighbors (KNN),logistic regression, support vector machines (SVM), decision trees, ensemble classifiers, and random forests, applied to SER. The research employs four datasets: Crema D, SAVEE, TESS, and RAVDESS. It focuses on extracting salient audio signal features like Zero Crossing Rate (ZCR), Chroma_stft, Mel Frequency Cepstral Coefficients (MFCC), root mean square (RMS) value, and MelSpectogram. These features are used to train and evaluate the models' ability to recognize eight types of emotions from speech: happy, sad, neutral, angry, calm, disgust, fear, and surprise. Among the models, the Random Forest algorithm demonstrated superior performance, achieving approximately 79% accuracy. This suggests its suitability for SER within the parameters of this study. The research contributes to SER by showcasing the effectiveness of various machine learning algorithms and feature extraction techniques. The findings hold promise for the development of more precise emotion recognition systems in the future. This abstract provides a succinct overview of the paper's content, methods, and results.

Keywords : comparison, ML classifiers, KNN, decision tree, SVM, random forest, logistic regression, ensemble classifiers

Conference Title : ICMLDA 2024 : International Conference on Machine Learning and Data Analysis

Conference Location : New York, United States

Conference Dates : June 03-04, 2024