

Speech Emotion Recognition

Introduction:

Speech Emotion Recognition (SER) is like teaching a computer to understand how people feel when they talk. In this project, we already have a bunch of recordings of people talking with their different emotions. The plan is to make a smart system that can listen to these recordings and figure out if the person is happy, sad, angry, or feeling other emotions.

Problem Specification:

- Leveraging a dataset from a source like RAVDESS, we extract acoustic and prosodic features for nuanced emotion analysis.
- Identifying relevant features from the dataset such as pitch, loudness, etc
- Exploring and evaluating various machine learning algorithms to determine the most suitable model for the task.

Features and Classification Models:

Speech Emotion Recognition employs acoustic features like pitch, energy, and formants, alongside prosodic elements such as intonation and stress. Deep acoustic features, including MFCCs and spectrograms, are also commonly used. Classification models range from traditional approaches like **Support Vector Machines and Decision Trees** to more advanced methods such as **Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and Convolutional Neural Networks (CNNs)**. Hybrid models, combining acoustic, prosodic, and linguistic features, can enhance performance. The selection of features and models depends on the specific requirements of the SER task and the nature of the dataset.

Expected Results:

The anticipated results in Speech Emotion Recognition involve achieving high accuracy, generalization to unseen data, real-time performance, and effective handling of variability in speech patterns and cultural nuances. Additionally, comparison with baseline models, consideration of ethical implications, and adaptability to evolving linguistic contexts are crucial aspects for evaluating the success of the SER system. The ultimate goal is to create a robust and ethically sound model that accurately captures and interprets emotional states in spoken language across diverse scenarios.

Team Information and Planned Contributions:

- **V Jahnavika Reddy- S20210020330**
 - **Contribution:** Data analysis, Feature extraction, and Classification using SVM.
- **K Sai Likhitha- S20210020289**
 - **Contribution:** Data analysis and classification using Decision trees and RNNs.
- **P Jasmitha Sai- S20210020314**
 - **Contribution:** Data analysis and classification using LSTM networks and CNNs.