

# **Prototype Document for Emotion Detection using Voice**

## **Approach**

The prototype uses audio-based features and machine learning techniques to detect aggression in audio data, involving data extraction, feature engineering, model training, and prediction.

## **Features**

**Audio Features:** MFCC (Mel-Frequency Cepstral Coefficients): These capture the spectral characteristics of audio.

**Chroma Features:** Represent the 12 different pitch classes.

**Mel Features:** Reflect the spectral characteristics of sound.

## **Libraries**

**Librosa:** Used for audio feature extraction.

**Soundfile:** Used for reading sound files.

**Numpy:** Used for numerical operations.

**Scikit-learn:** Used for machine learning.

**Pandas:** Used for data organization and visualization.

## **User Manual**

**Input:**

**Audio Files:** Users can provide their own audio files for aggression detection. The accepted audio file formats include:

**Analysis: Feature Extraction:** The prototype extracts MFCC, chroma, and mel features from the audio data.

**Training:** A Multi-Layer Perceptron (MLP) classifier is trained on the extracted features to classify emotions, with a focus on detecting aggression.

**Output:**

The prototype reports an accuracy score, indicating how well the model performs in emotion detection.

F1 Scores: F1 scores for individual emotions (e.g., 'calm,' 'disgust,' 'happy,' 'fearful') are provided.

Emotion Predictions: Users can predict emotions for their audio data using the trained model. A sample prediction ('disgust') is demonstrated in the code.

Model Serialization: The trained model is saved to a file for future use.

## **Conclusion:**

The prototype provides a solution for detecting aggression in audio data by utilizing audio features and machine learning to classify emotions, specifically focusing on detecting aggression, based on user input.