# Title: Speech Emotion Recognition with Librosa

## Introduction

In the digital age, understanding human emotions through speech can enhance numerous applications, from customer service to mental health monitoring. The **Speech Emotion Recognition with Librosa** project aims to leverage audio processing techniques to detect and classify emotions from speech using the powerful Python libraries Librosa, Soundfile, and sklearn. This project will serve as an introduction to audio processing, feature extraction, and machine learning, paving the way for more advanced applications such as deep learning-based emotion recognition systems.

## Objectives

- To load and process sound files using the Librosa and Soundfile libraries.

- To perform feature extraction from audio data.

- To train a Multi-Layer Perceptron (MLP) classifier model using the sklearn library to recognize emotions in speech.

- To gain foundational knowledge in audio processing and machine learning, enabling future advancements in the field of speech emotion recognition.

## Methodology

1. **Data Collection and Preprocessing**

   o   Collect a dataset of speech recordings labeled with corresponding emotions.

   o   Load and preprocess these sound files using the Librosa and Soundfile libraries to ensure they are in a suitable format for feature extraction.

2. **Feature Extraction**

   o   Extract relevant audio features such as Mel-frequency cepstral coefficients (MFCCs), chroma, and spectral contrast using Librosa.

   o   Compile these features into a structured format suitable for input into the MLP classifier.

3. **Model Training**

   o   Utilize the sklearn library to define and train a Multi-Layer Perceptron (MLP) classifier on the extracted features.

   o   Implement cross-validation and parameter tuning to optimize the model's performance.

4. **Evaluation and Analysis**

   o Evaluate the trained MLP classifier using appropriate metrics such as accuracy, precision, recall, and F1-score.

   o Analyze the results to understand the model's strengths and limitations in recognizing speech emotions.

5. **Future Work**

   o Discuss potential improvements and next steps, including the possibility of employing deep learning models such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) to achieve better accuracy.

## Tools and Technologies

- **Librosa**: For audio processing and feature extraction.

- **Soundfile**: For reading and writing sound files.

- **sklearn**: For machine learning model development and evaluation.

- **Python**: As the programming language to integrate these libraries and implement the project.

## Expected Outcomes

By the end of this project, you will have:

- A functional MLP classifier capable of recognizing emotions from speech recordings.

- A solid understanding of audio processing and feature extraction techniques.

- Hands-on experience with machine learning models for classification tasks.

- A foundation for advancing to more complex deep learning models for enhanced emotion recognition accuracy.