

Human Emotion Detection from Voice

Abstract

The ability to detect human emotions from speech is crucial in many applications, including mental health analysis, customer service, and human-computer interaction. This project aims to develop a system that can identify emotions from voice recordings using **the RAVDESS dataset**. By extracting audio features and training machine learning models, the system predicts emotions such as **Happy, Sad, Angry, Neutral, Fearful, Disgust, Calm, and Surprised**. The solution also provides a **Streamlit-based web interface** for real-time emotion detection.

Introduction

Emotions play a significant role in human communication. Recognizing emotions through voice can enhance the effectiveness of automated systems like **virtual assistants, call centers, and healthcare applications**. Traditional text-based emotion recognition often lacks accuracy as emotions are better expressed through tone and pitch.

This project leverages **audio processing techniques (MFCC, Chroma, Spectral features)** and **machine learning models** to predict emotions from speech.

Tools & Technologies Used

- **Programming Language:** Python 3.9+
- **Libraries:**
 - **Librosa:** For audio feature extraction
 - **Scikit-learn:** For machine learning model development
 - **Streamlit:** For creating a web-based interface
 - **NumPy & Pandas:** For data processing
 - **Soundfile:** For audio file handling
- **Dataset:** RAVDESS (Ryerson Audio-Visual Database of Emotional Speech and Song)
- **Version Control:** Git & GitHub

Steps Involved in Building the Project

1. Data Collection & Preparation:

- Downloaded the **RAVDESS dataset** from [Zenodo](#).
- Organized the dataset into the data/ directory for processing.

2. Feature Extraction:

- Extracted **MFCC (Mel-frequency cepstral coefficients)** for capturing spectral properties of speech.

- Added **Chroma, Spectral Contrast, Tonnetz**, and **RMS Energy** features for better accuracy.
- 3. Data Preprocessing:**
- Normalized extracted features.
 - Encoded emotion labels using **LabelEncoder**.
 - Split dataset into **training (80%)** and **testing (20%)**.
- 4. Model Training:**
- Trained a **Random Forest Classifier** using extracted features.
 - Tuned hyperparameters for optimal performance.
- 5. Model Evaluation:**
- Achieved high accuracy in predicting 8 emotion classes.
 - Evaluated using metrics like **accuracy score** and **confusion matrix**.
- 6. Web Application Development:**
- Created a **Streamlit web app** (app.py) for real-time predictions.
 - Users can upload audio files in .wav format and get:
 - ✓ Predicted emotion
 - ✓ Intensity level
 - ✓ Confidence score

Conclusion

The **Human Emotion Detection from Voice** project successfully demonstrates the ability to classify emotions from audio recordings using machine learning. The system provides a **user-friendly web interface**, real-time predictions, and robust performance using the **RAVDESS dataset**. This approach can be extended to real-world applications like **mental health monitoring, virtual assistants, and customer service automation**.