Group No.: 06
Members:
Md. Ahnaf Tajwar Khan (ID: 1931447042)
Md. Shadman Sakib (ID: 1911357042)
Zahid Al Razin (ID: 1921046042)
T M Redwan Shahriar (ID: 2021679642)

Project Update: Emotions on Audio Dataset

1. Introduction

The primary goal of this project is to analyze and classify human emotions from audio recordings using machine learning techniques. We are working with the 'Emotions on Audio Dataset' from Kaggle, which contains labeled speech data categorized into different emotions. The objective is to develop a model capable of identifying emotional states from voice inputs accurately.

2. Dataset Overview

The dataset consists of over 5000 labeled audio samples categorized into seven emotions: Angry, Disgust, Fear, Happy, Neutral, Sad, and Surprise. Features such as Mel-Frequency Cepstral Coefficients (MFCCs) and spectrograms are extracted from these recordings to train the model.

3. Methodology

Our approach consists of several steps:

- Data Preprocessing: Includes noise reduction and normalization.
- Feature Extraction: Using MFCCs and spectrograms.
- Model Selection: Training models such as CNN, RNN, and SVM.
- Performance Evaluation: Using metrics like accuracy and F1-score.

4. Progress So Far

- Successfully preprocessed the dataset and extracted necessary features.
- Implemented initial training of machine learning models (CNN & SVM).
- Currently fine-tuning models and evaluating performance.

5. Challenges Faced

- Imbalanced data across different emotion classes.
- Background noise affecting classification accuracy.
- Computational limitations in training deep learning models.
- Hyperparameter tuning for improving accuracy.

6. Future Plans

- Improve model performance through better tuning and feature selection.
- Implement data augmentation techniques to handle imbalance.
- Compare multiple machine learning classifiers to identify the best model.
- Deploy a real-time emotion recognition system.

7. References

- Kaggle Dataset: https://www.kaggle.com/datasets/tapakah68/emotions-on-audio-dataset
- Librosa Documentation: https://librosa.org/doc/main/index.html
- Scikit-Learn: https://scikit-learn.org/stable/
- Deep Learning for Audio Processing: https://arxiv.org/abs/1811.06621