Music Genre Classification and Recommendation System

-Adtiya Arun Nambiar

Abstract

This project demonstrates a music genre classification and recommendation system using deep learning techniques. By utilizing Mel Frequency Cepstral Coefficients (MFCC) for feature extraction and a Convolutional Neural Network (CNN) for classification, the system accurately identifies the genre of a given audio track. The model is trained on the GTZAN dataset, which consists of 10 distinct music genres. Additionally, a content-based music recommendation system is implemented using cosine similarity to suggest similar tracks. The project also includes spectrogram visualization for audio insights.

Introduction

Automatic music genre classification is an essential task in audio signal processing with applications in music streaming, playlist generation, and digital media organization. Manually tagging thousands of tracks is impractical, making automation crucial. This project explores how machine learning and signal processing can be integrated to build an end-to-end pipeline that not only classifies songs based on their content but also recommends similar ones. It bridges deep learning with traditional similarity-based filtering, creating a practical music analysis tool.

Tools Used

- **Python** – Core programming language
- **Librosa** – Audio analysis and MFCC feature extraction
- **TensorFlow/Keras** – Designing and training the CNN model
- **Scikit-learn** – Preprocessing, label encoding, and evaluation
- **Matplotlib** – Plotting training metrics and spectrograms
- **NumPy & Pandas** – Data manipulation and storage

Steps Involved in Building the Project

1. **Data Loading & Preprocessing**: Audio files were loaded from the GTZAN dataset. For each file, MFCC features were extracted to represent important audio characteristics.

2. **Label Encoding**: Music genres were encoded using LabelEncoder and converted into categorical format for model compatibility.

3. **Model Design**: A CNN with multiple convolution, pooling, dropout, and dense layers was designed to learn temporal patterns in the extracted features.

4. **Training**: The model was trained for 50 epochs with training and validation splits. Accuracy and loss were monitored to evaluate performance.

5. **Evaluation & Visualization**: The final model showed high validation accuracy. Accuracy curves were plotted for analysis.

6. **Prediction & Recommendation**: A genre prediction function returns the predicted genre with confidence. A recommendation engine uses cosine similarity to suggest top-k most similar tracks.

7. **Spectrogram Display**: An additional utility allows users to visualize the frequency spectrum of any given track.

Conclusion

This project successfully builds a hybrid music analysis system combining deep learning classification with content-based recommendation. The model generalizes well on unseen data and provides explainable visualizations through spectrograms. Such systems have real-world applications in music streaming platforms, DJ tools, and audio libraries. Future work can involve larger datasets, real-time prediction interfaces, or integrating collaborative filtering techniques.