

PROJECT REPORT-MUSIC GENRE CLASSIFICATION

Internship Project: HR Elevate Labs

Author: Obilisetty Vyshnavi

◆ Introduction

With the explosive growth in digital music, automatic genre classification has become crucial for music recommendation systems, streaming platforms, and media organization. This project focuses on predicting the **genre of a music file (MP3/WAV/MP4)** using **deep learning** techniques.

◆ Abstract

This project implements a **Convolutional Neural Network (CNN)** to classify music genres based on audio signal features. The model is trained using raw .wav/.mp3 audio files rather than preprocessed CSV features to improve real-world performance. Additionally, the app allows users to upload audio or record live from a microphone for real-time prediction. The system visualizes spectrograms and shows confidence scores across genres for transparency.

◆ Tools Used

- **Python**
- **Librosa** (audio processing)
- **TensorFlow/Keras** (model training)
- **Matplotlib/Seaborn** (plotting)
- **Streamlit** (web app UI)
- **Gradio or Pyngrok** (optional deployment)
- **Google Colab** (training & deployment)

◆ Steps Involved in Building the Project

1. Dataset Collection:

- GTZAN dataset and a custom dataset of Indian + Western music in .wav and .mp3 formats were used.

2. Preprocessing:

- Audio trimmed to 30s, converted to mono.
- MFCC and Mel Spectrogram features were extracted.

3. Model Development:

- CNN architecture with Conv2D, MaxPooling, and Dense layers.
- Trained on 80/20 split, evaluated with confusion matrix, precision, recall, F1-score.

4. App Development:

- Created Streamlit interface with file uploader & mic input.
- Displayed waveform, spectrogram, predicted genre, and confidence bar chart.

5. Deployment:

- Deployed via ngrok from Colab for live demo.
- Generated performance dashboard and PDF report.

◆ Conclusion

The Music Genre Classifier achieved over **80% accuracy** across 10 genres and supported real-time predictions using uploaded or recorded audio. The project showcases a complete ML workflow — from raw audio ingestion to model training, UI integration, and deployment. This system can be further improved with larger, balanced datasets and multilingual audio data, making it ideal for recommendation engines and digital music apps.