Introduction

Music genre classification is the process of automatically sorting music into specific genres like rock, jazz, classical, pop, and more. With the fast growth of digital music platforms, automatic genre classification is important for music recommendation systems, playlist creation, and music organization. This project aims to create a machine learning model that can predict the genre of a music file by using audio feature extraction techniques and a neural network classifier.

Abstract

The aim of this project is to design a Music Genre Classification system based on Deep Learning algorithms. The GTZAN dataset with 10 different genres of music was employed for training and testing purposes.

Audio files were pre-processed using the Librosa library to obtain Mel-Frequency Cepstral Coefficients (MFCC), which are significant frequency features of audio signals. A fully connected Artificial Neural Network (ANN) was developed using TensorFlow and Keras to classify the features into 10 genres.

The network was trained on an 80-20 split and performed well on the test dataset. The system is able to correctly identify music genres from unknown audio files.

Tools Used

- Python
- Google Colab
- Librosa (Audio Processing)
- NumPy
- Scikit-learn
- TensorFlow / Keras
- Matplotlib

Steps Involved in Building the Project

Step 1: Dataset Collection

The GTZAN Music Genre dataset was downloaded from Kaggle. It contains 1000 audio files categorized into 10 genres (blues, classical, country, disco, hiphop, jazz, metal, pop, reggae, rock).

Step 2: Data Preprocessing

Audio files were loaded using Librosa at a sampling rate of 22050 Hz. MFCC features were extracted from each audio file to represent the frequency patterns.

Step 3: Feature Extraction

40 MFCC coefficients were extracted from each audio file. The mean of the MFCC values was computed to create a fixed-length feature vector for each sample.

Step 4: Label Encoding

Genre labels were converted into numerical format using LabelEncoder and transformed into one-hot encoded vectors.

Step 5: Train-Test Split

The dataset was split into 80% training data and 20% testing data to evaluate model performance.

Step 6: Model Building

A Deep Neural Network was built using Keras with:

- Dense layers (256, 128, 64 units)
- ReLU activation function
- Dropout layers for regularization
- Softmax activation in output layer

Step 7: Model Training

The model was trained for 50 epochs using categorical cross-entropy loss and Adam optimizer.

Step 8: Model Evaluation

The model performance was evaluated using:

- Accuracy score
- Confusion matrix
- Classification report

The trained model was saved in .h5 format for future use.

Conclusion

The Music Genre Classification system successfully classified audio files into 10 different genres using MFCC feature extraction and a neural network classifier. The model achieved good accuracy on unseen test data and demonstrated the effectiveness of deep learning in audio signal processing tasks.

Future improvements may include:

- Using CNN models for higher accuracy
- Using spectrogram images
- Increasing dataset size
- Hyperparameter tuning