Project Report: Music Genre Classification using Machine Learning

1. Introduction

Music genre classification stands as a significant endeavor in the landscape of music analysis and information retrieval. It encompasses the process of categorizing audio tracks into distinct genres such as jazz, rock, pop, country, and more. The automation of this classification process through the utilization of deep learning techniques holds immense potential, enabling applications such as music recommendation systems, playlist generation, and genre-based content organization. This project aims to leverage deep learning methodologies to achieve accurate and efficient music genre classification.

2. Dataset Description

The cornerstone of this project is the dataset utilized for training and evaluation. This dataset comprises audio recordings from diverse music genres, with each audio clip represented by a comprehensive set of extracted features. These features serve as the basis for training machine learning models for genre classification. The dataset includes:

- Features : Extracted using the powerful Librosa library, the features encapsulate a wide array of audio characteristics, including spectrograms, spectral centroids, chroma features, and zero-crossing rates. These features provide rich insights into the underlying structure and nuances of each audio clip.
- Genres : The dataset encompasses a diverse collection of music genres, with each genre represented as a label for classification purposes. Common genres include country, jazz, rock, pop, electronic, etc.
- Size : With a substantial volume of data, the dataset comprises 9990 audio samples, each characterized by a comprehensive set of 60 features.

3. Steps for Machine Learning Project

3.1 Data Preprocessing

Data preprocessing serves as a crucial initial step in preparing the dataset for model training. This involves a series of essential tasks, including:

- Loading Audio Files : Leveraging the robust capabilities of the Librosa library to load audio files in numerical format, facilitating subsequent feature extraction and analysis.
- Feature Extraction : Employing Librosa to extract pertinent audio features from each audio clip. This process involves the extraction of various acoustic features such as spectrograms, chroma features, spectral centroids, and zero-crossing rates, which collectively capture the essential characteristics of the audio.

- Categorical Encoding : Encoding categorical variables, such as music genres, into numerical form using LabelEncoder from the scikit-learn library. This ensures compatibility with machine learning models and facilitates seamless classification.

3.2 Model Selection and Evaluation Metrics

For the task of music genre classification, a deep learning approach is adopted, leveraging the powerful TensorFlow and Keras libraries. The chosen model architecture comprises multiple dense layers with dropout regularization to mitigate overfitting and enhance generalization.

Evaluation Metrics :

- Loss Function : Sparse Categorical Crossentropy is employed as the loss function, measuring the disparity between predicted and actual genre labels.
- Metrics : Accuracy serves as the primary evaluation metric, quantifying the model's ability to correctly classify audio clips into their respective genres.

3.3 Model Training

Following data preprocessing, the model is trained on the preprocessed audio features alongside their corresponding genre labels. Training entails iteratively optimizing the model parameters to minimize the loss function and improve classification accuracy.

3.4 Model Evaluation

Upon completion of training, the model's performance is evaluated on a distinct test set to assess its efficacy in genre classification. Performance metrics such as accuracy are computed to gauge the model's ability to generalize to unseen data.

4. Observations

During the evaluation phase, several observations are made regarding the classifier's performance across different music genres. Additionally, the significance of specific features in classification may be analyzed, shedding light on the underlying patterns and characteristics that contribute to successful genre classification. Any encountered challenges during training and testing are meticulously documented, providing valuable insights for future iterations and improvements.

5. Conclusion

In conclusion, this project represents a concerted effort to automate music genre classification through the application of cutting-edge deep learning methodologies. While the model demonstrates promising accuracy, there remains ample scope for further optimization and feature engineering to enhance its performance. This project serves as a foundational endeavor

in leveraging machine learning for music genre classification, offering valuable insights into model development, evaluation, and potential avenues for future research and enhancement within the domain of audio processing.