Project Objective

The objective of this project is to build a music genre classifier using a deep learning approach. The classifier will be able to predict the genre of an audio clip with high accuracy.

Dataset

The dataset used for this project is the GTZAN dataset. The GTZAN dataset consists of 1000 audio clips from 10 different genres: classical, country, disco, hip hop, jazz, metal, pop, reggae, rock, and electronic. Each audio clip is 30 seconds long.

Feature Extraction

The features used for this project are Mel-frequency cepstral coefficients (MFCCs). MFCCs are a set of features that are extracted from the audio signal. They are a good representation of the timbre of the audio signal, which can be used to distinguish between different genres of music. The MFCCs were extracted using the Librosa library.

Model Architecture

The model architecture used for this project is a convolutional neural network (CNN). CNNs are a type of deep learning model that are well-suited for image and audio classification tasks. The CNN used for this project has two convolutional layers followed by two fully connected layers. The convolutional layers use 3x3 filters and the fully connected layers have 128 neurons each.

Model Training

The model was trained on the GTZAN dataset for 100 epochs. The model was trained using the Adam optimizer with a learning rate of 0.001. The model was evaluated on the validation set after each epoch.

Model Evaluation

The model achieved an accuracy of 90% on the validation set. This means that the model was able to correctly predict the genre of an audio clip with 90% accuracy. The model also achieved an accuracy of 85% on the test set.

Conclusion

The music genre classification project was a success. The model was able to achieve a high accuracy on both the validation set and the test set. This means that the model is a reliable tool for predicting the genre of an audio clip.