**Music Genre Classification**

Akash Bose (Univ. Roll: 12200217061, Reg. No.:171220110067)

Anirban Pal (Univ. Roll: 12200217055, Reg. No.: 171220110073)

Dr. Arijit Ghosal (Associate Prof.) Dr. Ranjit Ghoshal (Assistant Prof.)

**Abstract**

**1. Objective:**

The primary aim of this project is to find out suitable features for music genre classification and to create a web application that can predict a music genre from an audio file input using pre-trained machine learning models.

**2. Dataset Used:**

For this project we used GTZAN dataset that consists of 10 genres and each genres consists of 100 audio files of length 30 sec each. The tracks are all 22050Hz Mono 16 bit audio files in .wav format. In our project we worked with five genres namely Blues, Classical, Jazz, Metal and Pop.

**3. Feature Extraction:**

To increase the size of the dataset we segmented each of the 30sec audio files from each genre to 10sec and 3sec audio files and saved them separately so that each genre has 300 audio files and 1000 audio files respectively. Using Librosa we extracted 10 features from each of the audio files which are chroma stft, rmse, spectral centroid, spectral bandwidth, spectral rolloff, zero crossing rate, harmony, percussive, tempo and mel frequency cepstral coefficients. We saved these 10 features in a csv file for each audio file in the dataset and labeled them accordingly for supervised learning.

**4. Training Different Classification Models:**

We trained different classifiers each yielding varying degree of accuracy in prediction. Total seven classifiers have been used for this project and those are k-nearest neighbors, logistic regression, support vector machines, decision tree, random forest, artificial neural network (keras regression model), cnn. Since cnn accepts an image as an input so we have used 286x430 sized image of a mel-spectrogram of an audio signal in this case.

**5. Testing Different Classification Models:**

After training the various classification models we did a comparative analysis of the accuracy achieved from different classifiers. We then created a custom dataset consisting of 29 songs of 30sec duration each and of .wav format. We extracted the above features and then tested our best model to see what predictions it gives for unknown data i.e. audio files external to the dataset. We found out that Knn and Keras models give the best results.

**6. Deployment of Web Application:**

Finally we created a web application using HTML, CSS, Flask, Jinja2 and pre trained models. The web application uploads a .wav music files as an input and predicts its genre. The web application is also deployed on Heroku. [Genre Detection App (music-genre-detection.herokuapp.com)](https://music-genre-detection.herokuapp.com/) .