Classification of Audio Signals according to Genre

# Dataset

In our previous abstract, we had detailed three different datasets which suited the needs of our Machine Learning problem. Out of these datasets, after taking into considerations of dataset size and distinctness in the genres of audio files available in the datasets, we used the **GTZAN Genre Collection** as our training dataset. We have used an existing dataset only because of its standardized nature. Due to the nature of our problem, creating our own dataset is a trivial task.

# Feature Extraction

For representing the time domain audio waveforms, we have used Mel Frequency Cepstral Coefficients (MFCC). Existing research in Speech Recognition and Music Classification pointed us in their direction. Broadly speaking, MFCCs are used to convey the general frequency characteristics important to human hearing. Existing research [1] shows that MFCCs can be used to capture frequency characteristics of Music Files as well. For our assignment, we have considered 15 Mel Features.

# Algorithm Implemented

We used the k-Nearest Neighbour Supervised Learning Algorithm to tackle our problem in this project. The reason for choosing this algorithm is it’s relatively simpler implementation and comparable result to other much complex models such as Multi-Class Support Vector Machine and Neural Networks. The distance between any two songs is measured using Kullback-Lieber Divergence.

For the results obtained by using this algorithm, refer to the ‘Results.xlsx’ file included.

# References

1. Logan, *Mel Fequency Cepstral Coefficients for Music Modeling*, 2000, http://ismir2000.ismir.net/papers/

logan\_paper.pdf