CS 464 Project Proposal
Music Genre Classification from Mel Spectrogram Diagrams

Members

Mustafa Çağrı Durgut CS, 21801983
Muhammed Şamil Arınç EE, 21803683
Yusuf Miraç Uyar CS, 21802626
Oğuzhan Özçelik CS, 21802194
Melih Berk Yılmaz EE, 21803702

Data Description

We are planning to work on genre classification for the project. Free Music Archive Small dataset will be used for this purpose, which consists of 8000 tracks each 30 second long audio files. There are 8 genres consisting of 1000 music data each. For each data, there are 30 sec long wav files. We are planning to generate Mel Spectrogram png files using audio files.

Link to our data: https://github.com/mdeff/fma (We will use fma-small)

Besides, there is another dataset that contains more features already calculated; however, there are lots of projects done using it. So we are planning to use the first dataset we mentioned. Yet, in the future we might utilize this dataset too. This dataset is called GTZAN dataset which consists of 1000(number)x30(sec) audio files. There are 10 genres consisting of 100 music data each. For each data, there are 30 sec long wav files and 432x288 png files showing Mel Spectrogram. Also, there is a csv file which includes 57 different statistical parameters such as variance, mean etc.

Link of the second dataset:
https://www.kaggle.com/andradaolteanu/gtzan-dataset-music-genre-classification

Project Description

Our goal is to classify music genres by using different types of machine learning algorithms. Models will try to predict genres from Mel Spectrogram diagrams. First we will convert our audio dataset files into Mel Spectrogram png files. Then our classifier models will take png files as an input and output a probability for each genre that the audio might belong to.

Milestones

We are planning to compare different machine learning algorithms accuracy scores of classification for this specific dataset.
Our first task is to convert our audio dataset into Mel Spectrogram png files. We are planning to use Logistic Regression and Naive Bayes classifiers on the dataset. Until the first demo, our plan is to preprocess our data and implement Logistic Regression and Naive Bayes classifiers. After that, we will train a CNN model on the dataset. At the end of the final report, we are planning to tune those algorithms if applicable and compare their accuracy rate with each other.