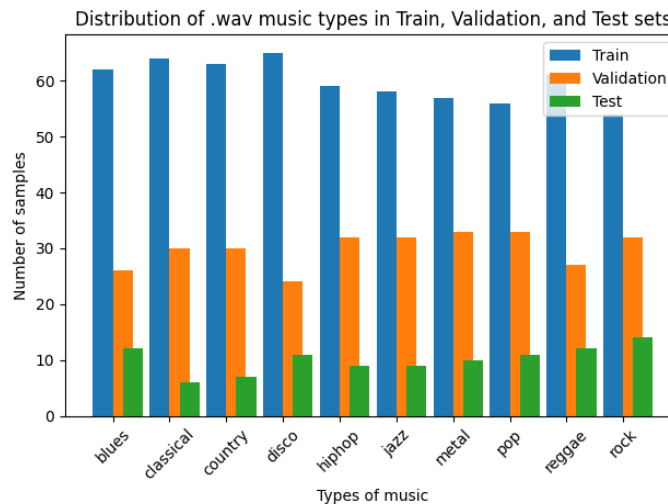


# Assignment 3 Proposal

Nathan MacDiarmid 101098993

## Dataset:

I will be using the GTZAN audio file dataset to train a model to recognize different types of audio genres as I believe it will yield more accurate results. It contains 1000 samples of audio files from genres hip hop, blues, metal, reggae, disco, rock, country, jazz, pop, and classical music. It contains 100 samples from each genre except for jazz, to which it only contains 99 as one of the audio files is corrupted.



## Model architecture:

I will develop a model like that produced in assignment 2 except instead of being a CNN model with convolution layers, I will be developing a Cross Gradient Boost (XGBoost) model with decision trees and gradient boosting, aiding in the continuous optimization of model's accuracy.

## Baseline model:

After evaluating various types of pretrained models for predicting the genre of a given music sample, I found that XGBoost was the most accurate at 90%. Although a KNN can be made equivalent to the accuracy of XGBoost by decreasing its  $n$  nearest neighbours, it runs the risk of poorly generalizing and overfitting the dataset.

## Performance measuring:

To measure the performance of each model, I will be using the f1-score as it yields the most accurate results, considering precision and recall of the model on a given dataset. I predict that my fine-tuned XGBoost model will perform better than the pretrained XGBoost model.