

# Team Dynamo - Music Classification

Drew Hoffman, Bao Nguyen, Jishnu Raychaudhuri

## Problem Space

The goal is to create a classifier that predicts the genre of various music files to provide more insight as to why certain songs belong to certain genres. We are going about this by testing different popular music classification algorithms and seeing how they perform. We believe that by testing different algorithms we will have a better understanding of how to make an effective classifier. From this we hope to create a more interpretive and personalized music recommender.

## Dataset

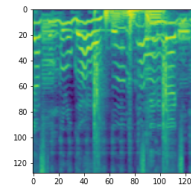
The dataset we used was the widely popular GTZAN dataset. There are 100 samples from 10 genres each, giving us a total of 1000 entries in the form of .wav files. There is also a .csv included containing sound engineering data on each of the songs. In total, there are 59 features in this set, containing discrete and continuous variables.

## Approach

### K-Nearest Neighbors

KNN is a supervised machine learning algorithm that is usually used to solve classification problems. After training, test data is supplied and the model classifies a sample based on the majority class of its k nearest neighbors.

### Convolutional Neural Network



Deep Learning neural network is commonly used in visual recognition and classification. For this method, the audio files were converted into spectrograms, a visual representation of the different sound frequencies and wavelengths of an audio recording.

### K-Means

k-Means is an unsupervised method assigning random initiated points in the data as centroids. This method then calculates repetitively to optimize the centroid positions. We used this algorithm on the .csv data to try and find hidden patterns in the data.

## Results

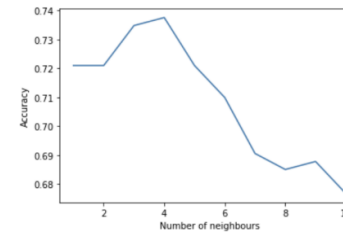


Figure 1: Accuracy vs k value of the KNN model

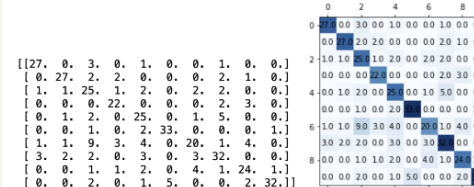


Figure 2: Raw confusion matrix and colour coded confusion matrix

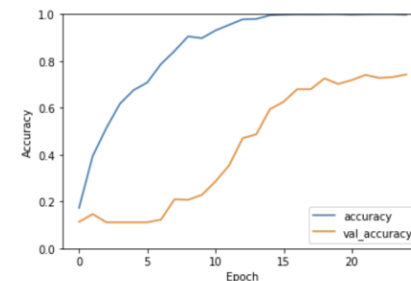


Figure 3: Plot of Accuracy over Epochs during training

## Discussion

For the KNN model, the classifier was most effective at classifying jazz music, and most often misclassified rock music. Genre pairs that were often confused for each other include classical and jazz, and hip hop and reggae.

The CNN model most often misclassified songs as blues or hip hop, and was adept at classifying these genres correctly. The classifier rarely classified songs as "pop", and in the future we may want to use a smarter trimming system

K-means is well suited for large datasets due to the fact that it is an unsupervised algorithm, and it helped find hidden patterns in the data and similarities between genres.

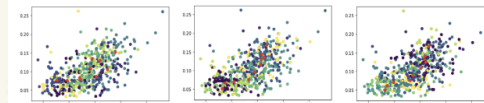


Figure 4: True labels, K-means predicted labels, K-means++ predicted labels

After obtaining our testing set accuracy we ran the following songs through each of our models to see what each classifier would predict. We thought that each of these songs has a distinct and recognizable genre to the human ear and wanted to see if the classifiers were able to capture that. The results for each mode are given below.

Song	Artist	KNN	CNN	Actual Genre
Sweet Child O'Mine	Guns N' Roses	Pop	Rock	Rock
Fur Elise	Beethoven	Country	Blues	Classical
Three Little Birds	Bob Marley	Pop	Blues	Reggae
N.Y. State of Mind	Nas	Hiphop	Hiphop	Hiphop
Stayin' Alive	Bee Gees	Pop	Hiphop	Disco
Mercy, Mercy, Mercy	Cannonball Adderley	Reggae	Blues	Jazz
22	Taylor Swift	Pop	Disco	Pop
Enter Sandman	Metallica	Pop	Rock	Metal
Take Me Home, Country Roads	John Denver	Pop	Reggae	Country
Billy's Blues	Billy Stewart	Hiphop	Blues	Blues

Table 1: First three entries and first eight columns of the data set

filename	temp	beats	chroma_stft	rmse	spectral_centroid	spectral_bandwidth	rolloff
blues.00081.au	103.359375	50	0.38026021	0.24826229	2116.94296	1956.61106	4196.10796
blues.00022.au	95.703125	44	0.30645087	0.11347541	1156.0705	1497.66818	2170.05354
blues.00031.au	151.999081	75	0.2534871	0.15157077	1331.07397	1973.64344	2900.17413