Music Genre Classification

# **Team**

TAARINI V-18BCS089 VISNU SANKER 18BCS112

PRIYADHARSHINI T-18BCS090 SITTHANADHAN K-18BCS106

KALAICHANDRAN C-18BCS100 SAKTHIVEL S 18BCS208

# ABSTRACT

In the area of music information retrieval (MIR), categorizing music according to their genre is a challenging task. We studied and implemented k-nearest neighbors , a classification algorithm admitting two different types of inputs.The model is trained end-to-end, to predict the genre label of an audio signal,with the help of spectrogram. Being able to instantly classify songs in any given playlist or library by genre is an important functionality for any music streaming/purchasing service.

**INTRODUCTION**

This project is primarily aimed to create an automated system for classification model for music genres. Genre usually assumes high weight in music recommender systems. Genre classification, till now, had been done manually by appending it to metadata of audio files or including it in album info. This project however aims at content-based classification, focusing on information within the audio rather than extraneously appended information.

Companies nowadays use music classification, either to be able to place recommendations to their customers (such as Spotify, Soundcloud) or simply as a product (for example Shazam). Determining music genres is the first step in that direction. Machine Learning techniques have proved to be quite successful in extracting trends and patterns from the large pool of data. The same principles are applied in Music Analysis also.

# **LITERATURE SURVEY**

Many companies nowadays use music Classification, either to be able to place Recommendations to their customers (such as Spotify, Sound cloud), or simply as a product (for example Shazam). Determining specific Music genres is a first step towards this goal. In this paper, we will present our successive Steps towards building different classifying Methods allowing us to identify a specific Genre from an initial audio file. We will first Describe how we collected data, and why this Choice of data was pertinent. Then, we offer a Possible conversion of this data into exploitable Information, and perform feature selection. We Will then progress onto presenting our various Algorithms and machine learning techniques Used for classification. The final output of our Algorithms is the prediction of the genre of Each input. We also quickly diverge into composer classification for classical music. Finally, we present Our results that we have obtained while studying this problem.

**DATASET DESCRIPTION**

Dataset used - GTZAN dataset

Repository – MARYSAS

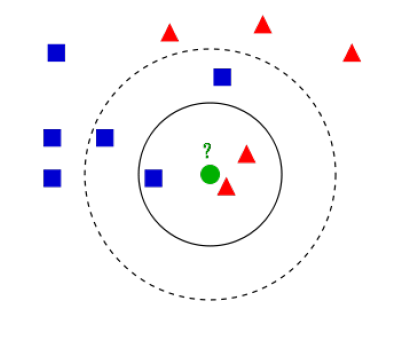
It contains 9 music genres, each genre has 100 audio clips in .wav format. The genres are – blues, classical, country, disco, pop, jazz,reggae,rock, metal. Each audio clips has a length 30 seconds, are 22050Hz Mono 16-bit Files.

The dataset incorporates samples from variety of sources like CDs, radios,Microphone recordings etc.We split the datset in 0.9 : 0.1 ratio and used 5-fold cross validation for reporting the results.

**ALGORITHM DESCRIPTION**

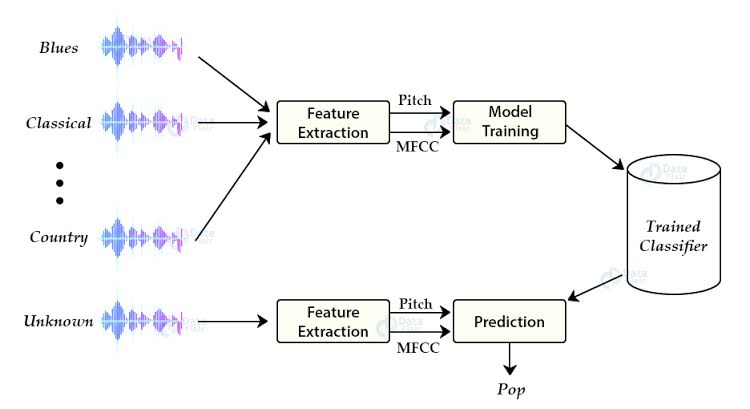
K-Nearest Neighbors Our last approach was the widely used technique of K-Nearest Neighbors. This is often one of the first techniques applied to classification problems because of its accessibility of implementation. We first applied Principal Component Analysis to our features vector to reduce the dimension of our features space to three, and then ran the K-NN algorithm. The idea here is to "place" our training set in space, coloring each example using our labels, pick k, and then find the k "nearest" neighbors of the testing example that we are trying to classify. We then make our prediction as the most represented class amongst the neighbors. The parameters in this model are the distance function, and k. The latter is chosen by training on different values of k, and picking the best one. The distance is often chosen to be the euclidean distance

**d(x1, x2) = |x1| |x 2 | 2**



**PROPOSED SYSTEM FLOW**

To classify our audio clips, we chose 5 features: Mel-Frequency Cepstral Coefficients, Spectral Centroid, Zero Crossing Rate, Chroma Frequencies, Spectral Roll-off. These 5 features are appended to give a 28 length feature vector. Then,we used dierent multi-class classiers and an ensemble of these to obtain our results.A flowchart is been provided below for better understanding:



**Results and Discussion**

The main quantitative metric which we used to judge our models is accuracy (that is, percentage of predicted labels which matched their true labels), and our main way of visualizing the performance of our best model is through the confusion matrices. Because the labeling was uniformly distributed on our data, cross-validation, and test sets, these confusion matrices offer not only a way to visualize our data, but more specific information than precision and recall values offer.We selected our hyper parameters based of empirical results and industry standards. Working with audio time-series data over 2 dimensions cause sparse gradient problems, similar to those often encountered in natural language or computer vision problems.

**REFERENCES**

[1] Hareesh Bahuleyan. Music genre classification using machine learning techniques

[2] G. Tzanetakis and P. Cook. Musical genre classification of audio signals. IEEE Transactions on Speech and Audio Processing, 10(5):293–302, July 2002.

[3] Y. Panagakis, C. Kotropoulos, and G. R. Arce. Music genre classification via sparse representations of auditory temporal modulations. In 2009 17th European Signal Processing Conference, pages 1–5, Aug 2009.

[4] Mingwen Dong. Convolutional neural network achieves human-level accuracy in music genre classification. CoRR, abs/1802.09697, 2018.