Machine Learning Engineer Nanodegree

Capstone Proposal

Rahul Mishra
Jan 2nd, 2018

Proposal

Building a music Genre classifier model.

1.0. Introduction

Classifying Music in CD stores and in online digital media has genre categorization at its base because of its ease to differentiate music in varieties. It has been considered a tradition in CD stores to differentiate shelf on the basis of genre to guide consumers into specific album of specific singer. Apart from that, Online digital media and personal song collection also has genre as one of its main categorization criteria. Humans are quite remarkable in distinguishing Music Genre. Usually a rough classification can be done by us after listening few seconds of music but with the improvement in Machine Learning algorithms, after 2010 many attempts has been made to design system for Automatic Music Genre Classification but the result were not very good. For those who have used the same dataset of Million Song, provided by LabROSA, comprised of 10 different music genres, has yielded an F1-score of 59%, which needs to be improved.

1.1. Motivation

Consumption of digital music is gaining popularity and its distribution over internet is increasing by significant amount day by day. Music genre are generally used to categorize digital music collection so as to facilitate navigation into it. Associating genres automatically to music has become so important firstly because of the increase in number of music unit and secondly because automatic music genre classification generates category independent of manual subjective category. In this digital era, millions of songs are being accessed by consumers. Also, with the technological advancement artist and producer are able to release and distribute songs instantly. This democratization of accessing digital media has arisen the requirement to develop efficient categorization systems. Humans have always been primary tool in attributing genre-tags to songs. Using machine to do the job is a more complex task, but is the requirement of present time. Since, Machine learning excels majorly in handling complex data very well, so, this project addresses learning of various Machine learning algorithm to be used for automatically classifying music files with improved prediction accuracy.

1.2. Advantages of Music Genre Categorization

Automatic classification based on genre will help creating music database in such a way that a general description like 90's classical is given by user and software based on the classifier does the file selection allowing user to create playlist of his own selection criteria. Apart from its practical implication, music genre classification is interesting field of study as well. Rigorous study of the classification and machine based music processing will certainly enhance our knowledge about the human perception of music.

1.3. Abstract

Music has some really powerful effect on our emotions, and Human ear is quite incredible at predicting the genre accurately, although music genre is a relatively complex concept that sometimes the music industry itself feels confused in assigning genre to some of the songs. Classification of genre by Machine is one of such technical approach which has a major drawback of predicting genre accurately. Previously many attempts has been made to design system to categorize music but the prediction result was not overwhelming. So, the problem continued that can machine predict genres of songs with better accuracy results? The Project discusses various Machine Learning concepts of classification and improving the accuracy of those classification techniques to be used for Genre based Music Categorization is the goal of our Project. In this project Million Song Dataset, made available by **LabROSA**, containing 30 summary features for each music files one of 10 genre, being used as training and testing data. The project focuses on implication of Classification algorithms like, **Random Forest**, **Gaussian Naive Bayes**, **Support Vector Machine**, **Decision trees**, **K-Nearest Neighbor**.

Lyrical modelling concept has been the second approach improvising Bag of Words technique over the dataset. We aimed to study and apply various machine learning concepts to solve a real world problem like Music Categorization based on Genre and come out with an improved result of **F1-score** while using Multiclass classification technique and a different F1-score accuracy while applying Lyrical Modelling concept. For the fact that many songs are unlabeled and much more songs are being released daily, the classification technique predicting genre with an **F1-score** better than 58% with which we started is quite bit of help to us.

Domain Background

Historically, attempts made by others to build music genre classification systems have yielded fine but not extraordinary results. Some who used the same dataset that we used for our project, which comprises one million songs each belonging to one of ten different genres, often chose to reduce the number of genres they attempted to identify, while those who decided to work with all ten genres only achieved accuracy of 40%.

The hardest part of music classification is working with time series data. The cutting edge of machine learning research is looking into more effective ways of extracting features from time series data, but is still a long way off. The most promising attempt is done by Goroshin et. al in their paper Unsupervised Feature Learning from Temporal Data.

The problem with temporal data is that the number of features varies per each training example. This makes learning hard, and in particular makes it hard to extract information on exactly what the algorithm is using to “learn”. Algorithms that extract temporal features can be thought of as extracting a more reasonably sized set of features that can be learned from. To address the difficulties of working with temporal data, we used the idea of an acoustic fingerprint to extract a constant sized number of features from each song, improving our accuracy on the testing set substantially.

Problem Statement

The problem we are trying to address here is, Can machine Learn Genre of a song by analysing the various feature of songs. Can we match and/or break the break benchmark set by the benchmark model.

Datasets and Inputs

For this project the Million Song Genre Dataset has been used, that is been made available by **LabROSA** at **Columbia university**, for educational study purpose.This dataset consists of about million songs of various genres, to be precise of 10 genres, namely **classic pop and rock**, **classical**, **dance** and **electronica, folk, pop, hip-hop, jazz and blues, punk, metal & lastly soul and reggae**.The distribution of songs into various genres can be recognized in Dataset Composition. This dataset is a well-organized classification dataset as it consist the features of each of the song in the dataset too. The Million Song Dataset has about million songs with related metadata and audio analysis features, In this dataset, we have each song as a training example containing **34 summary features along with the genre of the track**, which would be of use during training as well as testing the classifiers for predicting the genre of particular combination of features.

The dataset contains various features that needs little description.

genre: is the music category of the particular track in the dataset.
track_id: is the musicmatch track id of the track.
artist_name: is the song's artist name.
title: title of the song.
loudness: is the property of a sound that is primarily a psychological correlation of physical strength (amplitude).
tempo: The tempo is the speed of the underlying beat for a piece of music. Tempo is measured in BPM, or Beats/Minute. The top number represents how many beats there are in a measure, and the bottom number represents the note value which makes one beat.
key: The key of a piece is a group of pitches, or scale upon which a music composition is created.
mode: Refers to a type of scale, coupled with a set of characteristic melodic behaviors.
duration: song duration (in Secs).
avg_timber1 - 12: Timbre is then a general term for the distinguishable characteristics of a tone.
var_timber1 – 12: Timbre is mainly determined by the harmonic content of a sound and the dynamic characteristics.

Apart from this, we will be using **Lyrical Modelling concept**, i.e. to classify genre from the lyrics of the song. We will be collecting available lyrics from the internet and implementing bag of words approach on them.

Solution Statement

We will be using different classification problems which best addresses this problem, and we are provided with a benchmark **F1-score** of 0.59 in the paper, we will try to match and/or better that.

As a addon to this, we will be using Lyrical Modelling concept to classify songs based on genre.

Benchmark Model

Benchmark Model used in this project is the Random Forest model with bayesian optimisation that gave the highest **F1-score** of 0.59.

Evaluation Metrics

Evaluation Metrics used for this project is F1-Score, which is suitable for multi-Class Classification problems.

Genre	Precision	Recall	F1-score	Support
Classic pop and rock	0.61	0.69	0.65	4752
Punk	0.49	0.53	0.51	617
Folk	0.67	0.59	0.63	2680
Pop	0.27	0.23	0.25	327
Dance and electronica	0.61	0.49	0.54	1029
Metal	0.60	0.70	0.65	415
Jazz and blues	0.60	0.46	0.52	846
Classical	0.66	0.71	0.69	367
Hip-hop	0.22	0.53	0.32	98
Soul and reggae	0.42	0.37	0.39	789
Average / Total	**0.59**	**0.59**	**0.59**	**11920**

This is an example of how the mathematical representation of our results will look like where will be calculating and showing ,Recall and Precisions for each genre label.

Project Design

Starting with **Million Song Dataset**, It is huge. I will be using a subset of this dataset.
I will then preprocess this data. Since, the genre distribution is quite unstable, we reshape our database with taking equal number of labels from most available Genres.

The dataset will look something like this from its original composition.

Redistributed dataset Composition	
Genre Name	Number of Songs
jazz and blues	2001
classical pop and rock	2001
classical	1874
punk	2001
metal	2001
pop	2001
dance and electronica	2001
soul and raggae	2001

Preprocessing on Dataset.

- **Feature scaling** needs to be performed over the dataset so that the features are scaled between values (-1 to 1). So, that a single feature doesn't dominates the prediction.
- All the features are needed to be converted to floating types and not mixed type, because if latter is the case numpy arrays generated by us, are converted to string type and that could lead to unexpected results.
- Using **PCA** and **other dimensionality reduction techniques** to reduces the training time and complexity.
- In the **Bag-of-word** scheme, stemming should done before the bagging process as it could reduce the number of bags and could produce more meaningful results.
- Splitting the dataset into training and testing set is little tricky but is quite well handled by sklearn library in a single line split them in 80:20 percentage respectively that too the elements are taken randomly to form the two sets so that no biasing occurs.

Now, Over this DataSet we will be running various Classification Algorithms, tuning there hyperparameters and improving the F1-score over the time.

References

[1] From Classical to Hip-hop: Can Machine Learn Genre? A student publication by Aaron kravitz,Eliza Iupone, Ryan Diaz

[2] Music Genre
https://en.wikipedia.org/wiki/Music_genre

[3] Analysis of the organizational and informational value of links in psychology and geology popular science hyper articles,
http://www.scielo.cl/scielo.php?pid=S071809342008000300004&script=sci_arttext

[4] MSD genre dataset." [Online]. Available:
http://labrosa.ee.columbia.edu/millionsong/sites/default/files/AdditionalFiles/msd genre dataset.zip

[5] Lyrics Retrieving
http://lyrics.wikia.com/

[6] Introduction to Machine Learning,
https://www.udacity.com/course/intro-to-machine-learning--ud120

[7] Preprocessing data
http://scikit-learn.org/stable/modules/preprocessing.html#preprocessing

[8] Python Libraries
http://scikit-learn.org/stable/tutorial/basic/tutorial.html