Halifax, Nova Scotia, Canada

CSCI6505: Machine Learning

Project Proposal

By Team Members:

Tanvi Paruthi - B00875949
Tasnim Khan - B00882598
Nikita Kothari - B00893710
Naga Pavan Srivathsav Chevuru – B00893685

Table of Contents

Proposed Project Topic

Use supervised classification algorithms to identify moods of the music audio.

Project Motivation

While exploring ideas on which we can work for our course project, we came across how we can leverage the music data and its features to generate the moods of the tracks and, as a result, relate to the emotions of the people. Also, our love for music would keep us motivated to work and look for more by associating it with the algorithms that we learn throughout the course. That is how we thought of working on this idea. A general music clip or an audio clip can belong to multiple classes or labels such as happy and energetic, sad and lonely and so on. Hence, we plan to implement a multi-label classification algorithm so that we can classify a song into multiple labels to get better insights. After a while, we could use this learning and behavior to determine the population mood at given geolocation. We want to cite this article that played a crucial role in motivating us to take this as our course project.

Project Outline

This project essentially focuses on performing a supervised classification algorithm using a Spotify Audio Features dataset. This dataset is not readily available, and we will prepare the dataset using this dataset. This dataset would have audio features that fall into two categories: rhythmic and timbre. We will explore the dataset in detail and perform the features selection. We will classify the songs into labels like angry and aggressive, amazed and surprised, and so on. We will analyze results based on all the results of the algorithms.

Main Steps

Step 1: Gathering audio features dataset using this dataset

In the first step of an ML project, we will extract the data from the mentioned dataset to prepare the relevant, comprehensive, and absolute dataset and make it more suitable for the ML algorithms. The dataset will contain audio features which we will explore further.

Step 2: Exploratory analysis of the dataset

Once data collection is complete, we will analyze the dataset and create a relationship between the data to use it more effectively. Also, we will gain a better understanding of all the features in the dataset.

Step 3: Feature selection and normalization of the dataset

After understanding the dataset, we will reduce the dataset by removing the missing, erroneous, or less representative values to target the main attributes and features that have a real significance to the problem statement and add accuracy to the prediction.

Step 4: Feature analysis and identifying proper labels for classification

To start the implementation of the ML algorithms, we will target a few attributes to define and label them clearly. Also, we may create new features out of the existing features to refine the data labelling for a new modelling task.

Step 5: Data splitting – Training, Validation and Test
After labelling the data for classification, we will split the data into three subsets, training, validation and test, where 50%, 30% and 20% will be the allocated percentage to those subsets.

Step 6: Implementing Supervised Learning Algorithm
Implementing supervised learning algorithm to predict mood of the audio track

Step 7: Result analysis
After implementing the algorithms and gaining a better understanding of the predictions, we will analyze the results.

Step 8: Optimization and improvement in results based on our analysis
Based on the results of the predictive values, we will then try to optimize the algorithms, define more features or labels, re-group the dataset, or maximum or minimum function evaluations to improve the results and re-define the model.

Relevant Paper/References

[1] A. Edmans, A. Fernandez-Perez, A. Garel, and I. Indriawan, "Music sentiment and stock returns around the world," *SSRN Electronic Journal*, 2021.

[2] K. Napier and L. Shamir, "Quantitative sentiment analysis of lyrics in popular music," *Journal of Popular Music Studies*, vol. 30, no. 4, pp. 161–176, 2018.

[3] M. T. Quasim, E. H. Alkhammash, M. A. Khan, and M. Hadjouni, "Emotion-based music recommendation and classification using Machine Learning with IOT framework," *Soft Computing*, vol. 25, no. 18, pp. 12249–12260, 2021.

[4] T. Petri, O. Lartillot and T. Eerola, "Exploring relationships between audio features and emotion in music", Frontiers in Human Neuroscience, vol. 3, 2009. Available: 10.3389/conf.neuro.09.2009.02.033.

[5] F. Weninger, F. Eyben, B. Schuller, M. Mortillaro and K. Scherer, "On the Acoustics of Emotion in Audio: What Speech, Music, and Sound have in Common", Frontiers in Psychology, vol. 4, 2013. Available: 10.3389/fpsyg.2013.00292.

Description of the Dataset

We will use this dataset to get the multi-labeled audio features, for example, happy-pleased, relaxing-calm, quiet-still along with other timbre features that would include the first 13 attributes calculated from the Mel Frequency Cepstral Coefficients (MFCCs). The dataset consists of 593 songs with 30 seconds of sound clips from each of the following 7 genres: jazz, reggae, rock, pop, hip-hop, techno, and classical. At an initial phase after normalizing the dataset, we will start with a small dataset. We will go through this paper to explore the correlation between audio features and human emotions.

Also, we are planning to keep the volume of the dataset used for our model agile while doing feature selection and exploratory analysis.

Implementation

In the initial implementation, we plan to gather the raw dataset, identify the appropriate features and implement supervised learning algorithms and analyze the results and relate them with our understanding of the topics.

Anticipated Challenges

1. Identifying tracks of similar/closer human emotions segregation could be a challenge, for example, happy and energetic or sad and slow.
2. Considerable time to identify the correlation between the audio features that we extract from the Spotify dataset to human emotions would be challenging.
3. Classification of emotions could be ambiguous and may have cultural background impact.