Project Proposal: Classifying Song Genres from Metadata and Compressed Song Data

Sam Bennion

Level of Interest: 5/10

1. Description of the project

One project I've always wanted to work on is to build a mobile app that can classify all the Spotify songs I have downloaded into playlists based on genres. I started building the app over winter break, but in order to sort the songs, I need their genre. I'm curious how effective features of a song: tempo, time-signature, artist, year, location, and so on are for classifying genres and organizing music. This is a difficult task, especially because songs can fall into multiple genres, so this is is a multi-classification problem.

2. What features the data set might include

In order to make this project feasible within a semester, I'm probably not going to use song lyrics as a feature, because of the complexity added in turning these into numbers and classifying across languages. I was hoping to turn the audio file into some sort of bitmap in order to quantify the audio, but I may have a hard time using them with the Million Songs Dataset I'm planning to get music data from, as explained in part 3. The other features I'm planning to use are below. The Artist Genre is the output or solution to my model.

There are packages for this

Song Name	Artist Name	Hotness	Latitude	Longitude	**Artist Genre**	Tempo	Time Signature	Year	Mode*	Energy	Danceability	Duration (seconds)
Never Gonna Give you Up	Rick Astley	0.5	41.888 1832	-87.623 177	**Jazz,Rock**	150	4/4	1987	1	0.5	0.2	212

*Mode describes the tone of the piece (major, minor)

3. How and from where would the data set be gathered and labeled

The Million Song Dataset (MSD) is one of the most well used datasets for machine learning research with music. It's a collection of all sorts of smaller datasets, with 54 different features and almost 300 Gigabytes in size. Because of this, only a subset of 10,000 random songs is available to download directly, which I hope to use for the project. Of this, audio data is not available for download, instead 12 dimensional "timbre" and "chroma" vectors for every beat in the song, which I could incorporate into my model inputs if necessary for added complexity. Each feature I'm planning to get from the MSD dataset will require some pre-processing, one the most significant being for "artist mbtags". This is a list of tags for the artist where some of the most popular tags are music genres [1], but not all. These tags will be used to provide the "y" value of my dataset after being preprocessed and cleaned. The Energy and Danceability features come directly from the MSD, calculated based on loudness, beat strength, segment duration [2], and other similar factors. I can either use these parameters or compute my own.

[1] https://ieeexplore.ieee.org/abstract/document/8323324
[2] https://github.com/ronakdpatel/ai-msd