Music Subgenre Classification: Techno, House, Trance, & Drum and Bass

Executive Summary

OVERVIEW

Music genres are essential for organizing and categorizing music, making it easier for listeners to discover, enjoy, and connect with styles that resonate with them. Genres also carry historical, cultural, and sonic significance. Playlists, which often focus on a single subgenre, have become an increasingly popular way to discover new music.

We address the multi-label classification problem to identify a song's genre(s) using acoustic features extracted from audio files. We train a variety of supervised learning models to determine genre. Rather than focusing on broad genres (e.g., jazz, hip hop, electronic), we concentrate on four subgenres of electronic music: techno, house, trance, and drum and bass. While these subgenres are distinct and well-defined, they can be challenging to differentiate.

DATA COLLECTION AND METHODS

We used a subset of the AcousticBrainz data set, which contains a total of about 7.5 million unique songs. Each song can be identified with its unique MusicBrainz ID (MBID), which comes from MusicBrainz, a public database consisting of metadata on music. Due to the massive size of the data set, extracting the entire data set was not practical. Instead, we used an API to query MusicBrainz a list of MBIDs for each subgenre, then we used another API to extract the data from AcousticBrainz with the given MBID. After preliminary data cleaning, our data set had about 37,000 data points and about 2,500 acoustic features.

AcousticBrainz does not store any audio files. Rather, audio characteristics, such as loudness, dynamics, spectrum, beats, and chords, are extracted using Essentia and stored in AcousticBrainz. Many features are split based on bands of frequency, and then various statistics within each band (e.g. mean, variance), resulting in the many features. Genre labels were obtained from MusicBrainz, which are user submitted and then voted by the community. We consider the genre labels to be reasonably accurate.

APPROACH

We used a mixture of correlation analysis and PCA on our features to reduce our features down to 476. We then trained a variety of machine learning models, including k-Nearest Neighbors, Random Forest Classifier, and XGBoost using MultiOutputClassifier, or sometimes OneVsRest, to handle the multi-label nature of our data. Furthermore, we also trained a neural network on both the 476 feature data and the 2500 feature raw data.

RESULTS

XGBoost and our neural network outperformed all other models. Against each other, they had similar performances. XGBoost had higher accuracies in trance (88%) and house (79%), while the neural networks tended to have higher accuracies in techno (86%) and drum and bass (92%). We note that training our neural network was fast and only took about 30 seconds for 50 epochs.

FUTURE WORK

There are several promising future directions to explore. Improving the accuracy of MusicBrainz genre labels by combining sources could enhance confidence in ground truth labels. Investigating additional subgenres in various music styles or further subgenres of electronic music would be valuable. For more lyrical genres, incorporating lyrics as a feature alongside acoustic information could provide deeper insights. While we collected around 40,000 songs from AcousticBrainz, this wasn’t exhaustive and one could obtain more songs. However, this might lead to an imbalanced dataset, especially with the complete collection of drum and bass recordings.