MUSIC GENRE CLASSIFIER

EMILY SCHOOF
SPRINGBOARD CAPSTONE 2
SUMMER 2019

MUSIC BEATS

MB

DISTINGUISH BETWEEN MUSIC GENRES BASED ON SONG AUDIO FEATURES

A DATA SCIENCE ANALYSIS BY EMILY SCHOOF

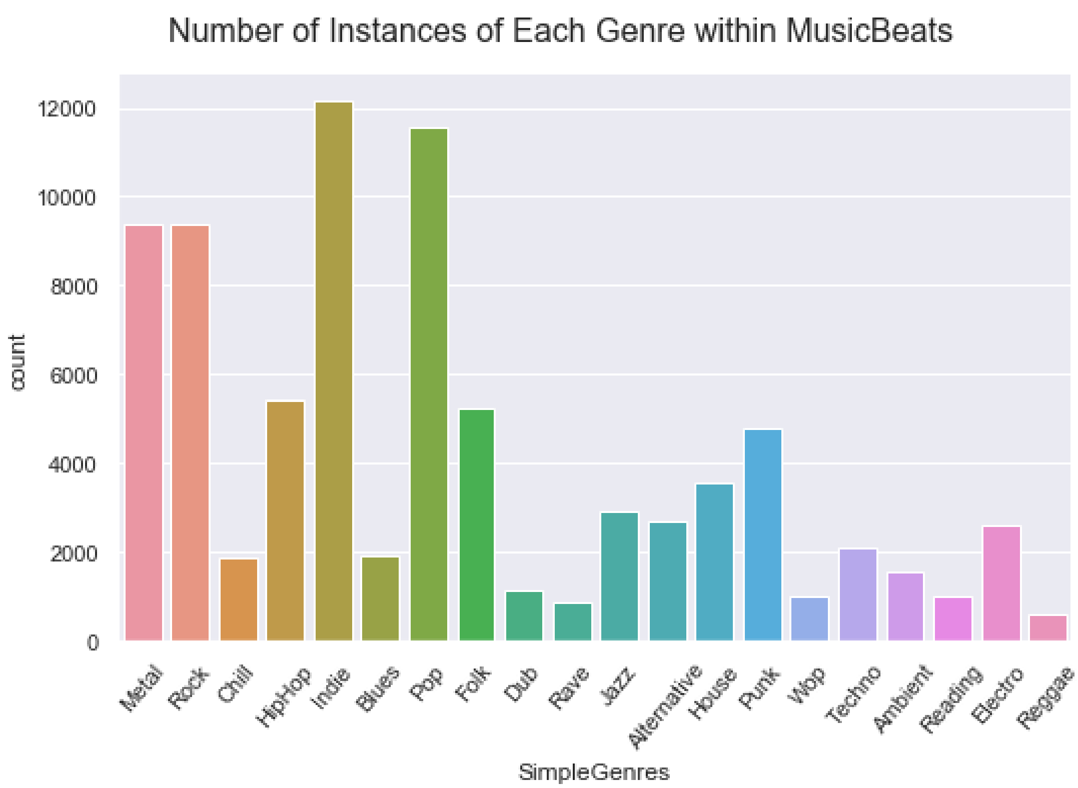
Music Genre Classification is one of the many branches of Music Information Retrieval, which encompasses beat tracking, music generation, recommender systems, track separation, and instrument recognition etc. Music analysis is a diverse field as each music session represents a distinct moment for the user, which makes describing and quantifying this moment an interesting challenge in the Data Science field.

The project data sourced from a Kaggle page, "Spotify music genre list and 80k songs/tracks", that contained over 131580 song instances with 2800 music genres, and 80k extracted audio features (songDB.tsv).

PROJECT IN REVIEW

Natural Language Toolkit (nltk) was used to determine Music Genre Frequency Distribution in order to create a list of 20 root Genre Categories:

- Alternative
- Electro
- Wop
- Reading
- Metal
- Chill
- House
- Indie
- Rave
- Ambinet
- HipHop
- Punk
- Rock
- Reggae
- Blues
- Pop
- Techno
- Dub
- Folk
- Jazz



A SQL query was then used to group each music genre within the sourced dataset in to it's respective Genre category.

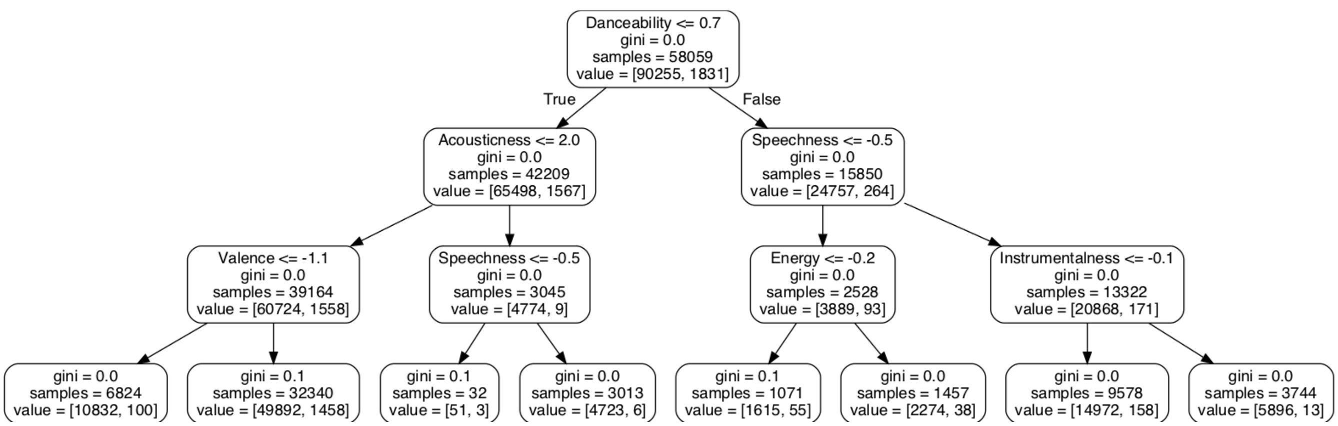
MODEL SELECTION AND SET-UP

The combination of Decision Tree classifiers with a K-Nearest Neighbors (kNN) Machine Learning Model was used to correctly music songs by genre, given the audio features within the ssongDB.tsv dataset.

While other classification models, like Logistic Regression, Support Vector Machines (SVM), and Naive Bayes Classifier, would have worked in place of kNN, I chose to go with the non-parametric model since the distribution of the underlying data would not be accounted for. This was important since the particular procurement of the audio features was not discussed by the owner of the dataset, and the overall distribution of each audio feature appeared to be relatively skewed. Thus, for the sake of this study, a kNN model would produce the most accurate predictions given the features provided.

I. RANDOM FOREST DECISION TREE

Prior to the kNN analysis, I used a Random Forest Decision Tree classifiers since useful for identifying which features within the dataset have the most Information Gain (IG).

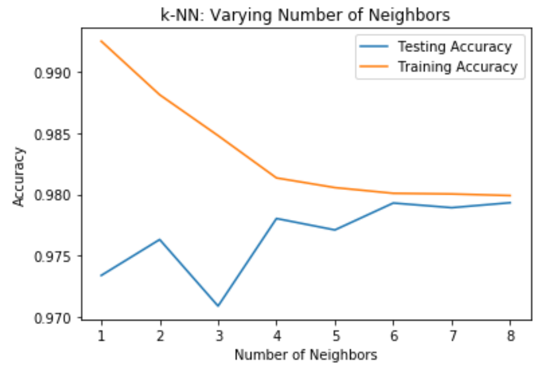


According to the resulting decision tree, 6 variables, 'Danceability', 'Acousticness', 'Speechness', 'Valence', 'Energy' and 'Instrumentalness', were required to separate the genres with a Mean Absolute Error of 2.0%. It is also worth noting that there are repeat splits of 'Speechness', once after the main split at 'Danceability', which appears to helps distinguish between two major classes of genre groups, and once within the left group of genre groups. This suggested that this classifier could not distinctly separate genres by features, alone, so a kNN assessment would potentially be more accurate and more reusable for future data.

II. K-NEAREST NEIGHBORS

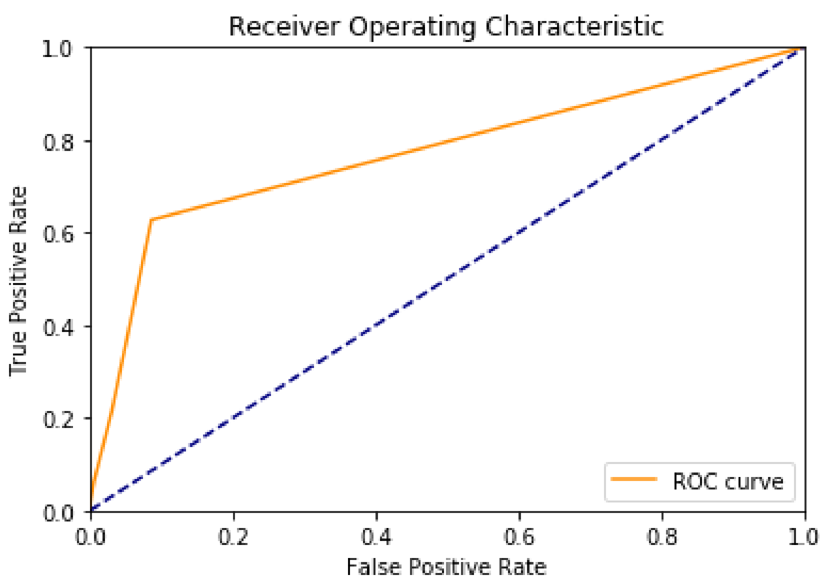
TARGET VARIABLE: SIMPLIFIED GENRES

Prior to the creation of the kNN instance, it was proven that creating a kNN model with at least 6 nearest neighbors would result in the most accurate model for both the training and testing datasets.



PROJECT CONCLUSION

The kNN classification model was successfully able to predict music genre based off of the audio features found in the dataset with 97.9% accuracy.



Further investigation into the procurement of the around audio features should be performed as their values and procurement were not heavily discussed in the Kaggle-sourced dataset. Until this is addressed, there is no way of determining if this model will accurately predict a music genre's by the same audio features if they are not recorded/converted/processed in the same manner. When sufficient proof has been made to support the claim that the audio feature extraction is repeatable, a less computationally dense classifier can replace the kNN model, such as SVM, Logistic Regression (Classification), or a combination in a neural network featuring several layers of classification and feature extraction as part of pipelined predictive model.

This model, when used in the manner described above, could be the music analyzer of a multi-dimensional music recommendation system that works in sync with additional datasets assessing how various users choose music genres.