Spring 4-27-2022

Exploring Music Genres: A Study of Optimal Differentiation by Feature

Rebecca Stetler
stetler@bgsu.edu

EXPLORING MUSIC GENRES:

A STUDY OF OPTIMAL DIFFERENTIATION BY FEATURE

REBECCA STETLER

HONORS PROJECT

Submitted to the Honors College

at Bowling Green State University in partial fulfillment of

the requirements for graduation with

UNIVERSITY HONORS

APRIL 2022

Dr. Umar Islambekov, Mathematics and Statistics, Advisor

Dr. Jeremy Harmon, Music Education, Advisor

1. Introduction

Understanding what makes music popular is of interest to listeners, artists, and producers alike. Listeners are likely interested simply due to curiosity, while artists and producers may be interested in how to create the next hit song. Several researchers have constructed models to predict music popularity, some aiming for the highest prediction accuracy, while others have focused on which variables are most influential in achieving prediction accuracy. In this study, the focus is on identifying variables with the most influence on popularity.

Intuitively, features often vary significantly with genre, which was explored by Hridi in their master's thesis (23). In this study, songs are subset by genre to identify which variables are influential for each genre separately. This also allows for comparison of influential features across genres. This will address the first guiding research question: What features are most influential toward the popularity of songs within individual genres?

Askin and Mauskapf introduce optimal differentiation in their research, which is the idea that there is an optimal level of balance between similarity to and difference from other songs that increases their likelihood of becoming popular (915-928). This study will explore whether optimal differentiation is evident for the features determined to be influential to popularity. This will address the second guiding research question: Where on their genre's feature distribution do hit songs tend to fall?

2. Literature Review

Many researchers have aimed to create music popularity prediction models and genre classification models. An accurate popularity prediction model would help music producers to create songs that successfully become hit songs when received by consumers (Gao 2). Efficient genre classification models are primarily of interest for the goal of predicting songs to users on streaming platforms (Chettiar and S 158). Genre has proven difficult to quantify, as it is an abstract concept based on many features, including social, cultural, and historical factors (Marino 241-249). Several researchers express interest in genre classification and exploration by genre, though, as genre is a mediator between various elements of music and a universally understood music description (Marino 239-240).

Several researchers have created popularity prediction models using various methods and features, and these models have resulted in varying levels of accuracy. Features used for creating models can generally be divided into two categories: acoustic and audio[1]. Lee and Lee built prediction models using audio features, achieving modest accuracy, ranging from 41% to 61% among the proposed models (3182). Gao proposes prediction models using acoustic features for their interpretability, suggesting that this allows for the results to be more readily applicable for music producers and artists (3). Using acoustic features, prediction accuracy was often around 80% (Gao 11-12). Interiano et al. and Middlebrook and Sheik proposed prediction models using acoustic features as well as a feature for the artist's previous popularity. These models achieved accuracies slightly better than with acoustic features alone. Interiano et al. demonstrate about 10% higher accuracy when including the artist past performance feature than without it (14-15). Middlebrook and Sheik achieved approximately 88% accuracy by using acoustic features and a measure of artist past performance (5). Zangerle et al. explore what they describe as "high-level and low-level audio features" as they construct popularity prediction models, aiming to optimize

[1] For the purposes of this analysis, a distinction is made between acoustic and audio features to avoid confusion. Throughout, acoustic features refer to metadata characteristics of the sound of a song. These are abstract features that are interpretable to the general population. Examples of acoustic features are danceability and valence. Audio features are measures taken directly from songs' audio signals. Examples of audio features are Mel-frequency cepstral coefficients and Fourier transform frequencies. Audio features are more direct measures of the music but are generally less interpretable.

prediction performance using high-level features, low-level features, or both. The best performance, about 75% accuracy, is achieved using all features (324). It should be noted, however, in comparing accuracy levels of models, that the way authors define and measure "popularity" varies significantly.

Audio features, perceived acoustic features, and artist past performance have all been demonstrated to influence music popularity by the relative success of the prediction models previously assessed. However, there is significant room for improvement in these models, suggesting that popularity is somewhat unpredictable by nature. This is explored in an experiment conducted by Salganik et al., who hypothesized that social influence is contributing to the unpredictability of music popularity. Their experiment reveals this to be the case, as they find that a participant's perception of a song's quality was influenced when provided information about other participants' opinions (854-856).

By exploring how similarity between songs is related to song popularity, Askin and Mauskapf describe the similarity-differentiation tradeoff, finding that there is an "optimal differentiation" in song typicality that gives a song the best chance for achieving high popularity. That is, if a song is too similar to or different from other songs, it is less likely to become a hit (Askin and Mauskapf 915-928). As demonstrated in several of the aforementioned articles studying music popularity prediction, hit song prediction is not exactly a science, likely due to both frequently changing trends (Zangerle et al. 324) and social influence (Salganik et al. 854-855); however, there are underlying audio and acoustic features that have been able to account for much of the variation in song popularity.

Recent research on classifying music into genres has expanded as streaming platforms, such as Spotify, expand. Grouping songs by genre allows users to more easily find new songs they might enjoy. Research studying genre classification models has taken a similar path to popularity prediction models regarding feature selection; that is, the groupings of audio and acoustic features are prevalent in this research as well. Using audio features, Chettiar and S. achieved up to 70% classification accuracy with their models (160). Goulart et al. achieved around 80% accuracy using various measures of entropy, an audio feature (61-62). In their master's thesis, Hridi used acoustic features to construct a genre classification model with 78% accuracy (34-35). Exploratory data analysis of acoustic features for each genre also suggested varied distribution of features by genre (Hridi 23). Setiadi et al. achieved 80% accuracy by selecting a subset of acoustic features. They also identify the most important features to their genre classification model as acousticness, instrumentalness, popularity, and energy, in order of importance (3-4), which may suggest these features vary significantly between genres.

3. Data

This study uses a dataset of over 140,000 unique tracks from Spotify and measures of each song's characteristics. In total, the dataset includes over 200,000 tracks, as many of the songs are repeated because they belong to more than one genre classification. In addition to each song's name, id, artist, and genre, several other variables measuring characteristics that have been calculated for each song and made available through the Spotify Web API are described in Table 1 below ("Spotify Web API"). As seen in the description for the popularity variable, a song's popularity measure is dependent on the time the data was acquired from Spotify. This dataset was obtained from Spotify in approximately July 2019. This dataset was acquired for this research via Kaggle, a website with several thousand open datasets to be used by data scientists (Hamidani).

Variable	Description
acousticness	"A confidence measure from 0.0 to 1.0 of whether the track is acoustic. 1.0 represents high confidence the track is acoustic."
danceability	"Danceability describes how suitable a track is for dancing based on a combination of musical elements including tempo, rhythm stability, beat strength, and overall regularity. A value of 0.0 is least danceable and 1.0 is most danceable."
duration_ms	"The duration of the track in milliseconds."
energy	"Energy is a measure from 0.0 to 1.0 and represents a perceptual measure of intensity and activity. Typically, energetic tracks feel fast, loud, and noisy. For example, death metal has high energy, while a Bach prelude scores low on the scale. Perceptual features contributing to this attribute include dynamic range, perceived loudness, timbre, onset rate, and general entropy."
instrumentalness	"Predicts whether a track contains no vocals. "Ooh" and "aah" sounds are treated as instrumental in this context. Rap or spoken word tracks are clearly "vocal". The closer the instrumentalness value is to 1.0, the greater likelihood the track contains no vocal content. Values above 0.5 are intended to represent instrumental tracks, but confidence is higher as the value approaches 1.0."
key	"The key the track is in."
liveness	"Detects the presence of an audience in the recording. Higher liveness values represent an increased probability that the track was performed live. A value above 0.8 provides strong likelihood that the track is live."
loudness	"The overall loudness of a track in decibels (dB). Loudness values are averaged across the entire track and are useful for comparing relative loudness of tracks. Loudness is the quality of a sound that is the primary psychological correlate of physical strength (amplitude). Values typically range between -60 and 0 db."
mode	"Mode indicates the modality (major or minor) of a track, the type of scale from which its melodic content is derived."
popularity	"The popularity of the track. The value will be between 0 and 100, with 100 being the most popular. The popularity is calculated by algorithm and is based, in the most part, on the total number of plays the track has had and how recent those plays are."
speechiness	"Speechiness detects the presence of spoken words in a track. The more exclusively speech-like the recording (e.g. talk show, audio book, poetry), the closer to 1.0 the attribute value. Values above 0.66 describe tracks that are probably made entirely of spoken words. Values between 0.33 and 0.66 describe tracks that may contain both music and speech, either in sections or layered, including such cases as rap music. Values below 0.33 most likely represent music and other non-speech-like tracks."
tempo	"The overall estimated tempo of a track in beats per minute (BPM). In musical terminology, tempo is the speed or pace of a given piece and derives directly from the average beat duration."

valence	"A measure from 0.0 to 1.0 describing the musical positiveness conveyed by a track. Tracks with high valence sound more positive (e.g. happy, cheerful, euphoric), while tracks with low valence sound more negative (e.g. sad, depressed, angry)."

Table 1: Descriptions of variables from developer.spotify.com/ ("Spotify Web API")

4. Methods

4.1 Prediction Models

The data was divided into multiple datasets by genre. Analyses focus on six genres: classical, country, jazz, pop, rap, and rock. Each genre's dataset includes between 8500 and 9500 songs. Using R, a programming language frequently employed for statistical analysis, prediction models of various types were created for each genre, with popularity as the response variable and all other variables in Table 1 as predictors. Models were created using the caret package with repeated cross validation. The purpose of creating these models was to determine the most influential variables in predicting song popularity in each genre.

The first type of model constructed were regression models, which result in equations of the form

$$y = ß_0 + ß_1 x_1 + \dots + ß_k x_k + \epsilon,$$

where y is the response (popularity in this case), x_1, x_2, …, x_k are the regressor variables, $ß_0$, $ß_1$, …, $ß_k$ are the regression coefficients, and ϵ is the random error term. The regression coefficients help demonstrate the impact each regressor has on the response variable. Simple linear regression and stepwise linear regression models were created for each genre.

Several models based on decision trees were also created. These models were bagging, random forests, and boosting. Bagging improves the accuracy of a single decision tree by averaging the results of several predictions. Random forests are similar to bagging but tend to reduce correlation between trees by only choosing from some random selection of features at each split in a tree. Boosting also averages the decisions of many trees, but this approach works sequentially to intentionally improve performance of the model by using information from previous trees. Alternatively, bagging creates all trees independently, which reduces variance.

Models were also made using support vectors machines, which allow for nonlinear decision boundaries. Three support vector machine models were made for each genre, each corresponding to a different kernel: linear, polynomial, and radial.

4.2 Evaluating Model Accuracy

Each genre's dataset was split into training and test sets. In each genre, approximately 75% of the songs were assigned randomly to the training set, and the remaining 25% constituted the test set. For each genre, each model used the same training and test sets. The prediction models were created using the training set, and the test set was used to assess the accuracy of each model. Three accuracy measures were compared to evaluate the prediction accuracy of each model. These three metrics are root mean squared error (RMSE), simple R^2, and mean absolute error (MAE). These measures assess how close a predicted value is to the observed value.

4.3 Importance Measures

The prediction models created were used to determine the most influential variables in predicting popularity for each genre. Using the caret package, importance measures were calculated for each variable. These importance measures identify which variables are most useful to the model in making its predictions. This study uses scaled importance measures, so these measures range from zero to one hundred, with the most important variable to a model having a score of one hundred.

4.4 Evaluating for Optimal Differentiation

Based on the importance measures described above, the three most influential variables in predicting popularity for each genre were identified. For each of these variables, the notion of optimal differentiation was explored as follows. A density plot was created to display the distribution of the variable for all songs in the genre. This plot was overlayed with a density plot of the same variable for the top ten percent of songs, by popularity, in the genre. By comparing the distribution of the feature for all songs and the most popular songs, these tracks were evaluated for the concept of optimal differentiation of top songs from all songs.

5. Results

5.1 Popularity Prediction Models

5.1.1 Model Accuracy

Within each genre, each popularity prediction model achieved similar accuracy. The accuracy measures are shown below in Tables 2-7.

	Root Mean Squared Error	R^2	Mean Absolute Error
Boosting	12.399	0.228	9.249
Random Forests	**12.186**	**0.259**	**9.085**
Bagging	12.759	0.183	9.643
Linear Regression	13.124	0.136	9.893
Stepwise Linear Regression	13.171	0.129	9.900
Support Vector Machine – Linear Kernel	13.273	0.131	9.703
Support Vector Machine – Polynomial Kernel	12.818	0.184	9.349

Support Vector Machine – Radial Kernel	12.707	0.197	9.267

Table 2: Classical prediction models' accuracy ratings

	Root Mean Squared Error	R^2	Mean Absolute Error
Boosting	9.501	**0.027**	7.280
Random Forests	9.521	0.022	7.281
Bagging	**9.497**	0.027	7.259
Linear Regression	9.563	0.014	7.331
Stepwise Linear Regression	9.582	0.009	7.355
Support Vector Machine – Linear Kernel	9.666	0.014	7.250
Support Vector Machine – Polynomial Kernel	9.651	0.017	**7.240**
Support Vector Machine – Radial Kernel	9.669	0.015	7.248

Table 3: Country prediction models' accuracy ratings

	Root Mean Squared Error	R2	Mean Absolute Error
Boosting	**9.112**	**0.094**	6.678
Random Forests	9.114	0.094	**6.644**
Bagging	9.186	0.079	6.692
Linear Regression	9.393	0.038	6.791
Stepwise Linear Regression	9.394	0.038	6.738
Support Vector Machine – Linear Kernel	9.413	0.034	6.751
Support Vector Machine – Polynomial Kernel	9.288	0.061	6.746
Support Vector Machine – Radial Kernel	9.345	0.051	6.787

Table 4: Jazz prediction models' accuracy ratings

	Root Mean Squared Error	R^2	Mean Absolute Error
Boosting	**7.231**	**0.028**	5.494
Random Forests	7.261	0.023	5.511
Bagging	7.253	0.023	5.520
Linear Regression	7.262	0.020	5.527
Stepwise Linear Regression	7.256	0.021	5.513
Support Vector Machine – Linear Kernel	7.319	0.020	**5.469**
Support Vector Machine – Polynomial Kernel	7.326	0.018	5.479

Support Vector Machine – Radial Kernel	7.354	0.013	5.498

Table 5: Pop prediction models' accuracy ratings

	Root Mean Squared Error	R^2	Mean Absolute Error
Boosting	**8.135**	**0.037**	6.411
Random Forests	8.162	0.030	6.429
Bagging	8.171	0.028	6.421
Linear Regression	8.192	0.023	6.444
Stepwise Linear Regression	8.200	0.021	6.444
Support Vector Machine – Linear Kernel	8.287	0.021	6.395
Support Vector Machine – Polynomial Kernel	8.260	0.026	**6.368**
Support Vector Machine – Radial Kernel	8.277	0.023	6.382

Table 6: Rap prediction models' accuracy ratings

	Root Mean Squared Error	R^2	Mean Absolute Error
Boosting	7.574	0.009	5.725
Random Forests	**7.567**	**0.012**	5.732
Bagging	7.587	0.005	5.734
Linear Regression	7.595	0.005	5.731
Stepwise Linear Regression	7.596	0.003	5.737
Support Vector Machine – Linear Kernel	7.649	0.003	5.712
Support Vector Machine – Polynomial Kernel	7.635	0.004	**5.699**
Support Vector Machine – Radial Kernel	7.660	0.004	5.738

Table 7: Rock prediction models' accuracy ratings

5.1.2 Variable Influence

Measures of the relative importance of each variable to each prediction model are shown in Tables 8-13 below, each table corresponding to the genre indicated. The far-right column is the mean importance measure for each variable of the boosting model, random forests model, bagging model, linear regression model, stepwise linear regression model, and support vector machine model with a linear kernel. Since the support vector machine models with each kernel determined the same variable importance measures, only one support vector machine model is used in calculating the mean importance measure to avoid unreasonably weighting the scores in favor of the support vector machine models. This mean value in the far-right column of each table was the score used to identify the relative importance of each variable in predicting the popularity of songs of the respective genre. Using the scaled variable importance metrics of each feature to each of the created prediction models, the top three most determinant predictors of popularity for each genre were identified as noted below each table.

	Boosting	Random Forests	Bagging	Linear Regression	Stepwise Linear Regression	Support Vector Machine – Linear Kernel	Support Vector Machine – Polynomial Kernel	Support Vector Machine – Radial Kernel	Mean
Acousticness	50.649	70.391	46.792	39.397	55.054	55.054	55.054	55.054	**52.890**
Danceability	19.802	58.754	13.449	7.380	50.562	50.562	50.562	50.562	33.418
Duration (ms)	77.203	84.497	62.306	14.349	19.845	19.845	19.845	19.845	46.341
Energy	30.762	73.281	41.927	32.279	20.627	20.627	20.627	20.627	36.584
Instrumentalness	100.000	100.000	68.758	83.768	10.191	10.191	10.191	10.191	**62.151**
Liveness	80.541	94.571	100.000	100.000	100.000	100.000	100.000	100.000	**95.852**
Loudness	83.793	85.421	65.533	9.632	18.670	18.670	18.670	18.670	46.953
Mode	3.494	4.360	8.057	16.120	0.000	0.000	0.000	0.000	5.338
Speechiness	28.715	70.817	30.597	43.471	68.032	68.032	68.032	68.032	51.611
Tempo	27.102	59.512	12.923	22.337	0.399	0.399	0.399	0.399	20.446
Valence	36.350	63.220	33.086	35.627	0.832	0.832	0.832	0.832	28.324
Key	0.866	1.435	7.155	5.739	0.255	0.255	0.255	0.255	2.618

Table 8: Classical: Liveness, instrumentalness, and acousticness are the top three predictors of popularity for classical music.

	Boosting	Random Forests	Bagging	Linear Regression	Stepwise Linear Regression	Support Vector Machine – Linear Kernel	Support Vector Machine – Polynomial Kernel	Support Vector Machine – Radial Kernel	Mean
Acousticness	33.381	93.562	3.415	18.496	0.641	0.641	0.641	0.641	25.023
Danceability	53.249	93.405	12.624	10.217	55.402	55.402	55.402	55.402	46.716
Duration (ms)	74.604	96.038	49.721	34.754	1.980	1.980	1.980	1.980	43.180
Energy	57.344	96.218	40.569	38.617	81.574	81.574	81.574	81.574	65.983
Instrumentalness	32.808	53.349	0.000	11.832	0.000	0.000	0.000	0.000	16.331
Liveness	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	**100.000**
Loudness	45.680	96.550	15.631	3.598	1.478	1.478	1.478	1.478	27.402
Mode	4.845	4.691	11.341	28.683	0.600	0.600	0.600	0.600	8.460
Speechiness	86.674	97.384	35.253	50.754	73.911	73.911	73.911	73.911	**69.648**
Tempo	75.449	99.716	31.661	37.206	0.566	0.566	0.566	0.566	40.861
Valence	67.680	96.558	59.879	36.911	82.909	82.909	82.909	82.909	**71.141**
Key	3.391	2.467	15.915	15.570	2.540	2.540	2.540	2.540	7.070

Table 9: Country: Liveness, valence, and speechiness are the top three predictors for country music.

	Boosting	Random Forests	Bagging	Linear Regression	Stepwise Linear Regression	Support Vector Machine – Linear Kernel	Support Vector Machine – Polynomial Kernel	Support Vector Machine – Radial Kernel	Mean
Acousticness	65.597	84.054	31.942	1.897	2.964	2.964	2.964	2.964	31.570
Danceability	57.091	85.278	66.395	61.493	56.342	56.342	56.342	56.342	63.823
Duration (ms)	100.000	100.000	52.964	8.777	5.098	5.098	5.098	5.098	45.323
Energy	53.658	84.474	53.505	22.689	100.000	100.000	100.000	100.000	**69.054**
Instrumentalness	75.053	94.990	100.000	100.000	10.449	10.449	10.449	10.449	65.157
Liveness	36.316	74.226	9.292	19.719	85.538	85.538	85.538	85.538	51.771
Loudness	49.829	81.731	24.205	4.335	1.025	1.025	1.025	1.025	27.025
Mode	2.145	6.808	1.916	0.000	0.000	0.000	0.000	0.000	1.812
Speechiness	84.125	96.236	65.253	39.000	95.387	95.387	95.387	95.387	**79.231**
Tempo	54.011	83.285	12.867	8.873	1.053	1.053	1.053	1.053	26.857
Valence	45.916	80.185	49.428	48.100	84.532	84.532	84.532	84.532	**65.449**
Key	0.664	2.878	2.757	7.963	0.766	0.766	0.766	0.766	2.632

Table 10: Jazz: Speechiness, energy, and valence are the top three predictors for jazz music.

	Boosting	Random Forests	Bagging	Linear Regression	Stepwise Linear Regression	Support Vector Machine – Linear Kernel	Support Vector Machine – Polynomial Kernel	Support Vector Machine – Radial Kernel	Mean
Acousticness	58.110	90.037	54.289	26.635	3.432	3.432	3.432	3.432	39.322
Danceability	100.000	100.000	100.000	100.000	64.429	64.429	64.429	64.429	**88.143**
Duration (ms)	78.306	97.938	68.203	9.881	5.467	5.467	5.467	5.467	44.211
Energy	53.471	88.804	55.158	49.650	70.982	70.982	70.982	70.982	**64.841**
Instrumentalness	44.089	53.270	38.065	42.888	0.000	0.000	0.000	0.000	29.719
Liveness	34.500	85.221	14.710	1.977	93.251	93.251	93.251	93.251	53.818
Loudness	71.236	90.833	53.517	62.917	1.109	1.109	1.109	1.109	46.787
Mode	5.294	8.041	12.927	23.129	0.687	0.687	0.687	0.687	8.461
Speechiness	43.099	86.685	16.568	21.498	100.000	100.000	100.000	100.000	**61.308**
Tempo	74.644	91.002	47.059	14.084	2.192	2.192	2.192	2.192	38.529
Valence	31.823	85.266	8.783	1.186	75.912	75.912	75.912	75.912	46.480
Key	1.110	2.899	13.679	6.590	1.125	1.125	1.125	1.125	4.421

Table 11: Pop: Danceability, energy, and speechiness are the top three predictors for pop music.

	Boosting	Random Forests	Bagging	Linear Regression	Stepwise Linear Regression	Support Vector Machine – Linear Kernel	Support Vector Machine – Polynomial Kernel	Support Vector Machine – Radial Kernel	Mean
Acousticness	39.385	90.838	2.820	14.430	0.185	0.185	0.185	0.185	24.641
Danceability	68.471	93.479	38.383	25.425	56.887	56.887	56.887	56.887	56.589
Duration (ms)	45.079	89.993	0.000	2.880	1.065	1.065	1.065	1.065	17.776
Energy	82.046	92.241	98.267	80.874	61.849	61.849	61.849	61.849	**79.521**
Instrumentalness	15.898	41.058	9.589	0.000	0.000	0.000	0.000	0.000	11.091
Liveness	38.121	88.364	10.853	20.198	90.694	90.694	90.694	90.694	56.487
Loudness	100.000	100.000	100.000	100.000	2.801	2.801	2.801	2.801	**67.600**
Mode	5.064	9.122	17.021	15.517	0.317	0.317	0.317	0.317	7.893
Speechiness	64.639	94.824	60.799	50.616	100.000	100.000	100.000	100.000	**78.480**
Tempo	47.681	91.139	0.000	12.435	0.953	0.953	0.953	0.953	25.527
Valence	38.882	88.616	34.954	14.479	89.106	89.106	89.106	89.106	59.190
Key	0.945	4.236	13.279	14.172	0.927	0.927	0.927	0.927	5.748

Table 12: Rap: Energy, speechiness, and loudness are the top three predictors for rap music.

	Boosting	Random Forests	Bagging	Linear Regression	Stepwise Linear Regression	Support Vector Machine – Linear Kernel	Support Vector Machine – Polynomial Kernel	Support Vector Machine – Radial Kernel	Mean
Acousticness	24.377	90.445	0.000	37.645	0.000	0.000	0.000	0.000	25.411
Danceability	71.745	95.515	24.986	31.029	58.589	58.589	58.589	58.589	**56.742**
Duration (ms)	84.204	100.000	14.891	6.285	1.880	1.880	1.880	1.880	34.857

Energy	4.091	88.296	0.000	1.056	67.833	67.833	67.833	67.833	38.185
Instrumentalness	100.000	75.375	100.000	100.000	2.772	2.772	2.772	2.772	**63.487**
Liveness	82.291	94.105	41.646	36.273	100.000	100.000	100.000	100.000	**75.719**
Loudness	36.099	95.182	34.368	9.065	0.433	0.433	0.433	0.433	29.263
Mode	6.506	6.230	22.520	28.407	0.012	0.012	0.012	0.012	10.614
Speechiness	54.118	96.910	24.678	1.508	74.445	74.445	74.445	74.445	54.351
Tempo	38.168	96.118	19.810	12.171	0.233	0.233	0.233	0.233	27.789
Valence	23.239	93.056	23.051	4.090	82.571	82.571	82.571	82.571	51.430
Key	2.649	1.880	23.602	29.043	2.008	2.008	2.008	2.008	10.198

Table 13: Rock: Liveness, instrumentalness, and danceability are the top three predictors for rock music.

Notably, liveness is the top predictor of popularity for classical, country, and rock music. Classical and rock share the same second-most influential predictor as well, instrumentalness. Speechiness is one of the top three predictors in country, jazz, pop, and rap music. Furthermore, another of the top three predictors for jazz, pop, and rap music is energy, so these three genres all share two of their top three predictors of popularity.

5.2 Optimal Differentiation

In order to explore the notion of optimal differentiation within each genre, the following density plots overlay the feature's distribution for all songs in the genre with the feature's distribution for the 10% of songs in the genre with the highest popularity. These density plots have been created for each of the top three predictive features per genre, as determined above.

As shown below in Figure 1, classical music tends to have low values for liveness. The top ten percent of classical songs follow approximately the same trend in distribution for liveness as all classical songs, with the top songs being even more likely to have a liveness value of around 0.1. Interestingly, Figure 2 shows two clear spikes in classical songs' instrumentalness distribution. These two curves appear to have a higher standard deviation for the top songs. This suggests optimal differentiation: the most popular classical songs more frequently fall just off of the most common trends in instrumentalness for all classical songs. Similarly, Figure 3 shows that, with two spikes in the acousticness plot, the most popular songs' distribution has larger standard deviations, with more of the most popular songs having acousticness measures slightly above or below the most common acousticness measures for classical music.

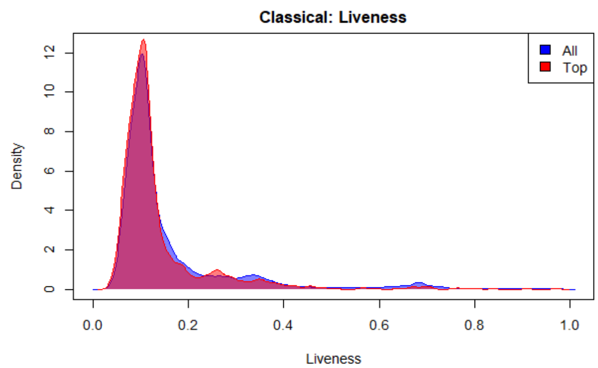

Figure 1

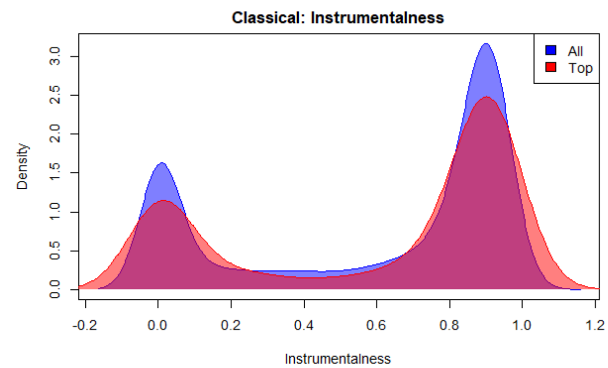

Figure 2

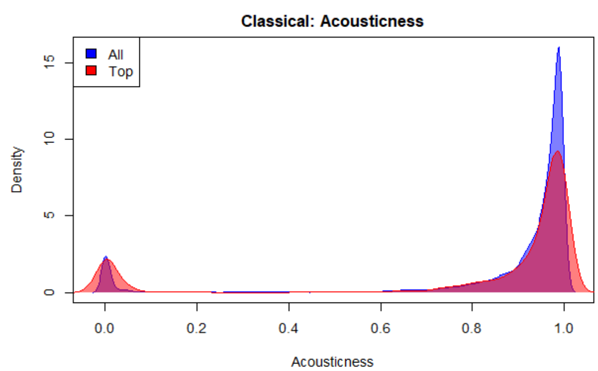

Figure 3

For country songs, liveness is identified as the most influential predictor of popularity by every type of prediction model, yet the distribution of liveness is almost identical for the most popular 10% of country songs as for all country songs. Valence, on the other hand, is more likely to be a bit lower for the most popular country songs. Speechiness, much like liveness, appears to have a very similar distribution between the two sets of country songs.

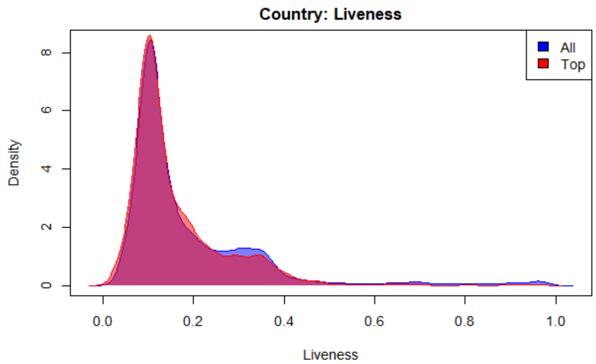

Figure 4

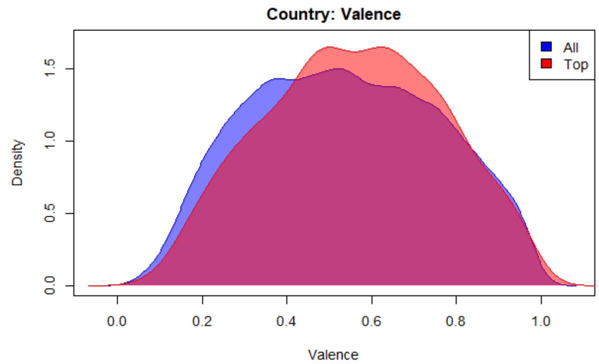

Figure 5

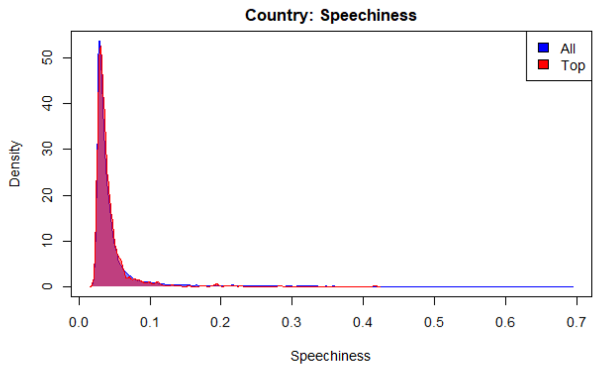

Figure 6

For jazz songs, the distribution of speechiness values is very similar for all songs and the most popular songs, with the most popular songs marginally more likely to have a slightly higher or lower value than the most common value for all songs. The distribution of energy for the top ten percent of jazz songs is more obviously offset from that for all jazz songs. The most popular jazz songs tend to have lower energy measures. The top jazz songs have a spike in valence measures around 0.2, where the set of all jazz songs does not.

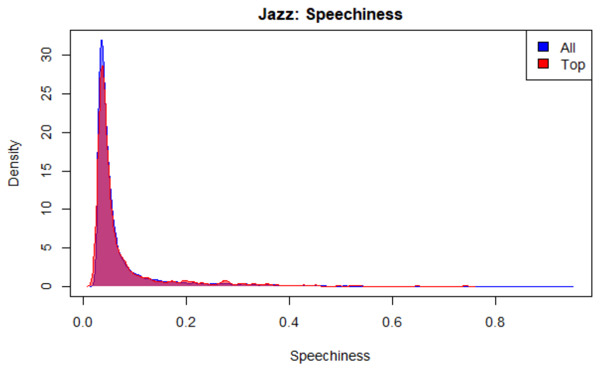

Figure 7

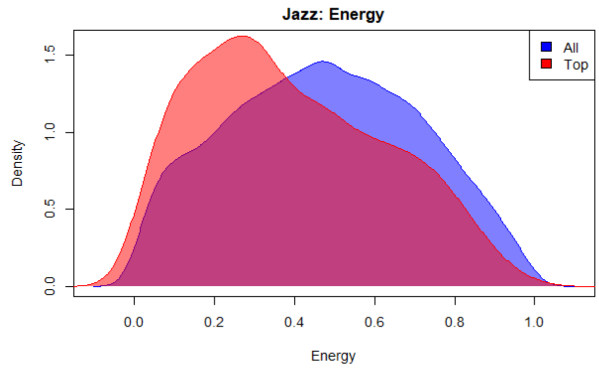

Figure 8

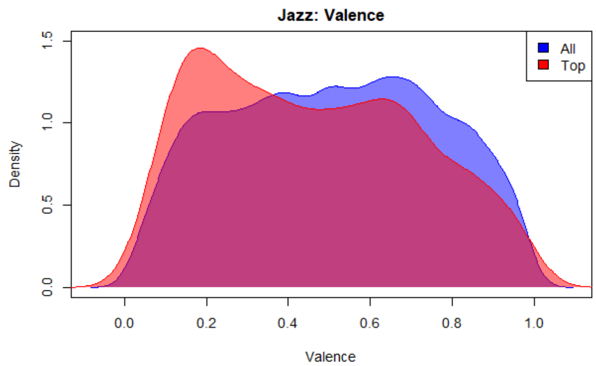

Figure 9

The danceability distributions for pop music demonstrates the concept of optimal differentiation, with the most popular pop songs more likely to have a danceability measure that is higher than the average pop song. The distributions follow a similar shape, but the top songs are concentrated slightly to the side, suggesting an optimal level of differentiation of danceability in pop music from the average pop song. The density plot for energy in pop songs is promising for the argument for optimal differentiation as well, with a concentration of popular songs both to just the right and to just the left of the most common energy measure for all pop songs. Speechiness in pop music also suggests optimal differentiation: the set of top songs are less likely to have the most common values for speechiness and more likely to be slightly higher or lower than these most common values than the set of all songs.

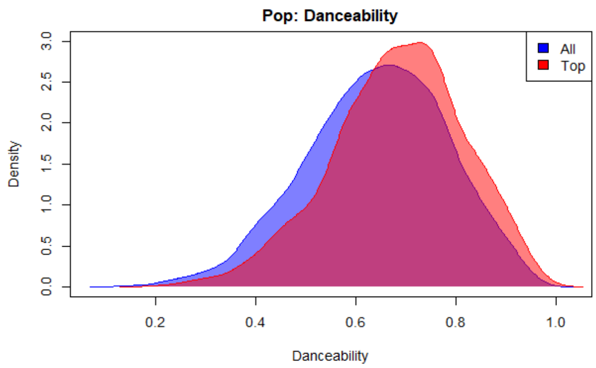

Figure 10

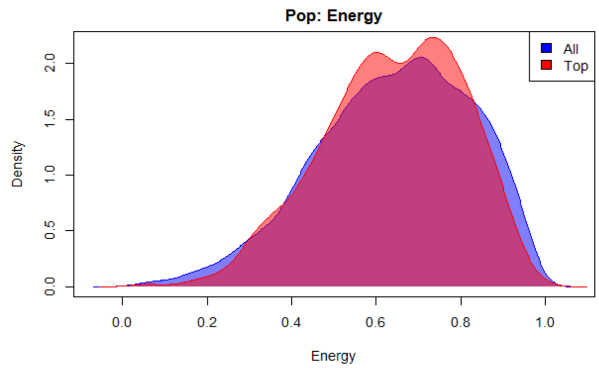

Figure 11

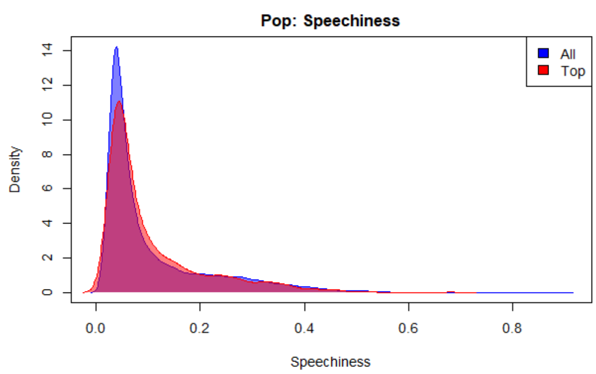

Figure 12

The density plots for the top predictor of popularity in rap music according to the created prediction models, energy, show that the most popular rap songs are more likely to have energy measures a little lower than all rap songs. In the distribution of speechiness for all rap songs, there are two humps, one much higher than the other. The distribution of this variable for the most popular ten percent of songs suggests that rap music's optimal level of speechiness differentiation lies to the right or left of the higher hump, with fewer songs near the lower hump. Loudness follows a very similar trend in all rap songs and top rap songs.

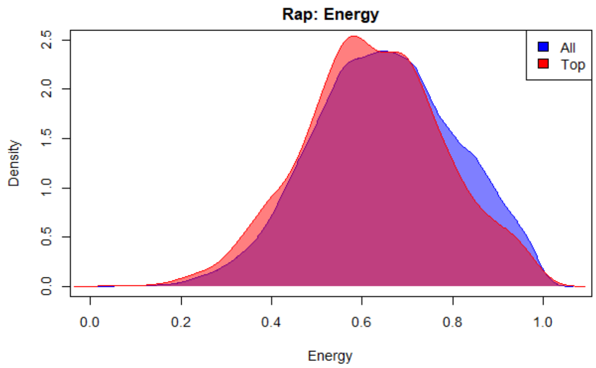

Figure 13

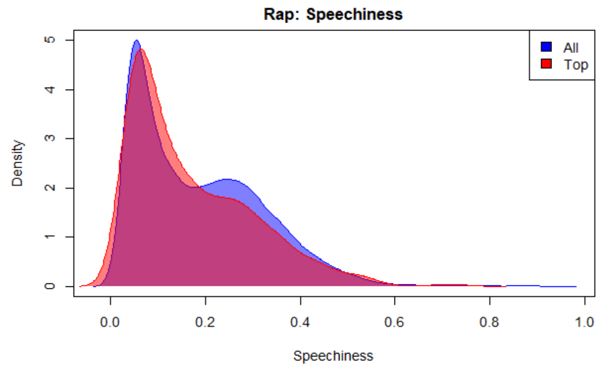

Figure 14

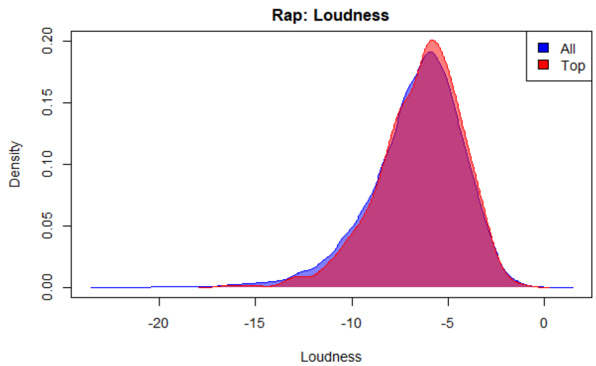

Figure 15

Liveness measures are a bit more spread out in top rock songs than in all rock songs, but overall liveness follows a very similar distribution pattern in both sets. Most of the rock songs in this dataset have an instrumentalness value of zero or close to zero – with 8,260 out of 9,272 total rock songs having an instrumentalness value of less than 0.1. This trend is true for the most popular rock songs as well. Danceability in rock music tends to be a little higher for the top songs than for all songs.

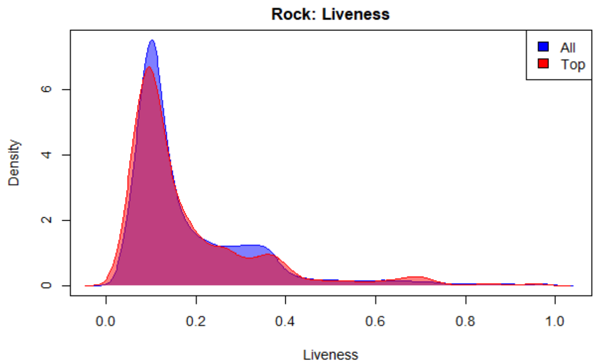

Figure 16

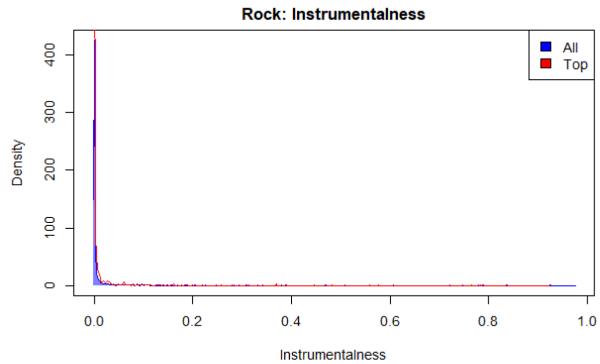

Figure 17

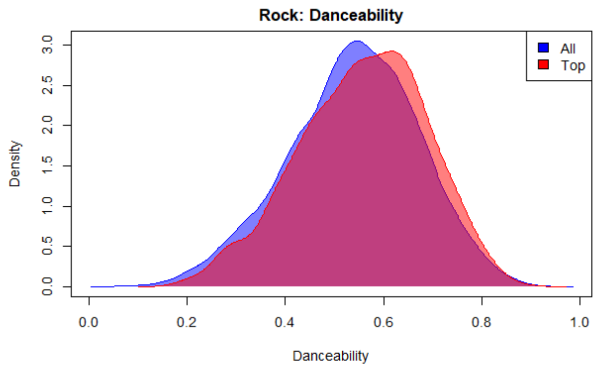

Figure 18

6. Limitations and Future Research

Of course, songs may become popular even if they do not follow the pattern of other popular songs in their genre. Likewise, a song could be engineered to have the same measure for every feature as an extremely popular song, but this would not guarantee it to be a hit. Music popularity is not a strict science, and as such, error will continue to persist in popularity prediction. However, the popularity prediction models in this study could likely be improved with the addition of variables that take into consideration societal and other factors not captured in this study's dataset.

As expressed by Marino, many types of features influence the popularity of a song, including various social, cultural, and economic factors not represented in this dataset (241-249). The accuracy of the prediction models created in this study could likely be improved by taking into consideration factors such as marketing, social media, and cultural events. Audio features, lyrics, existing popularity of the artist, amount spent on advertising, and many other factors are not included in this dataset. The addition of any of these features may improve the accuracy of the prediction models. Furthermore, any of these variables that prove to be useful in predicting popularity could be assessed for the presence of optimal differentiation in future research.

Future research may also consider the release dates of songs. Music trends are constantly changing, as explained by Zangerle et al. (324). This dataset measures song popularity at the particular time the data was extracted, approximately July of 2019, and this study suggests an optimal level of differentiation to be present. This could imply that timing may be an important factor when considering optimal differentiation. Optimal differentiation from other songs released around the same time could be explored in future research.

7. Conclusions

The results of this study agree with the findings of Askin and Mauskapf (931), with evidence in support of the presence of optimal differentiation in music. The above plots and analyses show either nearly matching distributions of influential features between all songs in a genre and the most popular songs in the genre or a pattern suggesting the existence of optimal differentiation. This is especially evident looking at pop music, where all three of its most influential variables in predicting popularity – danceability, energy, and speechiness – demonstrate optimal differentiation in their density plots. Pop music appears to be an effective case study in favor of optimal differentiation. Several other variables' density plots support the idea of optimal differentiation; these are instrumentalness and acousticness in classical music, energy and valence in jazz music, energy and speechiness in rap music, and danceability in rock music. With eleven of the eighteen density plots in this study indicating optimal differentiation, this study provides strong evidence for optimal differentiation at the feature level within music genres.

Works Cited

Askin, Noah, and Michael Mauskapf. “What Makes Popular Culture Popular? Product Features and Optimal Differentiation in Music.” *American Sociological Review*, vol. 82, no. 5, 2017, pp. 910–44, doi:10.1177/0003122417728662.

Chettiar, Gautam, and Kalaivani S. “Music Genre Classification Techniques.” *International Journal of Engineering Research & Technology*, vol. 10, no. 11, 2021, pp. 158–61.

Gao, Andrea. “Catching the Earworm: Understanding Streaming Music Popularity Using Machine Learning Models.” *E3S Web of Conferences*, *253*, EDP Sciences, 6 May 2021, doi:10.1051/e3sconf/202125303024.

Goulart, Antonio Jose Homsi, et al. “Exploring Different Approaches for Music Genre Classification.” *Egyptian Informatics Journal*, vol. 13, no. 2, 2012, pp. 59–63, doi:10.1016/j.eij.2012.03.001.

Hamidani, Zaheen. “Spotify Tracks DB.” *Kaggle*, 23 July 2019, www.kaggle.com/datasets/zaheenhamidani/ultimate-spotify-tracks-db.

Hridi, Anurata Prabha. *Mining User Personality from Music Listening Behavior in Online Platforms Using Audio Attributes*. 2021. Clemson University, Master’s thesis.

Interiano, Myra, et al. “Musical Trends and Predictability of Success in Contemporary Songs in and out of the Top Charts.” *Royal Society Open Science*, vol. 5, no. 5, 2018, doi:10.1098/rsos.171274.

Lee, Junghyuk, and Jong-Seok Lee. “Music Popularity: Metrics, Characteristics, and Audio-Based Prediction.” *IEEE Transactions on Multimedia*, vol. 20, no. 11, 2018, pp. 3173–82, doi:10.1109/TMM.2018.2820903.

Marino, Gabriele. “’What Kind of Genre Do You Think We Are?’: Genre Theories, Genre Names and Classes within Music Intermedial Ecology.” *Music, Analysis, Experience: New Perspectives in Musical Semiotics*, edited by Constantino Maeder and Mark Reybrouck, Leuven University Press, 2015, pp. 239-54, doi:10.2307/j.ctt180r0s2.21.

Middlebrook, Kai, and Kian Sheik. "Song hit prediction: Predicting billboard hits using spotify data." *arXiv preprint arXiv:1908.08609* (2019).

Salganik, Matthew J., et al. “Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market.” *Science*, vol. 311, no. 5762, 2006, pp. 854–56, doi:10.1126/science.1121066.

Setiadi, Rosal Ignatius Moses, et al. “Effect of Feature Selection on The Accuracy of Music Genre Classification using SVM Classifier.” *2020 International Seminar on Application for Technology of Information and Communication (iSemantic), Semarang, Indonesia, 19-20 Sept. 2020*, IEEE, doi:10.1109/isemantic50169.2020.9234222.

“Spotify Web API.” *Spotify for Developers*, 2022, developer.spotify.com/documentation/web-api/reference/#.

Zangerle, Eva, et al. "Hit Song Prediction: Leveraging Low-and High-Level Audio Features." *ISMIR*. 2019.