USING TWITTER TO RANK MUSICAL ARTIST POPULARITY

Shane Storks
Michigan State University
College of Engineering
428 S Shaw Ln #3115
East Lansing, MI 48824
+1 (517) 355-5210
storkssh@msu.edu

Andrew Schmidt
Michigan State University
College of Engineering
428 S Shaw Ln #3115
East Lansing, MI 48824
+1 (636) 300-7874
schmi651@msu.edu

ABSTRACT
The Billboard Artist 100 chart is a relatively new and popular gauge for artist popularity. It ranks musical artists using radio airplay, streaming, and sales data combined with data from several social media platforms. We demonstrate, however, that the Artist 100 can be predicted to fairly high accuracy with various decision tree models trained only on aggregated Twitter data and past chart data, suggesting that all of these data sources are highly correlated, and social media engagement is highly crucial to artist success. We also include a brief investigation into using sentiment analysis to predict the Artist 100.

Keywords
Artist 100, music industry, data mining, prediction, ranking classification, sentiment analysis

1. INTRODUCTION

1.1 Background
The rapid growth of social media platforms has caused musical artists to take advantage of their data to promote music. In 2008, an early adopter of this idea was Lady Gaga, whose team mined data from her Twitter and Facebook followers to identify a network of over 1 million fans who were targeted with promotions for her own website, where fans could engage with Gaga and buy music and merchandise [1].

Next Big Sound (NBS), a data analytics platform for musicians, launched in the following year [2], and as social media became more and more pervasive in American culture over the following years, this platform used data from social media to track up and coming musical artists. After NBS predicted the success of several artists including Gotye and Macklemore sometimes years in advance [1], Billboard partnered with them in 2014 to introduce a new chart: the Artist 100, the first-ever weekly ranking of artist activity, which produced rankings by blending Billboard's radio airplay, streaming, and sales data with social media data from NBS [3].

"There will continue to be an evolution in the platforms people are using, but it's going to be an integral part of managing a music career," claimed Liv Buli of NBS in 2013 [1]. As social media platforms have continued to grow and spread since then, we are interested in determining exactly how crucial these platforms are to the success of a musical artist in 2018. And if they are crucial, we hope that further research will shed more light on optimal ways to use them to increase popularity and profits.

1.2 Approach
If social media engagement is highly crucial to the success of an artist, we expect that music sales are highly correlated with some social media statistics. As such, we attempt to predict the Billboard Artist 100 using only data from Twitter and past charts, rather than the highly multidimensional blend of various types of sales and social media data claimed to be used by Billboard. Using simple classification models, we achieve an RMSE of 18.4 with 100% overlap in predicting a chart of artists in a domain of familiar artists, and an RMSE of 27.7 with up to 96% overlap in predicting a chart including out-of-domain artists, suggesting that social media engagement is indeed highly crucial to an artist's success.

1.3 Contributions
In this work, our contributions are as follows:

(1) A novel and succinct collection of features representative of current musical artist popularity.
(2) A ranking classifier which uses several models to accurately predict the Artist 100, which altogether exceed a simple baseline model in several performance metrics.
(3) An accurate sentiment classifier for tweets about musical artists.
(4) An interactive Web prototype showcasing our collected and simulated data, as well as performance metrics of tested models.

1.4 Roadmap
Following this section, Section 2 describes related work. Section 3 provides a more formal definition of our problem. Section 4 outlines our methodology. Section 5 presents our experimental design and results compared to a simple baseline method. Section 6 provides final conclusions, and Section 7 lists references.

2. Related Work
No published works have yet studied the Artist 100, but past works have attempted to predict the Billboard Hot 100, a chart launched

in 1958 which ranks songs based on radio airplay, streaming, and sales [4].

An early investigation of the Hot 100 occurred in [5], where the authors attempted to predict chart movements based on several features from past chart movements using an elaborate Bayesian model. While their experiments shed light on some movement patterns on the chart, they ultimately did not reach competitive accuracy basing their predictions only on past chart data. Since the experiments occurred in 2001, though, social media data like we use was not yet widely available.

Later, as social media platforms became popular, the authors of [6] use sentiment analysis techniques on comments from MySpace to rank artists. Since the Artist 100 did not yet exist, however, they compared their results to a list generated from artist occurrences on several weeks of the Hot 100. Though the artists in their predicted list had high overlap with the artists in the Hot 100-based list, the ranking was not accurate. They then perform a survey to demonstrate that users tend to agree with their predicted list more than the generated list. However, now that the Artist 100 exists and social media platforms have grown, it would be interesting to see a repetition of this experiment.

In [7] and [8] it is shown that the Billboard Hot 100 can be predicted with about 80% precision using file-sharing data from peer-to-peer networks where songs are shared and often pirated, demonstrating that piracy and song popularity are highly correlated. The authors use common data mining techniques and algorithms, similar to our work. They use several useful metrics in evaluating their predicted charts which we did not consider, however they do not compare their results to a baseline like we do.

In [9], authors show that social data from the music-based Last.fm platform can be used to classify songs as "hits" or "non-hits," where "hits" are songs which appear above a certain rank threshold on the Hot 100, and "non-hits" appear below the threshold. This threshold is moved from rank 1 to rank 50 on the chart, and their classifier achieves near 90% accuracy on some experiments. Similar to our work, high accuracy is achieved with common data mining algorithms, however they do not attempt to predict the rank of Hot 100 songs. Their task is a classification problem, while ours is a more challenging ranking problem.

In [10], it is shown that the path of a song through the Billboard Hot 100 is comparable to a time-series of Web searches for the song, as part of a larger work about predicting consumer behavior from Web searches. However, empirical results from the work indicate that overall, predictions of Hot 100 rank using this data are only slightly correlated with the actual rank (0.56).

Though many of these works aim to compare artists through the Hot 100 using various mathematical and data mining techniques, no works thus far have applied similar methods to predicting the Artist 100, a chart particularly suited for this task. We hope that the success we achieved in our fairly simple investigation inspires further research on this relatively new chart.

3. PROBLEM STATEMENT

Toward our goal of predicting the Artist 100 using only Twitter statistics and past charts, we aim to create a ranking classifier, i.e., a system which must predict a "win", "tie", or "lose" class for each pair of artists who may appear on the chart, and from these class labels predict the ordering of the chart.

To support this task, we collected the following publicly available data for our experiments:

(1) The names of artists who appeared on the Artist 100 from August 5, 2017 through November 10, 2018
(2) Artist 100 charts from August 25, 2018 through November 17, 2018, and the charts for December 1, 2018 and December 8, 2018
(3) Twitter follower counts and mention-per-day counts for each artist collected in (1) from August 14, 2018 through November 9, 2018
(4) Tweets about each artist on the Artist 100 for December 8, 2018 from November 27, 2018 through December 3, 2018

Artist 100 data was collected through the Billboard charts API by Allen Guo,[1] Twitter statistics were collected through the NBS Python API wrapper by Buck Heroux,[2] and tweets were collected through the Twitter API.[3]

For a pair of artists artist 1 and artist 2, we define our class labels as follows:

(1) **-1:** artist 1 charted higher[4]
(2) **0:** artist 1 and artist 2 tied, i.e., neither charted
(3) **1:** artist 2 charted higher

We plan to aggregate this data into several feature vectors per artist in (1) per chart week in (2). Then, per chart week, we plan to concatenate each pair of artists' feature vectors and train several models on the three-way classification task of predicting the label -1, 0, or 1. For testing, we will use a single chart week which is chronologically directly after the training chart weeks. Figure 1 visualizes this process.

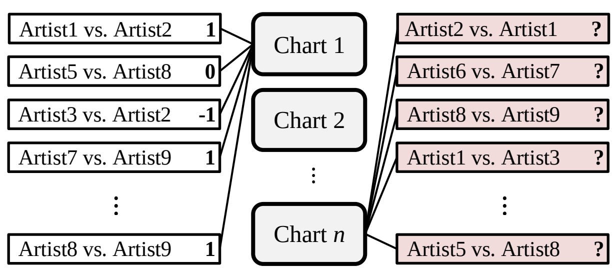

Figure 1: Training and testing process diagram.

[1] Billboard charts API available for download at https://github.com/guoguo12/billboard-charts.

[2] NBS API wrapper available for download at https://github.com/buckhx/NBS-API-Python.

[3] Twitter API documentation at https://developer.twitter.com/en/docs.html.

[4] Here, a lower chart position corresponds to charting "higher," as the top position on the chart is 1, also the lowest.

4. METHODOLOGY

After preprocessing our collected data, we trained and tested the following models from the Python scikit-learn[5] package on our pairwise classification task:

- Logistic regression[6]
- Depth-7 decision tree
- Ensemble of 50 depth-4 decision trees with bagging
- Ensemble of 50 depth-4 decision trees with AdaBoost [11]
- Gaussian naïve Bayes

Pairs of artists are chosen randomly during runtime to ensure there are no biases from the ordering of the data. To produce a ranking from the pairwise classification results, we populate and shuffle a list of the unique artists which our training vectors pertain to, then for each pair of artists (artist 1 and artist 2), we perform the following operations based on class label predictions on the test data:

(1) **-1:** Swap artist 1 and artist 2 if artist 2 is currently closer to the top of the list.
(2) **0:** Move artist 1 and artist 2 to the end of the list.
(3) **1:** Swap artist 1 and artist 2 if artist 1 is currently closer to the top of the list.

We perform two iterations of this swapping to reach convergence, resulting in an ordered list of artist names. The first 100 artist names in the list represent the model's predicted Artist 100 chart. We present the results of three best-performing models which compete with a simple baseline: single decision tree, ensemble decision tree with bagging, and ensemble decision tree with AdaBoost.

[6] uses sentiment analysis of social media comments to rank artists with minimal success. The tweet data from NBS does not take into account the sentiment of tweets about artists, however the sentiment of these tweets may be more reflective of sales. As such, we additionally perform a brief investigation into the utility of using the sentiment of tweets to predict the Artist 100. We create a lexical sentiment analysis system, and order the Artist 100 for a particular week based on its output.

Further, we present our results in an interactive Web prototype, which shows the actual and predicted charts from our in-domain experiments, as well as a summary of social media statistics per artist in each chart and the performance metrics for each predicted chart compared to the actual chart.

5. EXPERIMENTAL EVALUATION

5.1 Experimental Setup

5.1.1 Dataset Characteristics

Before preprocessing, our data consisted of:

(1) 568 artist names.
(2) 13 Artist 100 charts, each including the following fields:
    i. Chart date
    ii. Artist name
    iii. Peak position for artist on the chart
    iv. Previous position
    v. Total number of weeks on chart
    vi. Current position
(3) 85 artist name-date-follower count triples per artist in (1), excluding those which do not use Twitter.
(4) 85 artist name-date-mention count triples per artist in (1), excluding those which do not use Twitter.

We aggregated and integrated this data into two sets of feature vectors: an in-domain and out-of-domain set. The in-domain set only includes data from charts through November 10, 2018, i.e., the charts which the artists in (1) are pooled from. The out-of-domain set includes charts through November 17, 2018. The chart for November 17, 2018 has some artists which are not present in (1), and later serves as test data for an out-of-domain experiment.

The in-domain set contains 6,804 vectors, while the out-of-domain set contains 7,372 vectors. Each vector pertains to a chart date and artist, and has 16 attributes:

(1) **ChartDate:** Artist 100 chart date (always a Saturday)
(2) **ArtistName:** Artist name
(3) **Position:** Current chart position (101 if not on chart)
(4) **LastPosition:** Previous chart position (101 if not on chart)
(5) **PositionTrend:** Difference between previous position and position before that
(6) **WeeksOnChart:** Total weeks on chart (before chart specified by chart date)
(7) **DailyMentions:** Average daily Twitter mentions through chart's tracking week
(8) **TrackingWeekDailyMentions:** Average daily Twitter mentions during chart's tracking week[7]
(9) **DailyMentionsDiff:** Difference between (8) and (7)
(10) **DailyMentionsMult:** Ratio of (8) to (7)
(11) **AvgFollowers**: Average number of followers through chart's tracking week
(12) **MaxFollowers:** Maximum number of followers through chart's tracking week
(13) **DailyFollowers:** Average daily Twitter followers gained through chart's tracking week
(14) **TrackingWeekDailyFollowers:** Average daily Twitter followers gained during chart's tracking week
(15) **DailyFollowersDiff:** Difference between (14) and (13)
(16) **DailyFollowersMult:** Ratio of (14) to (13)

We then concatenated each 2-combination of artists' feature vectors per chart date, producing about 160,000 concatenated feature vector pairs per week, totaling to over 2 million examples in each of our two preprocessed data sets. Each example includes the following 11 attributes associated with two artists (artist 1 and artist 2), chosen from the original date-artist vectors based on a small pilot study:

(1) Artist 1 **LastPosition**
(2) Artist 1 **WeeksOnChart**
(3) Artist 1 **DailyMentionsMult**
(4) Artist 1 **DailyFollowersMult**
(5) Artist 1 **PositionTrend**

[5] scikit-learn documentation at https://scikit-learn.org/stable/supervised_learning.html#supervised-learning.

[6] Used the liblinear solver which is documented at https://www.csie.ntu.edu.tw/~cjlin/liblinear/.

[7] Artist 100 tracking week runs Friday through Thursday, from 15 through 9 days before chart date [12].

(6) Artist 2 **LastPosition**
(7) Artist 2 **WeeksOnChart**
(8) Artist 2 **DailyMentionsMult**
(9) Artist 2 **DailyFollowersMult**
(10) Artist 2 **PositionTrend**
(11) Class label:
   i. -1 if artist 1 **Position** < artist 2 **Position** (i.e., artist 1 charted higher)
   ii. 0 if artist 1 **Position** = artist 2 **Position** (i.e., if neither artist appeared on the chart)
   iii. 1 if artist **Position** > artist 2 **Position** (i.e., artist 2 charted higher)

For both the in-domain and out-of-domain experiments, we save the 160,000 examples from the latest chart week for test data, November 10, 2018 and November 17, 2018 respectively, while leaving the rest for training.

We perform an additional experiment for our investigation of sentiment analysis. We had also collected tweets about the 100 artists on the Artist 100 for December 8, 2018 between November 27, 2018 and December 3, 2018, resulting in about 100 tweets per artist. Due to Twitter API constraints, we were unable to pull tweets directly from the chart's tracking period, but the period we pull for is close to it, so the sentiment should be fairly similar. We use a word-based sentiment classifier to produce sentiment scores for each artist on the chart, and rank them based only on this.

5.1.2 Evaluation

For the purpose of comparing predicted charts to the actual Artist 100, we introduce six metrics which represent the distance between charts and the overlap of charts:

(1) **Chart RMSE:** RMSE of predicted and actual chart positions for the union of artists listed in the predicted and actual charts.
(2) **Chart RMSE (matches only):** RMSE of predicted and actual chart positions for the intersection of artists listed in the predicted and actual charts.
(3) **Chart overlap:** Proportion of artists in predicted chart which appear in actual chart.
(4) **Top 10 RMSE:** RMSE of predicted and actual chart positions for the union of artists listed in the predicted and actual top 10.
(5) **Top 10 RMSE (matches only):** RMSE of predicted and actual chart positions for the intersection of artists listed in the predicted and actual top 10.
(6) **Top 10 overlap:** Proportion of artists in predicted top 10 which appear in actual top 10.

In addition, we provide the exact-match classification accuracy for all in-domain artist pairs in the test data where applicable. For comparison purposes, we calculate a simple baseline by predicting the Artist 100 chart to be the same as the previous week.

Further, we evaluate our sentiment analysis system by manually labeling a small sample of its input tweets and comparing the results to its output.

5.1.3 Environment

Experiments are performed on a 64-bit Windows 10 machine using an x64-based 1.80GHz Intel® Core™ i7-8550U CPU with 16.0 GB of RAM. All code is in Python, with most data collection run from the Windows command prompt, and experiments run from the Spyder Python IDE.[8]

[8] More information at https://www.spyder-ide.org/.

5.2 Experimental Results

We ran three rounds of experiments with the following conditions:

(1) The classifier is familiar with all of the artists in the test chart (in-domain).

(2) The classifier is not familiar with all of the artists in the test chart (out-of-domain).

(3) The classifier is given all of the artists in the test chart and uses sentiment analysis to rank them (sentiment).

Results from selected experiments are available in our Web prototype at http://genesiscrypto.fund/twitterapp.php.

5.2.1 Experiment 1: In-Domain

In this experiment, we trained each model on feature vectors derived from the Artist 100 charts from August 25, 2018 through November 3, 2018, and tested on the Artist 100 chart for November 10, 2018. All artists on the testing chart are present in the classifier's training vectors, thus why we consider the experiment as in-domain.

Results for the three models are listed in Table 1.

	Decision Tree	Bagging Decision Tree	AdaBoost Decision Tree	Baseline
Chart RMSE	**18.42**	21.93	33.29	44.53
Chart RMSE (matches only)	18.42	21.93	33.29	**11.20**
Chart Overlap	**1.00**	**1.00**	**1.00**	0.82
Top 10 RMSE	74.06	84.64	92.27	**73.67**
Top 10 RMSE (matches only)	3.32	4.92	8.00	**2.74**
Top 10 Overlap	**0.60**	0.40	0.20	**0.60**
Pairwise Accuracy	**0.9951**	0.9931	0.9903	--

Table 1: Results from Experiment 1.
Best results for each metric in bold.

5.2.2 Experiment 2: Out-of-Domain

In this experiment, we trained each model on feature vectors derived from the Artist 100 charts from August 25, 2018 through November 10, 2018, and tested on the Artist 100 chart for November 17, 2018. Not all artists on the testing chart are present in the classifier's training vectors, thus why we consider the experiment as out-of-domain.

We expect performance to be worse here than in Experiment 1, although a slight advantage is gained from an extra week of training data. As the classifier has no way to predict artists it has not seen before, these experiments show how representative the classifier's pool of artists is, and the level of accuracy it gains from the guarantee that all artists on the actual chart are in the pool of artists.

Results for the three models are listed in Table 2.

	Decision Tree	Bagging Decision Tree	AdaBoost Decision Tree	Baseline
Chart RMSE	28.54	**27.69**	33.63	39.26
Chart RMSE (matches only)	19.33	17.91	28.43	**14.07**
Chart Overlap	0.95	0.95	**0.96**	0.85
Top 10 RMSE	79.70	88.35	95.12	**66.78**
Top 10 RMSE (matches only)	4.86	5.60	9.00	**3.42**
Top 10 Overlap	0.50	0.30	0.10	**0.70**
Pairwise Accuracy	**0.9951**	0.9946	0.9911	--

Table 2: Results from Experiment 2.
Best results for each metric in bold.

5.2.3 *Experiment 3: Sentiment*

In this experiment, we test a word-based sentiment classifier on ranking the artists in the Artist 100 for November 24, 2018, in hopes that this will shed some light on the utility of sentiment analysis in our problem.

Results are listed in Table 3; the Pairwise Accuracy metric is omitted as it is not applicable here.

	Sentiment	Baseline
Chart RMSE	42.72	**40.43**
Chart RMSE (matches only)	42.72	**14.95**
Chart Overlap	**1.00**	0.84
Top 10 RMSE	97.80	**66.38**
Top 10 RMSE (matches only)	--	**1.31**
Top 10 Overlap	0.00	**0.70**

Table 3: Results from Experiment 3.
Best results for each metric in bold.

5.2.4 *Sentiment Analysis Evaluation*

To check how representative the results from Experiment 3 are of the utility of sentiment analysis alone in ranking the chart, we manually annotate a subset of 992 randomly sampled tweets that the system classified, and compare our results to the system output. In performing this, we found that our annotations agreed with up to 44.0% of the system's predictions for whether a tweet was positive, negative, or neutral. Data for this evaluation can be found in the Results folder of the appendix package.

5.3 Discussion

5.3.1 *In-Domain Results*

In Experiment 1, the decision tree-based classifier achieves the maximum performance in 4 out of the 7 performance metrics, beating or matching the baseline in 3 metrics. By achieving the minimum Chart RMSE, this classifier predicts an Artist 100 chart with minimum distance to the actual chart; see Table 5 for a comparison of the predicted and actual charts. Interestingly, all three of the models predict all of the artists which appear on the actual chart, but not in the correct order. The ensemble decision tree models predict with higher error than the single decision tree. It is also important to note that all models achieve over 99% pairwise classification accuracy, suggesting that extremely high accuracy is required to perform competitively, and small changes in predictions can lead to large changes in error. This is further supported by the fact that the classifier must make over 100,000 predictions to predict the Artist 100.

While the previous chart baseline achieves overall non-competitive results, it is interesting to see that it was best at predicting the top 10 of the Artist 100. This may be because there tends to be less chaotic movement in the top 10 positions compared to the lower positions, i.e., it is much more difficult to reach the top of the chart than the bottom, thus there are few artists that can, and less change occurs there. The baseline also had superior performance in the "matches only" metrics, however the baseline achieved fairly low Chart Overlap, so this should not be considered a general indicator of superior performance.

If we take a closer look at the decision tree-based model, we see that of the 59 rules of the depth-7 tree, the LastPosition field appears in 33, the WeeksOnChart field appears in 19, the DailyMentionsMult appears in 5, the DailyFollowersMult appears in 16, and the PositionTrend field appears in 2. This may be an indicator of how influential each field was in predicting the class, however this requires further investigation. The produced decision trees from each experiment are included in the Results folder of the appendix file.

5.3.2 *In-Domain vs. Out-of-Domain*

We note that from the in-domain to the out-of-domain experiments, error increases in our models as expected. A closer look at the shows that only 4 out-of-domain artists appear on the testing chart in the out-of-domain experiments. This is a demonstration of the sensitivity of the classifier performance when it misses just one or a few artists. However, our models still beat the baseline in Chart RMSE and Chart Overlap, potentially the most representative metrics of performance, and achieve high classification accuracy on the pool of artists.

Interestingly, the ensemble decision tree models did not suffer as great of a performance loss between the experiments as the decision tree model did, an indicator of the robustness of these models. The ensemble decision tree model with AdaBoost actually predicted all 96 in-domain artists on the chart, while the other two models predicted 95.

5.3.3 *Sentiment*

Though the results of Experiment 3 suggest that we cannot compete with a simple baseline using tweet sentiment in predicting the Artist 100, we cannot draw any definitive conclusions from this investigation. Our sentiment classifier could only predict with accuracy slightly better than a random guess, and worse than a majority class baseline (choosing neutral for all tweets leads to 51.3% accuracy, while our classifier achieved 44.0% accuracy). Consequently, it is improper to draw conclusions based on its chart predictions. However, classifying unfiltered tweets about musical artists is a difficult task. Many tweets included emojis which we were not capable of classifying, many were spam or from trolls, and some were in different languages. Slang and use of typically negative words as positive words further complicate the task.

Rank	Predicted	Actual	Rank	Predicted	Actual
1	Lady Gaga	Andrea Bocelli	51	DJ Snake	Grateful Dead
2	BTS	Lady Gaga	52	NF	Tyga
3	Drake	Drake	53	Thom Yorke	6ix9ine
4	Ariana Grande	Post Malone	54	Steve Aoki	Nicki Minaj
5	Halsey	Cardi B	55	Bazzi	Tory Lanez
6	Khalid	Imagine Dragons	56	Robyn	Florida Georgia Line
7	XXXTENTACION	Bradley Cooper	57	Nicki Minaj	NF
8	Post Malone	Juice WRLD	58	6ix9ine	Selena Gomez
9	Eminem	Halsey	59	Andrea Bocelli	Bazzi
10	Imagine Dragons	Ariana Grande	60	Gunna	Josh Turner
11	Travis Scott	XXXTENTACION	61	Carrie Underwood	Jason Aldean
12	Ella Mai	Travis Scott	62	Bobby "Boris" Pickett And The Crypt-Kickers	Bastille
13	Bradley Cooper	Joji	63	Joji	lovelytheband
14	Maroon 5	Khalid	64	Sheck Wes	Thomas Rhett
15	Disturbed	Eminem	65	Lukas Graham	Kendrick Lamar
16	5 Seconds Of Summer	Queen	66	Rihanna	Gunna
17	Luke Combs	BTS	67	Justin Bieber	Future
18	Lauren Daigle	Lauren Daigle	68	lovelytheband	DJ Snake
19	Juice WRLD	Ed Sheeran	69	Josh Turner	Eric Church
20	Greta Van Fleet	Maroon 5	70	Kanye West	Carrie Underwood
21	Ed Sheeran	Luke Combs	71	Maren Morris	Robyn
22	Panic! At The Disco	Panic! At The Disco	72	Normani	The Chainsmokers
23	Lil Wayne	Disturbed	73	Cher	Lauv
24	Queen	Ella Mai	74	Swae Lee	Jimmie Allen
25	Kane Brown	5 Seconds Of Summer	75	Jimmie Allen	Quavo
26	Bruno Mars	Lil Baby	76	John Legend	Brett Young
27	Cardi B	Michael Jackson	77	Tory Lanez	Swae Lee
28	Future	Shawn Mendes	78	NCT	Bobby "Boris" Pickett And The Crypt-Kickers
29	twenty one pilots	Lil Wayne	79	Brett Young	Steve Aoki
30	Jason Aldean	twenty one pilots	80	Grateful Dead	YoungBoy Never Broke Again
31	Lil Baby	John Legend	81	Lil Pump	NCT
32	Marshmello	Dan + Shay	82	Migos	Rihanna
33	Shawn Mendes	Marshmello	83	Pentatonix	Lukas Graham
34	Dua Lipa	Kane Brown	84	P!nk	Maren Morris
35	Michael Jackson	Bruno Mars	85	Lynyrd Skynyrd	Zedd
36	Bebe Rexha	Camila Cabello	86	Tyga	Lil Pump
37	Bastille	Kodak Black	87	The Chainsmokers	Lynyrd Skynyrd
38	Selena Gomez	Taylor Swift	88	Luke Bryan	Adele
39	Kendrick Lamar	Kanye West	89	Zedd	Metallica
40	Chris Stapleton	Chris Stapleton	90	Adele	Migos
41	Taylor Swift	Greta Van Fleet	91	EXO	The Weeknd
42	Eric Church	P!nk	92	YoungBoy Never Broke Again	J Balvin
43	Bad Bunny	Cher	93	Lauv	Normani
44	Florida Georgia Line	EXO	94	Mitchell Tenpenny	Justin Bieber
45	Thomas Rhett	Billie Eilish	95	Ozuna	Luke Bryan
46	Kodak Black	Pentatonix	96	Metallica	Mitchell Tenpenny
47	Camila Cabello	Bebe Rexha	97	J Balvin	Ozuna
48	Dan + Shay	Dua Lipa	98	Kenny Chesney	Offset
49	Quavo	Sheck Wes	99	The Weeknd	Kenny Chesney
50	Billie Eilish	Bad Bunny	100	Offset	Thom Yorke

Table 4: Artist 100 for November 10, 2018 predicted by decision tree model compared to actual Artist 100 for November 10, 2018.

6. CONCLUSIONS

From our results, we conclude that it is indeed possible to predict the Artist 100 to fairly high accuracy using only past charts and social media data rather than the highly multidimensional blend of metrics used in generating the Artist 100 every week.

This conclusion has two consequences. First, as sales data was completely excluded from our experiments, we have shown that social media engagement is indeed highly crucial in reaching high positions on the Artist 100 chart, an indicator of an artist's general popularity and success. The second consequence is that the activity levels on all social media platforms for a particular artist are correlated to each other, as we made predictions using only Twitter rather than several platforms like the Artist 100 ranking algorithm does.

We hope that our results from this simple investigation encourage future research on the Artist 100. Future work on this particular problem may include a deeper investigation into the utility of sentiment analysis. API call and storage limitations prevented us from implementing sentiment analysis metrics throughout our data and, thus preventing us from augmenting our classifier with them. It is possible that better results may be achieved in a larger scale investigation.

As Billboard charts are reflective of artist popularity in the US, it may also be important to consider the country in which social media attention comes from, data we do not currently have access to. It appears that the Artist 100 draws from global social media data, but US sales are likely better represented by only US social media data.

Further, our decision tree approaches appeared to use a more brute-force method on each individual feature by comparing to different numbers over several levels. Improvements may come from adding features to our concatenated artist pair examples which directly compare the features of each individual artist's feature vector. For example, for each feature in the individual artist vectors, we could add the ratio of a feature from artist 1 to the corresponding feature of artist 2 as attributes of the concatenated vectors.

Lastly, it would obviously be more useful to artists to be able to attempt to predict the Artist 100 more than one week in advance, especially since the delay of chart releases from tracking periods allows for predictions at least two weeks in advance.

Overall, though, the current performance of our models is a sign that we have reached a time where social media engagement is indeed necessary, or at minimum incredibly advantageous to the success of artists, and thus satisfies our goals for this work. We hope that further research will shed more light on the optimal ways to use social media to increase popularity and profits.

7. REFERENCES

[1] C. Morris, "Rock music's new heroes: Lady Gaga and...big data?," *CNBC*, 11-Oct-2013. [Online]. Available: https://www.cnbc.com/2013/10/10/-new-heroes-lady-gaga-andbig-data.html. [Accessed: 26-Nov-2018].

[2] Next Big Sound, "About," *Next Big Sound*. [Online]. Available: https://www.nextbigsound.com/. [Accessed: 26-Nov-2018].

[3] G. Trust, "2014: The Year in... Billboard Chart Headlines," *Billboard*, 09-Dec-2014. [Online]. Available: https://www.billboard.com/articles/events/year-in-music-2014/6386057/2014-billboard-chart-headlines. [Accessed: 26-Nov-2018].

[4] G. Trust, "The Hot 100," *Billboard*, 2018. [Online]. Available: https://www.billboard.com/charts/hot-100. [Accessed: 26-Nov-2018].

[5] E. T. Bradlow and P. S. Fader, "A Bayesian Lifetime Model for the 'Hot 100' Billboard Songs," *J. Am. Stat. Assoc.*, vol. 96, no. 454, pp. 368–381, Jun. 2001.

[6] J. Grace, D. Gruhl, K. Haas, M. Nagarajan, C. Robson, and N. Sahoo, "Artist Ranking Through Analysis of On-line Community Comments," *Knoesis Publ.*, Apr. 2008.

[7] N. Koenigstein and Y. Shavitt, "SONG RANKING BASED ON PIRACY IN PEER-TO-PEER NETWORKS," *Poster Sess.*, p. 6, 2009.

[8] N. Koenigstein, Y. Shavitt, and N. Zilberman, "Predicting Billboard Success Using Data-Mining in P2P Networks - IEEE Conference Publication," 2009. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/5366058. [Accessed: 26-Nov-2018].

[9] K. Bischoff, C. S. Firan, M. Georgescu, W. Nejdl, and R. Paiu, "Social Knowledge-Driven Music Hit Prediction," in *Advanced Data Mining and Applications*, 2009, pp. 43–54.

[10] S. Goel, J. M. Hofman, S. Lahaie, D. M. Pennock, and D. J. Watts, "Predicting consumer behavior with Web search," *Proc. Natl. Acad. Sci.*, Sep. 2010.

[11] Y. Freund and R. E. Schapire, "A Short Introduction to Boosting," p. 14, 1999.

[12] Billboard Staff, "Billboard to Alter Chart Tracking Week for Global Release Date," *Billboard*, 24-Jun-2015. [Online]. Available: https://www.billboard.com/articles/columns/chart-beat/6605842/billboard-alter-chart-tracking-week-global-release-date. [Accessed: 02-Dec-2018].