CAN MICROBLOGS PREDICT MUSIC CHARTS? AN ANALYSIS OF THE RELATIONSHIP BETWEEN #NOWPLAYING TWEETS AND MUSIC CHARTS

Eva Zangerle, Martin Pichl, Benedikt Hupfauf, Günther Specht
Department of Computer Science
University of Innsbruck, Austria
{eva.zangerle, martin.pichl, benedikt.hupfauf, guenther.specht}@uibk.ac.at

ABSTRACT

Twitter is one of the leading social media platforms, where hundreds of millions of tweets cover a wide range of topics, including the music a user is listening to. Such #nowplaying tweets may serve as an indicator for future charts, however, this has not been thoroughly studied yet. Therefore, we investigate to which extent such tweets correlate with the Billboard Hot 100 charts and whether they allow for music charts prediction. The analysis is based on #nowplaying tweets and the Billboard charts of the years 2014 and 2015. We analyze three different aspects in regards to the time series representing #nowplaying tweets and the Billboard charts: (i) the correlation of Twitter and the Billboard charts, (ii) the temporal relation between those two and (iii) the prediction performance in regards to charts positions of tracks. We find that while there is a mild correlation between tweets and the charts, there is a temporal lag between these two time series for 90% of all tracks. As for the predictive power of Twitter, we find that incorporating Twitter information in a multivariate model results in a significant decrease of both the mean RMSE as well as the variance of rank predictions.

1. INTRODUCTION

The microblogging platform Twitter has long since become one of the leading social media platforms, serving 271 million active users, who publish approximately 500 million tweets every day [3]. On Twitter, users for instance share their opinions about various topics, post interesting articles, let the world know what they are currently doing, and interact with other users. Furthermore, users share what music they are currently listening to in so-called #nowplaying tweets. These tweets can be identified by one of the hashtags #nowplaying, #listento or #listeningto, and typically feature title and artist of the track the user is listening to (e.g. "#NowPlaying Everlong – Foo Fighters

http://spoti.fi/J1Gqhs"). The majority of such tweets are automatically generated by music players or music streaming platforms (as in the previous example tweet, Spotify).

#nowplaying tweets have already been analyzed in regards to the listening behavior of users from around the world [20], or to perform user-specific music recommendation tasks [24]. Also, Schedl and Tkalčič looked into the genre distribution within #nowplaying tweets, in particular, the use of social media of classical music enthusiasts [21]. Kim et al. present a prediction analysis of Billboard charts based on Twitter data [13]. The authors crawled #nowplaying tweets and the Billboard Hot 100 over the course of 10 weeks and extracted tweets that contain a song or artist already contained in the charts. The actual prediction was then performed by a classification task based on the number of tweets about the given track and artist, and the number of weeks the track has already been in the Billboard Hot 100. The results indicate that the future success of a track, which is already in the charts, can be predicted accurately. However, we argue that only using those tweets which feature artists or songs currently in the Billboard Hot 100, and using information about how long a certain song has already been in the charts limits this approach and its general applicability. Furthermore, the study by Kim et al. was limited to an analysis period of 10 weeks.

Following up on this research, we are interested in generalizing the hypothesis of chart prediction based on Twitter data, and look into how suitable #nowplaying tweets are when it comes to predicting the Billboard Hot 100. Therefore, we shed light on the following research questions (RQ) in this study:

- **RQ1:** To which extent do #nowplaying-tweets resemble the Billboard Hot 100?

- **RQ2:** How are #nowplaying tweets and the Billboard Hot 100 temporally related?

- **RQ3:** How can Twitter data be exploited for predicting music charts?

In this work, we model the Billboard charts data as well as the Twitter data as time series [10]. Based on this data, we look into three different aspects to answer the research questions: (i) the correlation of Twitter-activity regarding

musical tracks and the Billboard charts, (ii) the temporal relation of tweets and charts to investigate whether there is a timely offset between Twitter and the Billboard Hot 100 charts and (iii) the predictive power of #nowplaying-tweets with respect to the Billboard Hot 100.

The contribution of this work can be summarized as follows: To the best of our knowledge, this is the first paper providing a deep analysis on the relationship between #nowplaying tweets and charts. Furthermore, our study is based on data collected over two years, covering a time frame substantially longer than previous research and particularly, utilizing time series to perform the analyses. We find that the basic correlation of time series representing the Billboard Hot 100 and Twitter performance of all individual tracks is moderate. When performing a cross-correlation analysis to compute a time-agnostic correlation analysis, the correlation coefficient is 0.57. Analyzing the lag between the two time series shows that the temporal offset between musical tweets and charts would allow for a prediction from a temporal perspective for 41% of all tracks. Our prediction experiments show that Twitter data enhances the quality of charts rank predictions as the multivariate model incorporating both Billboard and Twitter data reduces the prediction error significantly and further also reduces the error variance significantly.

The remainder of this paper is structured as follows. Section 2 presents background information and approaches related to our analyses. Section 3 introduces the dataset underlying our analyses, and Section 4 describes the analyses methods we facilitate. Section 5 presents the results of our analysis, and Section 6 discusses our findings in the light of the posed research question. Section 7 concludes the paper.

2. BACKGROUND AND RELATED WORK

Research related to the analyses presented in this paper can be categorized into the following categories: (i) charts and hit prediction based on Twitter data, (ii) analyses of musical tweets and (iii) time series analyses in social media.

Kim et al. [13] present the first approach to predict the success of songs in terms of their Billboard Hot 100 rankings based on Twitter data. Therefore, the authors facilitate a dataset comprising 10 weeks of #nowplaying tweets, and the Billboard Hot 100 of the same time period. The authors use three different features for their further computations: a song's popularity on Twitter, an artist's popularity on Twitter, and the number of weeks a song was in the Billboard Hot 100. Based on these features, the authors compute the Pearson correlation coefficient between the chart-ranking of a particular song in the Billboard Hot 100, and each of the aforementioned features. They find that the song popularity obtains the highest correlation with the ranking of the given songs. In a second step, the authors build different regression models (linear, quadratic linear, and support vector regression models) to predict the ranking of a certain song. The authors find that using all three features with a support vector regression model produces the best prediction performance (r^2=0.75). As for detect-

ing whether a certain song will be a hit, the authors divide their dataset into hits and non-hits and try to predict whether a random non-hit song will be a hit by using random forest classification. The results show that hits (chart rank 1-10) can be predicted with a precision value of 0.92 and a recall value of 0.88. However, these prediction tasks are only performed for songs which are already in the Billboard Hot 100. Incorporating the number of weeks a song was in the charts requires knowledge of the charts and as a consequence, such a method cannot be applied to new songs which are possibly about to enter the charts.

Generally, #nowplaying tweets have been in the focus of researchers aiming to detect musical preference patterns from around the world. However, except for Kim et al., none of these approaches aim to predict charts based on this data. Hauger and Schedl extract genre patterns for regions of the world based on geolocation information of #nowplaying tweets [20]. Schedl et al. also analyse the listening behavior of Twitter users with a particular focus on geospatial aspects [19]. Following up on this research, Schedl and Tkalčič looked into the general genre distribution among tweets, with an emphasis on classical music [21]. Furthermore, Pichl et al. facilitate #nowplaying tweets from the streaming platform Spotify to recommend artists to users [16].

Social media data has been modeled as time series for a variety of analyses, e.g., for predicting the stock market based on emotion within social media [7]. Similarly, time series have been used for modeling mood information extracted from Twitter [6] or the detection of influenza epidemics via Twitter [8]. Further, Sakaki et al. facilitate time series for real-time event detection [18]. Huang et al. model the user of tags on Twitter as a time series to understand how hashtags are used on Twitter [12]. Also, Twitter models data as time series to monitor the health and success of their services as well as to detect anomalies in regards to spikes in user attention [4, 5].

3. DATA

In the following, we describe the data used for the performed analyses. In principle, we require data from two different sources, gathered over the same time frame: #nowplaying tweets, and information about the charts at the given time. As for the set of musical tweets, we choose to base our analyses on a dataset the #nowplaying dataset provided by Zangerle et al. [25] as it is the most extensive dataset of #nowplaying tweets publicly available, which is constantly updated. In total, we utilize all of the 111,260,925 #nowplaying tweets that are available for the years 2014 and 2015. These tweets form the basis of our analysis. To evaluate to what extent Twitter data can be used to predict the charts, we choose to use the Billboard Hot 100 as a reference, as they are one of the most influential indicators for the popularity of songs [2]. Analogously to the work by Schedl and Tkalčič [21], we are aware that the Billboard charts reflect the U.S. market only. Within the Billboard charts, the songs are ranked according to a score computed based on radio airplay, sales and streaming activity [1].

We crawled the top 100 of each week of the years 2014 and 2015 from the Billboard Website [2]. In total, we gathered 886 distinct songs. On average, a track stays within the Hot 100 for 11.74 weeks (SD=10.58), the minimum number of weeks for a track within the Hot 100 is one week and the maximum is 58 weeks.

4. METHODS

This section describes the methods implemented in our study. We propose three analyses to investigate whether the prediction of charts is possible (i) from a correlation perspective, (ii) from a temporal perspective with respect to the distribution and popularity of songs on Twitter and in the charts, and (iii) a prediction perspective analyzing to which extent Twitter data can contribute to predicting future charts for given songs.

To be able to compare #nowplaying tweets and the Billboard Hot 100, we quantify the overlap of musical tweets and charts by counting how many tweets refer to a song on the charts. We aim to match tweets against all tracks that have been on the charts since 2011 to ensure a maximum overlap. We iterate over all tweets and all songs that have been in the charts in nested loops and match title and artist independently. The track title is considered successfully matched, if it is a substring of the tweet (case insensitive). Matching the artist is more complex, as there are several formats, for instance "Michael Jackson" and "Jackson, Michael", in use. Moreover, there are many ways to list featured artists: "Bad Meets Evil Featuring Bruno Mars – Lighters", "Bad Meets Evil – Lighters Featuring Bruno Mars", or sometimes, featured artists are simply neglected. For this reason, the artist string is split at keywords and symbols, such as "feat.", "ft.", "&", etc. If more than half of the resulting tokens can be found in the tweet (again, case insensitive), the artist counts as successfully matched. We consider a track matched, if both track title and artist were matched successfully.

Based on the gathered and preprocessed data on both Twitter and the Billboard Hot 100, we model both of these as time series [10] to compute the analyses as described in the following.

4.1 Correlation of Rankings

In a first step we aim to analyze to which extent Twitter data and Billboard Hot 100 data correlate. We perform a detailed correlation analysis of the popularity of each track in both the Billboard and the #nowplaying dataset. Kim et al. provide a correlation analysis (using Pearson correlation) for (i) the log of the track's playcount and (ii) the log of the artist's playcount and (iii) the number of weeks the song was in the Billboard Hot 100. The authors define the playcount as the median number of mentions of the respective track or artist per day for a given week. We extend this analysis by aggregating the playcounts for a given week not only by using the median of the playcounts of the days of a week, but also computing the mean number of playcounts per track per day and the sum of play-

counts for a given week. If a track was not on the Billboard charts in a given week, we consider its rank as 0. Based on this extracted information, we compute the Pearson correlation coefficient for the track and artist playcounts and the Billboard time series [15]. As for the log of the track's playcount and log of artist's playcount, we rely on Spearman's ρ [22] due to the data's ordinal scale and the fact that the logarithm is a non-monotonic transformation of the data [11].

4.2 Temporal Relationship of Tweets and Charts

As we aim to analyze to which extent Twitter data resembles trends in the Billboard Hot 100 and hence, allows for a prediction of future Billboard charts, we are naturally interested in the temporal relationship between Twitter and the Billboard charts. I.e., we analyze whether Twitter does—from a timely perspective—represent trending tracks earlier than those are reflected in the Billboard Hot 100.

To investigate this matter, we propose to perform a cross-correlation analysis of the time series based on Twitter and the time series based on Billboard data for any given music track. In principle, cross-correlation is used to compute the similarity of two signals or time series as a function of the lag between these two signals [14]. This allows for a correlation analysis independent of temporal shifts between the time series while at the same time obtaining information about the lag between the time series. Along the lines of previous research we consider a time lag optimal if it maximizes the correlation of the two signals [23]. This allows to compute the maximum correlation between two time series independent of timely shifts (i.e., Twitter data and Billboard data) and determine the lag between these time series. I.e., if we determine a negative lag for a given track, this implies that Twitter data would allow to predict future charts from a timely perspective.

For the computation of the cross-correlation we rely on the median number of mentions of each track per week on Twitter as this measure has shown to provide the highest correlation value with the Billboard charts (cf. Section 5.1).

4.3 Prediction of Billboard Charts

For the prediction of future Billboard charts, we aim to forecast the performance of a given track in regards to its rank in the future based on past and present observations. Mostly, a regression approach facilitating univariate or multivariate time series is applied for such tasks [9]. Therefore, we propose to evaluate the quality of predictions based on Billboard charts only, Twitter charts only and a combined approach. In particular, we propose to compute three prediction models to evaluate the predictive power of the individual approaches: (i) an autoregressive time series model (AR) based on solely the Billboard time series (we utilize this method as a baseline), (ii) extract the lag from the cross-correlation analysis, shift the time base and compute an AR-model based on the difference between the Twitter and Billboard time series to be able to evaluate the Twitter time series in regards to predicting

Measure	Aggregation Method for Playcounts		
	Median	**Mean**	**Sum**
track playcount	0.50 (481)	0.49 (469)	0.49 (453)
log(track playcount)	0.49 (457)	0.48 (453)	0.47 (442)
artist playcount	0.37 (378)	0.37 (364)	0.37 (358)
log(artist playcount)	0.40 (398)	0.38 (389)	0.38 (389)

Table 1: Mean Correlation Coefficients ($p < 0.01$); Numbers in Parenthesis: Number of Significantly Correlating Tracks

the Billboard charts and (iii) utilize information from both Billboard and Twitter to perform a combined prediction approach using a vector autoregressive model for multivariate time series (VAR). Such a VAR model computes predictions based on multiple time series, in our case, Billboard data and Twitter data (particularly, the median number of tweets about each track). As for the training of the models, we rely on the Ordinary Least Squares (OLS) method [9]. We evaluate these three models in regards to their rank prediction accuracy using the standard evaluation measure root mean squared error (RMSE) [9].

5. RESULTS

In the following section we present the results of our analyses in the light of the research questions posed. We firstly present the results of the correlation analysis and subsequently provide the results of the cross-correlation analysis to look into the temporal relationship between Twitter and Billboard. Based on these findings, we present the results of the prediction analysis for Billboard charts.

5.1 Correlation of Tweets and Charts

Firstly, we analyze the correlation of the measures describing the performance of each track extracted from the #now-playing dataset (cf. Section 4.1). Therefore, we compute the correlation between the mean, median and the sum of playcounts of the track and its rank within the Billboard Hot 100. The results of this analysis can be seen in Table 1, where we list the mean correlation coefficient across all tracks within the dataset. Please note that we only consider those tracks for which we can find a significant correlation ($p < .01$). As can be seen, we observe the highest correlation values for the track playcount with a moderate correlation of 0.50 for 481 of 886 tracks (54.29% of all tracks in the dataset). Figure 1a shows the correlation distribution for all tracks with a significant correlation at the $p < 0.01$-level. We observe both positively correlated and negatively correlated tracks. We also observe that—in line with the findings by Kim et al. [13]—using the median value of all track counts of a given week to represent the performance of any given track on Twitter shows the highest correlation values throughout all configurations. Also, the level of correlation obtained is in line with those observed by Kim et al. as those reach a correlation value of 0.56 for the log of the track playcount [13]. To answer RQ1, we find that there is a moderate correlation between #nowplaying playcount numbers and Billboard charts.

Performing an explorative study on the time series representations of given tracks in the Twitter dataset as well as the Billboard dataset shows a timely lag between these two signals, which we investigate further in the next experiment.

5.2 Temporal Relationship of Tweets and Charts

To gain a deeper understanding for the temporal relationship between tweets and charts, we perform a cross-correlation analysis of the respective time series. As described in Section 4.2, this analysis allows for determining the lag between two time series. Figure 2a depicts the distribution of lags detected for all tracks in the dataset. We observe that the lag histogram shows its highest peaks at -1 and 1 weeks, implying that a substantial number of tracks are mentioned and trending on Twitter either one week before or after these occur and evolve on the Billboard Hot 100 charts. At the same time, the histogram also shows that there also is a substantial number of tracks with a positive time lag. I.e., maximum correlation is reached when shifting the Twitter signal to a later point in time. In total, 286 tracks feature a negative lag (41.09%), whereas 335 tracks (48.13%) feature a positive lag and 75 tracks (10.77%) feature no lag. Table 2 shows a five-number summary of the distribution of time lags (cf. row "all"). On average, the lag between Twitter and the Billboard charts is positive (1.47), whereas the median value is 0.

	Min	**Q1**	**Med**	**Mean**	**Q3**	**Max**
All	-17.0	-2.0	0.0	1.47	5.0	17.0
TF	-17.0	-2.0	0.0	0.97	4.0	17.0

Table 2: Five-Number Summary: Temporal Lag (TF refers to those tracks that first occur on Twitter)

We can now utilize the computed lag to shift the base of the time series such that they are maximally correlated. This cross-correlation analysis shows that the mean correlation for all tracks is now 0.57 in contrast to 0.50. The correlation distribution of the playcount measure after the base shift is shown in Figure 1b, where the improvement is clearly visible as the distribution is now shifted towards the right, now only containing positive correlation coefficients in contrast to Figure 1a, where we still observe negative correlation coefficients. The median correlation coefficient is slightly increased from 0.50 to 0.57 when computing the cross-correlation coefficient for all tracks.
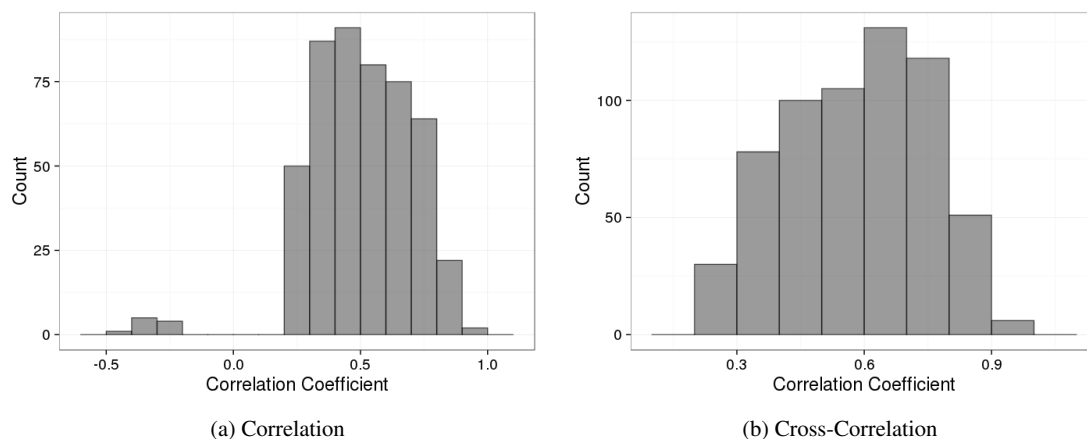
(a) Correlation

(b) Cross-Correlation

Figure 1: Histogram of Correlation Coefficients for Track Playcounts

In a second step, we repeat the experiment based on the set of tracks which first occur on Twitter before they actually appear on the Billboard Hot 100. I.e., we extract all tracks for which we observe that they first appear on Twitter and at a later point in time on the charts. This results in a total of 619 tracks. This experiment aims to look into tweets, which would actually allow for a prediction of Billboard charts as these tracks are featured and trending on Twitter *before* they actually appear in the Billboard Hot 100 charts. Figure 2b shows the lags resulting from a cross-correlation analysis of these tracks. As can be seen, the lag distribution is slightly shifted towards the left, i.e., this subset of tweets features lower lags. Table 2 shows the five-number summary of this distribution (cf. row "TF" for Twitter First). Directly comparing the distribution to the lag distribution of all tracks shows that the mean lag is lower, reaching 0.97 weeks. For 69 tracks (11.14%) we do not observe a lag between Twitter and Billboard. 286 tracks (46.20%) feature a positive lag, whereas 264 tracks (42.64%) feature a negative lag. I.e., for 264 tracks a prediction based on Twitter charts seems possible from a temporal perspective.

To answer RQ2, we observe that 89.23% of all tracks actually feature a temporal lag and that 41.09% of all tracks feature a negative lag. I.e., a prediction of charts based on Twitter data is possible from a temporal perspective. When shifting the base of the time series according to the lag, the correlation is increased to 0.57. Looking at the subset of tracks which appear on Twitter first, we observe a negative time lag for 264 tracks, which accounts for 42.64%, which would allow for a prediction.

5.3 Prediction of Charts

Based on the findings of the previous analyses, we aim to investigate the usefulness of Twitter data for the prediction of Billboard charts in the following. Therefore, we now present the results of a autoregression approach to predicting future Billboard charts. We applied the Augmented Dickey-Fuller test to confirm stationarity in the time series [17].

Table 3 shows the 5-point-summary of the results of the three proposed prediction models in terms of the RMSE values obtained: The Billboard-based autoregressive model (BB), the Twitter-based AR model (T) and the multivariate model combining Twitter and Billboard data (V). Please note that outliers are omitted (we consider all tracks that are more distant than 1.5 interquartile ranges from the upper or lower quartile as outliers). As can be seen, the autoregressive approach based only on Twitter data works substantially worse than the other two approaches, the median of the RMSE distribution being 116.1. In contrast, autoregressive prediction based on the Billboards model (and hence, the baseline) reaches a median RMSE of 26.8 and the VAR model reaches the lowest median RMSE with 12.6. Figure 3 shows a boxplot of the RMSE of the VAR and the Billboard autoregressive model. As can be seen, the VAR model also features a lower variance within the predicted ranks. The mean RMSE—indicated by a diamond in the boxplot—is also clearly lower for the VAR model (14.1 vs. 26.8, hence, a 48.38% lower value). Due to the non-normal distribution of the data, we apply a Mann-Whitney U test to test for significant differences in the prediction performance of the VAR and Billboard AR model and the result shows a significantly lower RMSE for the VAR-model ($p < 0.05$). A Levene's test of equality of RMSE variance shows that the VAR model reaches significantly lower error variance in terms of ranking predictions than the Billboard AR model ($p < 0.01$).

Regarding RQ3, we can therefore observe that combining Twitter and Billboard data enhances the quality of predictions. We show that the RMSE is significantly lower when incorporating Twitter information. Similarly, we show that the error variance of predictions is significantly decreased using a multivariate model.

6. DISCUSSION

In the following section, we discuss the results presented in the previous section, and put them into context in order to confirm or disprove our hypothesis that Twitter musical data can be used for predicting music charts.
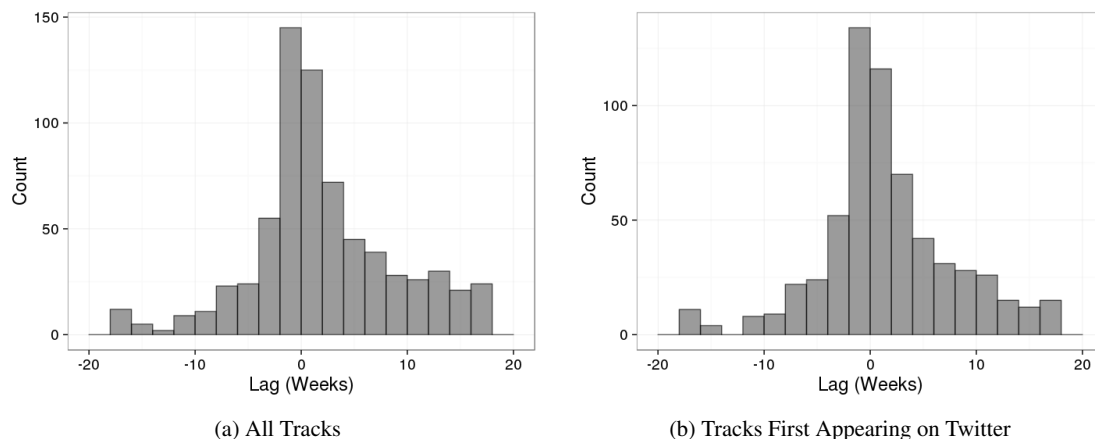
(a) All Tracks   (b) Tracks First Appearing on Twitter

Figure 2: Distribution of Temporal Lags between Twitter and Billboard Hot 100
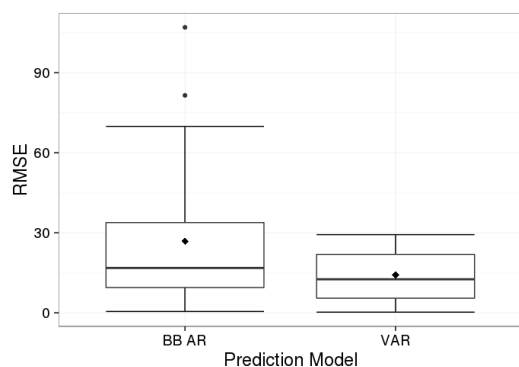

Figure 3: Boxplot of RMSE of Prediction Models (Diamond: Mean Value)

As for the correlation between the time series of the Billboard Hot 100 ranks of tracks and the time series of the tweet playcounts, we observe moderate correlation. In line with the findings by Kim et al. [13], who observe a correlation coefficient of 0.56 between the Billboard Hot 100 and the logarithm of the track playcounts, our dataset features a maximum correlation between Billboard and Twitter of 0.50. Our dataset represents a long-term view on the relationship between Twitter and the Billboard charts as we performed the analyses over the course of two years, whereas Kim et al. observed a window of 10 weeks.

Based on the performed evaluation, we find that from a temporal perspective, approximately 41% of all #nowplaying tweets can be used to predict music charts as these appear on Twitter before they actually appear on the charts.

	Min	Q1	Med	Mean	Q3	Max
T	16.3	84.5	116.1	148.6	178.2	388.4
BB	0.51	9.5	16.8	26.8	33.8	107.0
V	0.27	5.5	12.6	14.1	21.9	29.3

Table 3: Five-Number Summary: Charts Prediction RMSE (T: Twitter AR, BB: Billboard AR, V: VAR model)

However, the lag between the two time series is rather low and more importantly, the mean lag across all tracks is positive. Therefore, we argue that approaches exploiting this temporal shift for predictions does not seem promising. However, using Twitter data to enhance chart predictions based on Billboard data has shown to provide promising results. Our evaluation of autoregressive prediction approaches shows that incorporating Twitter data in the prediction process using a multivariate model significantly lowers the RMSE as well as the variance of the RMSE. I.e., we are able to predict the rank of tracks more accurately while at the same time providing a lower error margin for predictions. Hence, we come to the conclusion that #nowplaying data is able to contribute to better charts prediction and can be utilized as an additional sensor which allows for enhancing chart predictions. However, we have to note that solely relying on Twitter data for charts prediction has shown to be highly error-prone and performing substantially worse than both the Billboard AR and the VAR model presented.

Another limiting factor we are aware of is the demographic gap between the user base of Twitter and the average consumer in the U.S. music market (represented by the Billboard Hot 100 charts).

7. CONCLUSION

Based on a dataset gathered from Twitter and the Billboard charts over the course of 2014 and 2015, we analyzed the relationship between Twitter and the Billboard charts in regards to whether tweets could be utilized for predicting future Billboard charts. Therefore, we performed a three-fold analysis of the dataset. These experiments showed that in principle, Twitter and Billboard time series for tracks share a moderate correlation which is influenced by a timely shift between those two. We further find that there is a negative timely lag for 41% of all tracks. As for the predictive power of #nowplaying charts we find that a multivariate model incorporating both Billboard and Twitter data significantly reduces the prediction error while at the same time, the error variance is significantly reduced.

8. REFERENCES

[1] Billboard Charts Legend: `http://www.billboard.com/biz/billboard-charts\-legend`, 2016. (last visited: 2016-02-25).

[2] Billboard Hot 100: `http://www.billboard.com/charts/hot-100`, 2016. (last visited: 2016-02-25).

[3] Twitter: About Twitter `https://about.twitter.com/company`, 2016. (last visited: 2016-02-28).

[4] Twitter Blog: Introducing practical and robust anomaly detection in a time series `https://blog.twitter.com/2015/introducing\-practical-and-robust-anomaly-detect\-ion-in-a-time-series`, 2016. (last visited: 2016-03-01).

[5] Twitter Blog: Observability at Twitter `https://blog.twitter.com/2013/observability-at-twitter`, 2016. (last visited: 2016-03-01).

[6] Johan Bollen, Huina Mao, and Alberto Pepe. Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. In *Proceedings of the Fifth International Conference on Weblogs and Social Media, Barcelona, Catalonia, Spain, July 17-21, 2011*, volume 11, pages 450–453. The AAAI Press, 2011.

[7] Johan Bollen, Huina Mao, and Xiaojun Zeng. Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1):1 – 8, 2011.

[8] David A Broniatowski, Michael J Paul, and Mark Dredze. National and local influenza surveillance through twitter: an analysis of the 2012-2013 influenza epidemic. *PloS one*, 8(12):e83672, 2013.

[9] Chris Chatfield. *Time-series forecasting*. CRC Press, 2000.

[10] James Douglas Hamilton. *Time series analysis*, volume 2. Princeton university press Princeton, 1994.

[11] Jan Hauke and Kossowski Tomasz. Comparison of values of pearson's and spearman's correlation coefficients on the same sets of data. *Quaestiones Geographicae*, 30(2):87–93, 2011.

[12] Jeff Huang, Katherine M. Thornton, and Efthimis N. Efthimiadis. Conversational tagging in twitter. In *Proc. of the 21st ACM Conference on Hypertext and Hypermedia*, HT '10, pages 173–178, New York, NY, USA, 2010. ACM.

[13] Yekyung Kim, Bongwon Suh, and Kyogu Lee. #Nowplaying the Future Billboard: Mining Music Listening Behaviors of Twitter Users for Hit Song Prediction. In *Proc. of the First International Workshop on Social Media Retrieval and Analysis*, SoMeRA '14, pages 51–56, New York, NY, USA, 2014. ACM.

[14] Sophocles J. Orfanidis. *Introduction to Signal Processing*. Prentice-Hall, Inc., 1995.

[15] William R Pearson and David J Lipman. Improved tools for biological sequence comparison. *Proc. of the National Academy of Sciences*, 85(8):2444–2448, 1988.

[16] Martin Pichl, Eva Zangerle, and Günther Specht. Combining Spotify and Twitter Data for Generating a Recent and Public Dataset for Music Recommendation. In *Proc. of the 26nd Workshop Grundlagen von Datenbanken, Ritten, Italy*, 2014.

[17] Said E Said and David A Dickey. Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, 71(3):599–607, 1984.

[18] Takeshi Sakaki, Makoto Okazaki, and Yutaka Matsuo. Earthquake shakes twitter users: Real-time event detection by social sensors. In *Proc. of the 19th International Conference on World Wide Web*, WWW '10, pages 851–860, New York, NY, USA, 2010. ACM.

[19] Markus Schedl. Leveraging Microblogs for Spatiotemporal Music Information Retrieval. In *Advances in Information Retrieval - 35th European Conference on IR Research, ECIR 2013, Moscow, Russia, Proceedings*, pages 796–799. Springer, 2013.

[20] Markus Schedl and David Hauger. Mining Microblogs to Infer Music Artist Similarity and Cultural Listening Patterns. In *Proc. of the 21st WWW: 4th International Workshop on Advances in Music Information Research*, Lyon, France, 2012.

[21] Markus Schedl and Marko Tkalčič. Genre-based Analysis of Social Media Data on Music Listening Behavior. In *Proc. of the 1st ACM International Workshop on Internet-Scale Multimedia Management*, ISMM '14, pages 9–14, New York, NY, USA, 2014. ACM.

[22] Charles Spearman. The proof and measurement of association between two things. *The American journal of psychology*, 15(1):72–101, 1904.

[23] Mikalai Tsytsarau, Themis Palpanas, and Malu Castellanos. Dynamics of news events and social media reaction. In *Proc. of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 901–910. ACM, 2014.

[24] Eva Zangerle, Wolfgang Gassler, and Günther Specht. Exploiting Twitter's Collective Knowledge for Music Recommendations. In *Proceedings, 2nd Workshop on Making Sense of Microposts: Big Things come in Small Packages, Lyon, France*, pages 14–17, 2012.

[25] Eva Zangerle, Martin Pichl, Wolfgang Gassler, and Günther Specht. #nowplaying Music Dataset: Extracting Listening Behavior from Twitter. In *Proc. of the 1st ACM International Workshop on Internet-Scale Multimedia Management*, ISMM '14, pages 3–8, New York, NY, USA, 2014. ACM.