Using Twitter to Predict Chart Position for Songs

Eleana Tsiara, Christos Tjortjis[0000-0001-8263-9024]

The Data Mining and Analytics research group, School of Science & Technology,
International Hellenic University, 14th km Thessaloniki–Moudania, 57001, Thermi, Greece
c.tjortjis@ihu.edu.gr

Abstract. With the advent of social media, concepts such as forecasting and nowcasting became part of the public debate. Past successes include predicting election results, stock prices and forecasting events or behaviors. This work aims at using Twitter data, related to songs and artists that appeared on the top 10 of the Billboard Hot 100 charts, performing sentiment analysis on the collected tweets, to predict the charts in the future. Our goal was to investigate the relation between the number of mentions of a song and its artist, as well as the semantic orientation of the relevant posts and its performance on the subsequent chart. The problem was approached via regression analysis, which estimated the difference between the actual and predicted positions and moderated results. We also focused on forecasting chart ranges, namely the top 5, 10 and 20. Given the accuracy and F-score achieved compared to previous research, our findings are deemed satisfactory, especially in predicting the top 20.

Keywords: Social Media Analytics, Prediction, Classification, Regression.

1   Introduction

Social media have penetrated everyday life to the point that they constitute an integral part of our daily routines. The vast amount of data available through these services can be utilized in many domains including finance, marketing, and politics [8]. Social media can be exploited in a variety of cases, such as forecasting the commercial success of movies [1], election result predictions etc. [7], [9].

In this paper, we chose Twitter to generate predictions for the Billboard chart. After gathering chart data, including titles, artist names and rankings, as well as tweets related to the top 10 songs for each week, results showed a moderate correlation between the number of mentions of a song and its future performance, but no relation between the number of mentions of an artist and their imminent success.

This work has the following objectives: a) to acquire data from the Billboard chart, including rank, artist and song title of the top 100 songs at the current time. For that purpose we developed a method which extracts these parameters from the official site and saves them in a .json and a .csv file. b) to collect Twitter posts concerning the top 10 songs utilizing the Twitter Search API. c) to preprocess data into a homogenous, structured format removing redundant information. d) to perform sentiment analysis on tweets, categorizing each post as positive, negative or neutral. e) to assess the contribution of the features, extracted from the Billboard chart and collected posts, to forecast the chart for the week to come.

The aim is to predict the top N songs for the following week and evaluate the efficiency of the process. This requires a classifier to generate predictions. The optimal number of N, the highest performing classification algorithm and the best feature combination can be determined by experimentation and result comparison.

The remaining of this paper reviews background in section 2, describes our approach in section 3, discusses results in section 4 and concludes with directions for future work in section 5.

2   Background

This section reviews forecasting related literature. Twitter was employed to predict the commercial success of movies [1]. The choice of film was based on the number of discussions and the difference of opinion, as well as by obtaining financial information about it. After collecting about 3 million tweets, the authors used a linear regression model and performed sentiment analysis to conclude that there is a correlation between the fame of a movie prior to its release and the revenue it produces.

Sales prediction was targeted in [3]. The aim was to estimate the impact of each variable such as posts, comments and likes on a group of Nike's Facebook page, examining each variable and each page individually, as well as a combination of all variables and all pages. Moreover, the predictive impact of search query data and the relation between Nike's events and the subsequent Facebook activity was explored. Simple regression scored as high as Bloomberg forecasts in terms of accuracy for predictions pertaining to the near future.

Another study that used Twitter data aimed at predicting the last presidential elections in the USA [7]. Its first objective was to retrieve data from Twitter, while organizing the timing of data collection according to important dates, such as debates. It then used sentiment analysis on 277,509 tweets. Each tweet was given a polarity and subjectivity score, depending on being negative, positive or neutral. Naïve Bayes was used for classification. The results were accurate and in fact predicted the right candidate, in contrast to most polls.

With regards to the prediction of the Billboard chart. Yekyung et al. focused on two tasks: a) the relationship between Twitter activity regarding music and forthcoming sales and b) predicting hit songs for the next Billboard chart [10]. The authors gathered over 30 million tweets, searching for the keywords #nowplaying, #np and #itunes, presumably because these hashtags are used to indicate the song a Twitter user is currently listening to. They also collected information from previous Billboard charts over the span of 10 weeks, which resulted in a dataset of songs, each with its own title, artist, rank and the time period it stayed on the chart. Altogether, the information retrieved related to 178 songs and 134 artists. 3 distinct metrics were established: a) song popularity, i.e. the number of tweets related to a specific song, b) artist popularity, i.e. the number of tweets mentioning this artist, and c) the number of weeks a song appeared on the Billboard chart. Forecasting targeted predictions for the top 10 songs, since this was the range with the highest accuracy achieved. Using the Pearson correlation, artist popularity and number of weeks on the chart, were not found to be strong predictors for a song's ranking. However, considering song popularity if all 3 metrics are combined, it is possible to predict the imminent success of a song quite accurately.

Koenigstein et al. approached Billboard chart prediction through the exploitation of Gnutella peer-to-peer network [6]. In total, 185,598,176 query strings, originating from the USA, were gathered during a 30-week period. They used both M5 and C4.5 algorithms to predict the Billboard Hot 100 and the Billboard Digital Songs charts. Moreover, in some cases, they also considered a song's debut rank on the chart, while predicted positions were usually either the top 10

or 20. They verified that queries used in peer-to-peer services are strong predictors for a song's success. For the Billboard Hot 100 chart, precision surpassed 86%, while for the Billboard Digital Songs accuracy exceeded 89%.

Finally, Zangerle et al. examined the relationship between song-related tweets and their ranking on the Billboard Hot 100 [11]. Their objective was to investigate the resemblance of the amounts of Twitter data referring to Billboard tracks to the state of the actual chart, along with the temporal offset between them, so that it can be determined if tweets have a predictive value for the chart and not vice versa. The authors used an existing dataset which consists of 111,260,925 tweets that included the term #nowplaying, gathered during 2014 and 2015. They also collected data from the Billboard Hot 100 chart for the same time period, corresponding to 886 distinct songs. They observed that a song remains on the chart for an average of 11.74 weeks and may stay on the top 100 for up to 58 weeks. To determine the correlation of rankings, they calculated three metrics per song, based on tweets: the median of play-counts per week, the mean play-counts per day and the total play-counts for a whole week. The median achieved the highest correlation (0.5), for 481 songs or 54.29% of the dataset.

The temporal relationship between tweets and charts was investigated through cross-correlation analysis that increased the mean correlation to 0.57 compared to the value of 0.5 that was previously found. 89.23% of all the examined tracks appear to have a temporal lag in relation to the Billboard chart, while 41.09% of all tracks have a negative lag. This means that the last percentage is exploitable for providing predictions about the chart's ranking for the weeks to come. Moreover, the authors followed the same process for songs that were first noticed in Twitter data and appeared on the chart at a later point (619 tracks in total). Similarly, 42.64% of them featured a negative lag thus facilitating future chart forecasts.
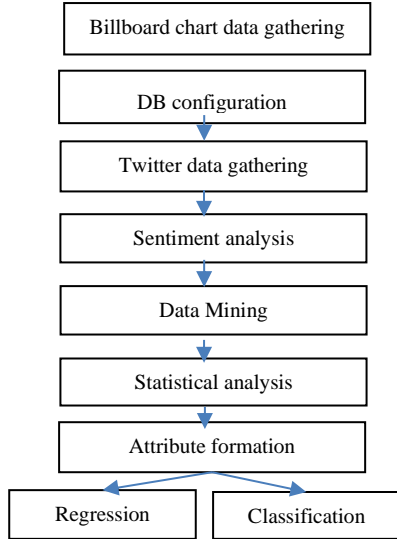
Regarding the predictive capability of Twitter data, they compared 3 models: one based exclusively on Billboard chart data, one relying only on tweets and the last one combining both. In terms of RMSE, the Twitter-based model had the worse score of 116.1, while the first one achieved an RMSE of 26.8. However, the multivariate model displayed a notable performance of 14.1. In conclusion a combination of data originating from both the Billboard chart and Twitter can significantly reduce forecasting error, thus song-related tweets can be useful for increasing the accuracy of ranking predictions. Summing up, about 41% of the collected tweets could be used for the prediction of the Billboard chart ranking if handled properly.

Having considered the literature, it can be hypothesized that the more attention a song gathers in Twitter, the more likely it is for people to listen to it and possibly end up purchasing it, further adding to its total popularity. Similarly, the whole publicity an artist gets and the image she presents, considering both her career and personal life, will probably motivate the public to check her work and increase her commercial success.

3 Approach

The methodology we followed is depicted in Fig. 1 and described below. The Search API was utilized for interacting with Twitter, using Python scripts for searching for tweets, storing posts and performing sentiment analysis. Database (DB) management tasks, including storing and retrieving Twitter data, were implemented using SQL Server. All useful information extracted from DB records was preprocessed and input to Weka [4] for mining.

3.1 Twitter Search API

Billboard chart data gathering

DB configuration

Twitter data gathering

Sentiment analysis

Data Mining

Statistical analysis

Attribute formation

Regression   Classification

Regarding song searching, tweets related to specific words were collected mainly to estimate their airplay count, either from radio stations or from individual users. These queries incorporate the term #nowplaying, which is joined via the AND operator with the titles of all songs, in brackets. Each title is placed in quotes and brackets, if it consists of multiple words. Quotes are necessary to ensure that the post contains all the words in the same order, and they are not just spread in the text. This would lead to the collection of many irrelevant tweets and would add noise to the dataset. Moreover, similarly to the queries about artists, all words of each title relate to the removal of spaces, and the hashtag symbol is inserted at the beginning of each concatenated word.

Figure 1: Proposed methodology

The strings for searching for songs do not include artist names, since they might exceed the 500-character limit and unnecessary complexity would be added to queries. Instead, artist names and titles are matched, when retrieving data from the DB where all tweets are stored. This ensures that the posts refer to the songs that need to be analyzed.

3.2 Sentiment Analysis

Overall, the VADER lexicon comprises useful scripts [5]. Compared to VADER, sentiment analysis with SentiWordNet [2] is not easy to implement, since it includes only the lexicon without any ready-to-use code. Therefore, it is up to the programmer to develop her own code to manipulate the lexicon and make the appropriate configurations to get the result she needs. The task is rather cumbersome because each word needs to be treated separately. It needs to be tagged based on the part of speech they represent, which can be noun, verb, adjective, adjective satellite (a subcategory of adjective according to WordNet) and adverb, and also each term has to be assigned with a score indicating its usage frequency.

For that reason, a handy script from [13] was utilized, featuring a class with methods for calculating different sentiment parameters for a given word or phrase. Most importantly, a procedure for estimating the total sentiment score of a sentence was used. This was the metric of interest we used to make comparisons with the respective result of VADER's sentiment analysis. The scoring method considers the following negation words: (not, n't, less, no, never, nothing, nowhere, hardly, barely, scarcely, nobody, none). Furthermore, there are 3 options for the generation of the sum of scores: average, geometric and harmonic.

The connection to the DB and the execution of queries was implemented in a similar way. However, in this script, there is a need for an extra query to fetch all the tweets that have not been analyzed yet. The table is updated with the compound column receiving the sentiment score through the compound dimension of the polarity_scores method with the extracted tweet text as

a parameter. If the compound column was not specified, the method would also return all other metrics, meaning the positive, negative and neutral scores. In order to set the sentiment column, a new method has been defined, which takes a chunk of text as a parameter and returns its semantic orientation (-1 for negative, 0 for neutral and 1 for positive) based on the common thresholds of the compound value.

3.3   Mining Information from the Database

In terms of preprocessing, tweet replication is treated with the removal of all records that have the same tweet_id leaving only one copy. Any records with a black text field are also redundant and therefore deleted. Twitter typically does not allow tweeting blanks and all tweets captured by the Search API are retrieved using specific keywords, which should be part of a post's text. Discarding posts with an empty string is just an added form of protection against unexpected behavior.

Duplicate tweets with a different tweet_id field, but identical text, were not removed because they were perceived as adding information and depicting extra attention towards a topic. Moreover, especially in the case of songs that are gathered using the #nowplaying term, the content of the posts is mainly fixed, as it usually includes the artist and title of the track, sometimes followed by a link.

Twitter's Search API captures all tweets, including retweeted posts. These can be sometimes recognized from their textual content, which begins with the pattern "RT @[username]", where [username] refers to the user being quoted. Nevertheless, this is not an official feature and it is not necessarily inserted in every retweet, so there is no accurate way to determine that a post captured via the API is definitely a retweet.

In order to produce a tradeoff between the number of occurrences for each distinct text value (that is the number of posts that have the exact same text) and the number of retweets, the greatest value of the two is chosen. The idea behind this is to get the biggest sample possible and exclude the difference between the two parameters, which is probably tweets that coincide.

For retrieving song tweets from the DB, the song title is used with and without spaces among words. For most tracks, the artist must also appear in the text. This is a measure to filter out any records with different songs that happen to have the same title or phrases which are absolutely irrelevant to the tracks under examination.

3.4   Attribute Description

In order to use Weka, all the information extracted from the Billboard chart and gathered tweets should be organized in attributes and formatted properly to create an .arff file. The file consisted of 80 instances in total with no missing values. Each instance describes a specific track for one week, during which the track was at one of the top 10 positions of the chart. So, the data set consists of 10 instances per week.

Since our hypothesis is that most of the attributes are positively correlated to the success of a song, a song will get to a higher position when attribute values increase. The Billboard chart ranks songs in ascending order, starting with the most popular ones. So, mainly for harmonization with the rest of the attributes, as it was seen in [10], the position of each track is inverted by subtracting it from 101. For instance, number 1 song will be in position 100, while number 10 in 91 and so on. This also applies to previous-position and position attributes.

Each instance of the data represents a different song for a single week and it has to be compared to all other instances, which are the remaining 9 tracks from the top 10 positions of the chart for

the same week and 10 more tracks for every week that has been monitored. Data gathering took place daily in a scheduled manner, but the total number of tweets captured on a weekly basis would differ each time, despite requesting a consistent number of items. Therefore, measuring every instance against all the others required some transformation on the specimen of tweets. For each week, the total number of posts for the top 10 songs was summed and the percentage of each track in comparison to the other 9 was calculated. Attributes song-play-count, artist-play-count and artist-tweets were tuned as described above.

The attributes selected are described below:

- Previous-position: The song position one week before.
- Chart-weeks: The number of weeks that the song has been in one of the 100 positions of the chart since its first entry. This value is set to 1 as soon as the song makes its chart debut.
- Top-weeks: The number of weeks that the song was in number 1.
- Song-play-count: The tuned number of song related tweets, including retweets, with the #nowplaying keyword.
- Artist-play-count: The tuned number of tweets, including retweets, which are related to the artist who sings this particular track, along with the #nowplaying keyword, not necessarily referring to this specific song. This is an estimation of the total number of play-counts for the artist, irrespective of the song being played. If a song belongs to more than one artist, the artist with the maximum number of tweets associated with her, represents the song.
- Artist-tweets: The tuned number of tweets, including retweets, which are related to the artist who sings this track. If a song belongs to more than one artist, the artist with the maximum number of tweets associated with her represents the song.
- Artist-sentiment-analysis-1: The average value of the score, produced by VADER via sentiment analysis on all tweets associated with the artist of this song. If there are many artists, the choice is made similarly to the previous cases.
- Artist-sentiment-analysis-2: This attribute is calculated like artist-sentiment-analysis-1 and the only difference is using SentiWordNet for score generation.
- Position: The position of the song for that week. This is the class attribute to be predicted.

Table 1 shows type and value restrictions for each attribute.

Table 1: Attribute type and values

Attribute	Type	Range
Previous-position	Integer	[0 , 100]
Chart-weeks	Integer	>= 1
Top-weeks	Integer	>= 0
Song-play-count	Real	[0 , 100]
Artist-play-count	Real	[0 , 100]
Artist-tweets	Real	[0 , 100]
Artist-sentiment-analysis-1	Real	[-1, 1]
Artist-sentiment-analysis-2	Real	[-1, 1]
Position	Integer	[0 , 100]

4 Results and Discussion

This section presents and discusses results

4.1   Results

Table 2 presents the attributes in Pearson's correlation value descending order.
Table 2: Pearson's correlation per attribute

Attribute	Pearson's Correlation Coefficient
Chart weeks	0.377
Previous position	0.370
Song playcount	0.292
Artist playcount	0.147
Artist sentiment analysis 2	0.134
Artist sentiment analysis 1	0.133
Top weeks	0.123
Artist tweets	0.013

Based on this output, the attributes can be divided in 3 categories with regards to their correlation (r) with the predicted class (position):

- **Moderately correlated attributes (0.2 < r < 0.7):** These are: *chart-weeks*, *song –playcount* and *previous-position*.
- **Weakly correlated attributes (0.1 < r < 0.2):** *Artist-play-count*, *artist-sentiment-analysis-2*, *artist-sentiment-analysis-1* and *top-weeks* are the attributes in this category.
- **Unrelated attributes (r~0):** *Artist-tweets* is the only attribute in this category with its correlation coefficient close to 0.

Regarding regression analysis, the following results were generated using 10-fold cross-validation. Support Vector Regression had the smallest Mean Absolute Error (MAE): 4.0515 and Random Forest the lowest Root Mean Square Error (RMSE): 8.8117. Bagging also achieved decent results with RMSE: 8.961 and the highest correlation coefficient: 0.5697.

Hit prediction refers to forecasting whether a song is going to be within a specific range of positions or not. This is a classification problem, as each instance should be matched to one of the two available classes. For this purpose, a new attribute should be created through Weka's preprocessing functionality. After experimenting with all available algorithms using 10-fold cross-validation, J48 and PART appeared to be the classifiers with the best performance in terms of both accuracy and F1 scores.

The best scoring algorithms in the case of top 5 hits are Filtered Classifier and Decision Table, achieving 90% accuracy and the same for precision, recall and F1 score. F1-scores was 0.875 for hits and 0.917 for non-hits. For the prediction of the top 20 hits of the following week, the classifiers that achieved high accuracy and F1 values were LMT and Simple Logistic.

4.2   Discussion

This study investigated the forecasting strength of social media, focusing on the prediction of the Billboard's Hot 100, based on data extracted from the chart and music-related Twitter posts referring to artists and songs on the top ten each week. In total, more than one million tweets were gathered, and data collection lasted for about 2 months, during October and November in 2018.

Twitter data appear to be quite representative when it comes to the number of total play-counts for a song. Considering the generally accepted ranges, there is a moderate correlation (0.2917) between the number of tweets that include the title of a song and the #nowplaying terms as a subset and the imminent success of the song on the Billboard Hot 100 chart for the following week. On the contrary, tweets that provide an estimation for an artist's total play-counts in general are not adequate to give an accurate picture of the future performance of a particular song, as the correlation coefficient was shown to be weak.

There appears no relation between an artist's publicity, positive or negative, expressed by the total number of tweet mentions, and his ranking on the chart, since the value of the correlation coefficient was close to 0 (0.0134). Furthermore, there is no evidence that the positive attention an artist gets, represented by a value that estimates how favorable the posts related to her are, is significantly correlated to the positioning of her tracks on the Billboard chart, at least in the short term. Findings concerning sentiment analysis cannot be generalized for other domains other than this particular chart, as many other studies have ascertained [1], [7].

Features derived from tweets combined with chart data, can provide results of noteworthy accuracy, but this happens only in the investigation of specific aspects of the problem. Regarding regression analysis, the best performance was achieved by Support Vector Regression with MAE 4.0515 and Random Forest with RMSE 8.8117. MAE value means that the position predictions were on average 4 ranks away from the actual values. Nevertheless, RMSE is, in most cases, considered to be more reliable, as it penalizes outliers. According to RMSE, the average derivation is approximately 8 positions.

Result evaluation depends on the type of problem that needs to be addressed. For example, if the goal is to predict all the Billboard chart for the following week, the values could be considered useful. On the other hand, if it is desired to predict a particular range of chart positions, especially if this is limited, such as top 10 hits for instance, then this model would not be able to provide an accurate prediction. The highest correlation coefficient was 0.5697, with Random Forest, and the squared correlation coefficient was approximately 0.325, which is low compared to the best result found in [10], with a squared correlation of 0.57, using Support Vector Regression and 5-fold validation.

Hit prediction, implemented via classification, yielded some promising results: top 10 songs could be predicted with an accuracy of 85% utilizing J48 and PART. The F-scores for PART was 0.906 for hits and 0.625 for non-hits, while the same values for J48 were 0.909 and 0.571. In [10], Random Forest with 5-fold cross-validation, achieved 90% accuracy and 0.901 and 0.899 F-scores for hit and non-hits, respectively.

Experimenting with different ranges for hits, specifically the top 5 and top 20, scores were higher than for the top 10 range. Filtered Classifier and Decision Table achieved an accuracy of 90% and F-scores of 0.875 for the top 5 hits and 0.917 for non-hits. For the prediction of the songs in the 20 highest positions Simple Logistic and LMT achieved accuracy of 96.25%. Specifically, the F-scores for hits and non-hits were 0.980 and 0.727, respectively. In [10], the accuracy for the same range (with the Random Forest Classifier and a 5-fold validation) was 88.2% and the F-score for hit songs was 0.885. However, they achieved a better F-score for non-hits, which was 0.879.

5   Conclusions and Future Work

This study gathered data from two sources: titles, artist names and rankings from the Billboard Hot 100 chart and tweets related to songs in the top 10, in order to test how their integration could

be used for forecasting future charts. Our findings suggest that there is a moderate correlation of 0.2917, between the number of mentions of a song and its chart performance. No significant relationship was observed between the attention an artist gets on Twitter, even via emotionally charged tweets, and their tracks' success.

Regarding regression analysis, the best score was a mean square error of 4.0515 and RMSE 8.8117. Both metrics show the average number of positions between the actual and predicted values, and their evaluation depends on the range of positions we want to predict. With the best scoring algorithm, hit prediction could be achieved with an accuracy of 80% and F-scores of 0.881 and 0.385 for hits and non-hits, for the 10 highest ranks. Similarly, for the top 5 hits these values were 90%, 0.917 and 0.875, respectively. Finally, for the top 20 songs, the results were competitive to previous research, achieving top accuracy of 96.25% and F-scores 0.980 for hits and 0.727 for non-hits.

Overall, the study has shown that mining Twitter data, extracting specific information and handling them properly can provide useful conclusions regarding the next Billboard chart, although it is not yet capable of perfect predictions.

5.1   Limitations and Future Research

This research gathers and examines Twitter and Billboard chart data for hit songs at the top 10. Tracks at lower ranks were not considered, thus it is impossible to predict their way up. So, future research could investigate the chart at a larger scale and consider all songs on the chart, and create a more comprehensive dataset, consisting of chart data for a period of one or more years, like the ones used in [10] and [11].

From a technical point of view, researchers are encouraged to try to capture tweets incorporating the #nowplaying term in general, without targeting specific songs or artists. Finding Twitter data referring to a set of 100 songs for each week would require a complicated methodology if the Search API is used, due to the search query limitations that were discussed in section 3.1. On the other hand, a more generalized approach would demand a lot of space for storing data and would capture a lot of irrelevant tweets that would need to be excluded at a later stage. Thus, there should also be extra preprocessing effort. Nevertheless, collecting tweets indiscriminately ought to give a better insight regarding airplay trends and could probably aid the discovery of songs that are not at the chart at the given moment, but would be introduced in the weeks to come.

Undoubtedly, an exhaustive study of the optimized way to predict the Billboard chart should rely on a model that considers all parameters that affect song ranking, such as streaming activity and song sales. During data gathering and experimentation, it was observed that many tracks made their debut into the chart by occupying one of the first 10 positions, consequently they could not have been monitored even if all 100 songs were accounted for. They also tended to follow a specific pattern: usually they dropped significantly after their successful entrance into the chart and their play-counts seemed to be quite low at that time. It is, therefore, speculated that their steep rise to the top is basically owing to music sales and streaming data. In order to get a better idea about this phenomenon, research should be extended beyond the limits of the Billboard chart and the airplay counts and consider the other dimensions that contribute to the formation of the chart each week.

With regards to sentiment analysis, the degree to which tweets constitute a trustworthy sample for determining the positive or negative disposition of the population towards an artist is still vague. There may be a relationship, but it may not be directly visible. This fact raises new questions about the predictive strength and limits of sentiment analysis in social media and is

something worth investigating in the future. It is suggested that researchers explore the impact of sentiment analysis in a longer term context. This essentially means, monitoring sentiment orientation for an artist and track the performance of her songs for many weeks, possibly considering some time lag, as the public's response to a likeable or unlikeable person may be revealed gradually.

Different angles can also be investigated, for instance, the relationship between a positive or negative bias towards an artist and the number of tracks belonging to him that are on the chart at the present time or the rate at which these tracks ascend and descend in the chart. Researchers could, additionally, consider improving the mechanism of sentiment analysis and attempt to enhance existing lexicons in order to make them more suitable for music-related data. For instance, some terms referring to music, like radio, album, concert, listen, etc., which would otherwise be characterized as neutral, could infuse a positive tone into a tweet, when the name of an artist is included in the text.

References

[1] Asur S., Huberman B., "Predicting the Future with Social Media", Web Intelligence and Intelligent Agent Technology (WI-IAT), 2010 IEEE/WIC/ACM Int'l Conf., vol. 1, 2010, pp. 492–499.

[2] Baccianella S., Esuli A., Sebastiani F., "SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining", Proc. Int'l Conf. on Language Resources and Evaluation, LREC 2010, 17-23 May 2010.

[3] Boldt L. C., Vinayagamoorthy V., et al., "Forecasting Nike's Sales using Facebook Data", IEEE Int'l Conf. on Big Data, 2016.

[4] Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J., *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, 2016.

[5] Hutto C.J., Gilbert E., "VADER: A Parisomonious Rule-based Model for Sentiment Analysis of Social Media Text", Proc. 8th Int'l AAAI Conf. Weblogs and Social Media, 2014.

[6] Koenigstein N., Yuval S., Zilberman N., "Predicting Billboard Success Using Data-Mining in P2P Networks", 11th IEEE Int'l Symposium on Multimedia, 2009.

[7] Oikonomou L., Tjortjis C., "A Method for Predicting the Winner of the USA Presidential Elections using Data Extracted from Twitter", 3rd IEEE SE Europe Design Automation, Computer Engineering, Computer Networks, and Social Media Conf., 2018.

[8] D. Rousidis, P. Koukaras, C. Tjortjis, "Social Media Prediction A Literature Review", Multimedia Tools and Applications, Springer, 79(9), 6279-6311, 2020.

[9] Beleveslis D., Tjortjis C., Psaradelis D. and Nikoglou D., 'A Hybrid Method for Sentiment Analysis of Election Related Tweets', 4th IEEE SE Europe Design Automation, Computer Engineering, Computer Networks, and Social Media Conf., 2019.

[10] Yekyung K., Bongwon S., Kyogu L., "#nowplaying the Future Billboard: Mining Music Listening Behaviors of Twitter Users for Hit Song Prediction", Proc. 1st int'l workshop on Social media retrieval and analysis, pp. 51–56, 2014.

[11] Zangerle E., Pichl M., Hupfauf B., Specht G., "Can Microblogs Predict Music Charts? An Analysis of the Relationship Between #Nowplaying Tweets and Music Charts", Proc. 17th Int'l Society for Music Information Retrieval Conf., 2016.

[12] twitter_search github.com/agalea91/twitter_search/blob/master/twitter_search.py (last accessed 2/20).

[13] Sentimentanalysis github.com/anelachan/sentimentanalysis (accessed 2/20).