Twitter Sentiment Analysis: Does the Audience Opinion Reflect on the Oscar Winners?

Igor Tannus Correa, Daniel Duarte Abdala, Rodrigo Sanches Miani, Elaine Ribeiro Faria
School of Computing
Federal University of Uberlandia
Uberlandia, Brazil
igoortc@gmail.com, {abdala, miani, elaine}@ufu.br

***Abstract*—This paper aims to perform the sentiment analysis of Twitter user posts related to the movies nominated for Best Picture of the 2017 Academy Awards. It compares the sentiment expressed on the social network with the Oscars results. By following the proposed methodology, it was possible to predict with a good accuracy which movie would be the winner of the ceremony and which would be among the less prestigious ones. The collected data was initially pre-processed, a significative number of tweets were manually classified to compose the training base and finally, it was evaluated by three distinct approaches, namely Naive Bayes, Distant Supervision Learning, and Polarity Function. It has also been noted that Twitter users prefer to use the network to post positive comments about movies rather than saying bad things about the ones they did not like. Furthermore, it was verified that award shows such as the Oscars cause a growth in the number of posts on Twitter.**

***Index Terms*—Sentiment Analysis, Twitter, Big data applications**

I. Introduction

Social networks are online platforms where different entities – such as users, groups or organizations – are able to create and share different kinds of content and also access publications of the other entities on the network [1]. Some of these platforms gather millions, or even billions of users, representing more than two thirds of the global online population [2].

The social network Twitter[1] was launched in 2006 proposing that each of the posts – called tweets – published by its users must not exceed 140 characters.[2] Since then, its popularity has been growing and its main goal has been the exchange of ideas and information among its users around world [3]. Twitter has more than 313 million monthly active users and the platform has been modified through time to keep its users online [3].

The film industry is one that benefits from social networks in order to promote movies and keep track of their clients' profiles and opinions about their movies. Besides the immense visibility provided by the social networks, they also provide to film distributors access to a wide range of opinions [4] that may influence future movie releases and how promotion is made.

Social networks and entertainment are usually strongly attached, therefore several users use their accounts to express their opinion, enthusiasm or disappointment on a movie, its cast, and production – especially the movies that are nominated for the Oscars, the most famous and respected award show in the film industry [5].

However, despite the numerous tweets and notable excitement about the theme, there are only a few studies that try to find correlations between the opinion of the Twitter audience and the votes of the Academy [4]. Some authors have tried to create mathematical models using critics' reviews and social networks like IMDb (Internet Movie Database) to predict Oscar winners [6]. Nonetheless, none of these authors have used only Twitter data in order to seek for this correlation.

Twitter is one of the social networks that is most used in Sentiment Analysis papers because it is an abundant source of personal opinions that come from the whole world [7]. Moreover, about 350 thousand tweets are published per minute [8] by users that belong to multiple social groups with different interests [7]. Therefore, it is relevant to find out if there is a correlation between the opinion of the Twitter audience and the Oscars result, i.e., discovering if the analysis of Twitter user data could predict the Academy Awards winners.

The aim of this paper is to verify if there is a correlation between the winners of the 2017 Oscars and the sentiment expressed by people through Twitter. In order to achieve this goal, a tweets database was created and preprocessing techniques were applied to conform the dataset. Subsequently, the tweets were automatically classified as positive, negative, or neutral. Finally, the sentiment of the tweets was compared to the Oscars results by different similarity measures.

Once the results were officially available, it was be possible to explore if there were any connections between the opinion of the Twitter audience and the winners chosen by the Academy of Motion Picture Arts and Sciences. The goal was to find out if it was possible to predict which movies would be winners or losers of the ceremony by only using the tweets related to the theme.

The contributions of this paper can be summarized as follows:

- Creation of a public database of tweets related to movies nominated to the Best Picture category of the

[1]Available on: http://www.twitter.com

[2]As of November 2017, Twitter expanded its posts' limit to 280 characters (Available on: http://bit.ly/TwLim). However, the data collection for this paper was conducted before this date, thus the original limit will be considered in this paper.

2017 Academy Awards containing 889,840 preprocessed tweets written in English.[3]

- Creation of a database of tweets related to movies nominated to the Best Picture category of the 2017 Academy Awards containing 3,235 tweets manually labeled as "positive", "neutral", or "negative". This database may be used as a training set for supervised learning algorithms.[4]
- Formulation of a measure to build an Oscars ranking based on the number of nominations and victories earned by each of the analyzed movies.
- Development of a public Java tool able to perform several preprocessing steps in a database, which can be used in different contexts. This tool also includes dictionaries for translation of emoticons, slangs, and abbreviations.[5]
- Comparison between different text classifiers to perform Twitter sentiment analysis, suggesting the Naive Bayes classifier as promising to perform tweet sentiment classification tasks.
- Analysis of the correlation between the sentiment expressed on Twitter and the movies nominated to the Best Picture category of the 2017 Academy Awards.

The organization of this paper is described as follows.

- **Section II – Related Work:** lists several works related to the theme of this paper.
- **Section III – Methodology:** explains the methodology used in this paper, describing the steps executed in order to perform the sentiment analysis of the collected tweets.
- **Section IV – Experiments:** describes several experiments that were made by analyzing the collected data and the results obtained from the chosen classifier.
- **Section V – Conclusion:** the conclusions are presented as well as suggestions of future work.

II. Related Work

The amount of work related to the field of Sentiment Analysis is abundant [9]. Several researchers have been exploring different methods and techniques that can be applied to data from social networks [4] [10]. However, there is a certain absence of works aiming to analyze if there is a connection between social networks and movies, especially the ones related to the Academy Awards.

Researchers have approached similar themes in [1] and found meaningful connections between the sentiment of social media users and the financial performance of movies. They reached that conclusion by extracting data from Facebook and Twitter related to movies that were not yet released, treating and classifying the data, and using Spearman's correlation to determine if there was a relation between the posts and the financial performance of the movies.

In [6], the authors were able to predict Oscar nominees according to user opinions published on IMDb. After collecting the data, the authors created a customized dictionary to measure the levels of positivity on a text, thus proving that it is possible to predict Oscar nominees based on movie reviews in online forums like IMDb.

The work proposed in [11] aimed to predict the winners of eight categories of the 2017 Academy Awards by using the machine learning platform BigML. A dataset was built by gathering information of previous Oscar nominees and winners, and other important movie awards between the years of 2000 and 2016. In addition to it, ratings and reviews written by IMDb users were included in the database. The models developed by the author were able to correctly predict the winners of five out of eight categories analyzed, supporting the fact that a proper dataset and a good machine learning model may bring relevant results.

In this paper, the tweets posted about movies nominated for the Best Picture on the 2017 Oscars were analyzed by performing several Sentiment Analysis techniques that allow to predict the sentiment of the tweets according to a classifier. Once the the analysis was done, it was possible to draw conclusions about the Oscars result and the result obtained from the analysis.

III. Methodology

In order to achieve the goal proposed by this study, a set of steps was executed to analyze the sentiment of the tweets.

A. Data Collection

The collection of Twitter data was made considering the time period between the day the Oscar nominees were announced (January 24th, 2017) and the day before the ceremony was presented (February 25th, 2017). The tool *GetOldTweets*[6] was chosen to perform this step, and nine queries were executed in order to search for tweets written in English.

The keywords used as parameters were the original titles of the analyzed movies: "arrival", "fences", "hacksaw ridge", "hell or high water", "hidden figures", "la la land", "lion", "manchester by the sea", and "moonlight".

The number of collected tweets is shown on the second column of Table III.

B. Labeled Database

In order to obtain a training set to build a supervised classification model and validate the best classifier, a labeled tweet database was created. A random sample of tweets was selected from the original database, and each tweet was manually labeled as "positive", "negative", or "neutral". Considerable amounts of tweets representing each of the three sentiments were included, even though this kind of sorting is one of the greatest challenges while building a labeled database.

Table I summarizes the data in the labeled database, which represents 0.36% of the original database after it was preprocessed.

After collecting the data for each movie and sorting them by week, the ground truth process listed below was followed in order to build the labeled database.

[3]The original and preprocessed databases used in this study are available on http://bit.ly/TCCIgor

[4]The labeled database used in this paper is available on http://bit.ly/TCCIgor

[5]The algorithm developed for this paper is available on http://bit.ly/TCCIgor

[6]Available on: http://github.com/Jefferson-Henrique/GetOldTweets-python

TABLE I
SUMMARY OF DATA CONTAINED ON THE LABELED DATABASE

Sentiment	Number of tweets
positive	1444
negative	1362
neutral	429
Total	3235

1) Random selection of tweets, ensuring that a significant amount was selected from each of the five weeks analyzed.
2) Each randomly selected tweet must have satisfied some conditions to be included in the labeled database:
   a) It is written in English.
   b) It definitely refers to the movie being analyzed.
   c) It contains relevant terms for the classification, i.e., it could not contain only links, stop words, etc.

   This procedure was useful for one of the preprocessing steps, because a list of unrelated words was created for some of the movies in order to remove from the database the tweets that do not refer to the movies being analyzed.
3) Once all of the conditions were satisfied, each tweet was manually labeled by one of the authors with one of the three sentiments: "positive", "negative", or "neutral".
4) The selected tweets and their labels compose the labeled database.

C. Preprocessing of Tweets

Once the tweets collection is done, the data needed to be preprocessed in order to discard what is irrelevant to the classification step [12]. A Java tool capable of performing several preprocessing steps was developed especially for this study and it can be used in different contexts.

The preprocessing steps conducted in this study, using the tool referred above, are listed below. The algorithm was executed individually for each movie.

- Conversion of upper case letters to lower case, for the purpose of standardizing the text.
- Removal of links, because they have no semantic value.
- Replacement of emoticons with matching words, with the intention of improving the classification model. In order to do this, a dictionary of emoticons and their respective translation was built based on a list of frequently used emotions[7].
- Removal of non-alphabetic characters and punctuation.
- Removal of user mentions (preceded by "@" on Twitter).
- Removal of the movie titles, since the presence of some of these terms could mislead the results generated by the classifiers. For instance, the word "hell" (in the movie title "Hell or High Water") usually expresses a negative feeling, which could corrupt a positive tweet's classification.
- Removal of repeated letters. Twitter users often repeat word letters to intensify a feeling, but those words with repeated letters are not recognized by classifiers – for this reason, the repeated letters have been removed. For example, "i looooved la la land" turns into "i loved la la land".
- Replacement of slangs and abbreviations with complete words or expressions. Twitter users generally write slangs and abbreviations considering that there is a character limit for each tweet. A dictionary containing 367 slangs and abbreviations was built to incorporate new terms to the tweets and ensure that the semantics of each tweet was preserved.
- Removal of stop words. These are words that are very common in a language and do not have relevant semantic value, like "a", "the", and "what". Therefore, they were removed from the tweets – for instance, "this movie is awesome" turns into "movie awesome". A list of stop words, part of the Onix Text Retrieval Toolkit [8] was used in this step.
- Removal of tweets that were not related to the theme. One of the challenges in Sentiment Analysis is to ensure that the data collected corresponds to the topic that is being explored. The titles of the movies "Arrival", "Fences", "La La Land", "Lion", and "Moonlight" refer to other words and expressions spoken in the English language, as shown on Fig. 1 . Besides, some movie titles can also refer to other movies, songs, etc. like the tweet on Fig. 2.

Fig. 1. Some of the tweets collected for the movie "Fences" referred to the wall that USA president Donald Trump intends to build to reduce the illegal entry of immigrants into the country.

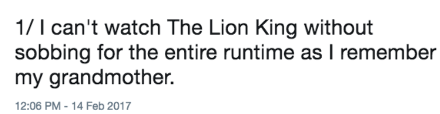

Fig. 2. The query for the movie "Lion" returned some tweets that made reference to Disney's *The Lion King*.

  To minimize the effects of this challenge and ensure that the tweets being analyzed indeed refer to these movies, for each of them a list of unrelated terms was created. A snapshop of the list of terms unrelated to movie "Fences" is shown on Table II. Tweets containing at least one of the terms in the list of its respective movie were removed from the movie database.

The number of tweets originally collected (before preprocessing) and the number of tweets contained in the resulting database (after preprocessing) is detailed on Table III.

[7] Available on: http://bit.ly/wikiEmot

[8] Available on: http://www.lextek.com/manuals/onix/stopwords1.html

TABLE II
EXCERPT FROM THE TERMS NOT RELATED TO THE MOVIE "FENCES"

picket	neighbor	garden
wall	refugee	yard
trump	border	government

TABLE III
NUMBER OF TWEETS IN THE DATABASE BEFORE AND AFTER THE PREPROCESSING STEPS

Movie	**Before preprocessing**	**After preprocessing**
Arrival	138,825	135,214
Fences	53,211	41,682
Hacksaw Ridge	54,689	48,740
Hell or High Water	14,919	13,320
Hidden Figures	145,868	137,151
La La Land	250,942	244,213
Lion	186,295	150,641
Manchester by the Sea	31,768	28,601
Moonlight	108,121	90,278
Total	**1,035,739**	**889,840**

D. Tweets Classification

In this paper, three different text classification approaches were considered: supervised learning, distant supervision learning, and polarity function.

The labeled databased was used to validate the algorithms. Each of these learning methods is briefly explained below.

1) Supervised Learning: The Naive Bayes is one of the most widely used supervised learning methods in the scope of Sentiment Analysis, because of its remarkable performance in text classification [13]. For that reason, it was the supervised learning algorithm chosen to be tested in this paper. It is a probabilistic algorithm that is based on prior knowledge of the problem and training examples to determine the probability of a document belonging to a certain class [14] [15].

2) Distant Supervision Learning: Distant supervision learning uses an alternative way to generate training data. In this strategy, an existing database is used to collect instances related to the relation to be analyzed. Then, these instances are used to automatically generate training sets [16].

Sentiment140 is an specific tool for Twitter Sentiment Analysis. It uses distant supervising learning and a Maximum Entropy classifier[9] to calculate the polarity of a sentence based in a database labeled according to the emoticons found on the tweets that are in it [17].

According to [17], the method used by Sentiment140 – using tweets that contain emoticons as a training set – has proven itself as a good technique to classify tweets, since classification algorithms like Naive Bayes, Maximum Entropy, and Support Vector Machines achieved excellent accuracy indices when tested.

3) Polarity Function: TextBlob[10] is a Python library for word processing that provides solutions to different tasks related to natural language processing [18]. This tool integrates with the NLTK (Natural Language Toolkit)[11] platform and with the web mining module Pattern[12].

One of the text classification methods available on TextBlob is the polarity function, which returns the polarity of a sentence given as input. This function uses an unsupervised learning algorithm to classify the sentences based on a lexicon built by a specialist and manually labeled according to its polarity strength, subjectivity, and intensity of each word [19]. The lexicon is a dictionary including frequent adjectives present in online product reviews [19].

E. Assessment of the Classifiers

Each of the classifiers was tested using the labeled dataset, and the model validation technique applied was the 10-fold cross validation.

After the classifiers were tested, confusion matrices were built derived from the real sentiment stated for each tweet on the labeled base and the classification result. Thus, the quality measures accuracy, precision, and recall were calculated – allowing us to evaluate the classifiers.

Once the assessment of the classifiers was done, the best one according to the measures calculated was chosen to classify the complete preprocessed tweets database.

F. The 2017 Academy Awards Ranking

There is no ranking that classifies the movies that competed on the 2017 Oscars – the only information available is the number of nominations and wins gotten by each movie. Thus, in order to facilitate the comparison between the result of the 2017 Oscars and the one obtained with the classification, a measure was created especially for this paper.

By visiting the official Oscars website[13], it can be noted that the visual representation of the winners is presented on three different sizes. The most relevant categories – for example, Best Picture and Best Animated Feature Film – are more prominent on the web page, therefore it was decided that they have more weight. Taking this into account, weights were assigned to each category based on how relevant they are according to the way they appear on the Oscars website.

Consequently, the Table IV was built. It shows all of the categories considered and the respective weight (w) that each one sums up on the score of the movies.

The categories were included only if at least one of the Best Picture nominees was competing in it. According to the Oscars website, the Best Directing category would have weight *1*, but for this paper it was decided that it would have weight *2*, because it is also a very relevant category as it is considered one of the "Big Five" Oscar awards [20].

The number of nominations and wins that each movie got was obtained from the official Oscars website. Then, the score of each movie was calculated according to 1, where i indicates the index of each of the 16 categories (described on Table

[9] Available on: http://help.sentiment140.com/api

[10] Available on: http://textblob.readthedocs.io/en/dev/

[11] Available on: http://www.nltk.org/

[12] Available on: http://www.clips.uantwerpen.be/pattern

[13] Available on: http://oscar.go.com/winners

TABLE IV
THE WEIGHTS OF THE CATEGORIES

i	Categories	*w*
1	Picture	3
2–6	Directing, Actor, Actress, Supporting Actor, Supporting Actress	2
7–16	Original Screenplay, Adapted Screenplay, Original Score, Original Song, Sound Editing, Sound Mixing, Production Design, Cinematography, Costume Design, Film Editing	1

IV). Besides, the weight to be multiplied by the number of nominations received by each movie is reduced by half, since the victories are more important.

$$score_{\text{movie}} = \sum_{i=1}^{16}(vic_{\text{i}} \times w_{\text{i}}) + (nom_{\text{i}} \times \frac{w_{\text{i}}}{2}) \qquad (1)$$

The terms of 1 are explained below.

- vic_i represents the number of victories that the *movie* has received on the category i.
- nom_i represents the number of nominations that the *movie* has received on the category i.
- w_i represents the weight of the category i.

IV. EXPERIMENTS

The objective of this section is to analyze the complete preprocessed database and its relation to the 2017 Oscars result. In order to do this, several comparisons were made.

1) Building the 2017 Oscars Ranking: For this paper, only the nine movies nominated for the Best Picture category were considered. Table V indicates the number of nominations and wins that each of these movies got on the 2017 Oscars.

TABLE V
NOMINATIONS AND WINS RECEIVED BY THE MOVIES

Movie	Nominations	Wins
Arrival	8	1
Fences	4	1
Hacksaw Ridge	6	2
Hell or High Water	4	0
Hidden Figures	3	0
La La Land	14	6
Lion	6	0
Manchester by the Sea	6	2
Moonlight	8	3
Total	**59**	**15**

Once having the information contained on Table IV and Table V, and knowing which categories each movie has won or has been nominated (available on the Oscars website), it was possible to build the ranking of the 2017 Academy Awards by using 1. The ranking is exposed on Table VI.

TABLE VI
RANKING OF THE 2017 ACADEMY AWARDS

Position	Movie	Score
1	*La La Land*	17.5
2	*Moonlight*	12.5
3	*Manchester by the Sea*	9
4	*Hacksaw Ridge*	7
5	*Arrival*	6.5
6	*Fences*	6
7	*Lion*	5
8	*Hell or High Water*	3.5
9	*Hidden Figures*	3

2) Number of Tweets Collected for Each Movie: This first experiment aimed to analyze the amount of tweets collected for each movie throughout the weeks and draw conclusions about the data.

The graph on Fig. 3 shows the number of tweets in the complete preprocessed database according to the movie and the week the tweets were posted on Twitter. It can be noted that the biggest frequency of tweets happened on the first week of analysis, i.e., the week when the 2017 Academy Awards nominees were announced.

It is possible to perceive that there is a relation between the movement on the social network and the 2017 Oscars, once the amount of tweets posted about the movies in the week when the nominees were announced was bigger than the other weeks. Besides, during the fifth week of analysis – closer to the date of the ceremony – the number of tweets posted about most of the movies has gradually increased comparing to the third and fourth weeks.

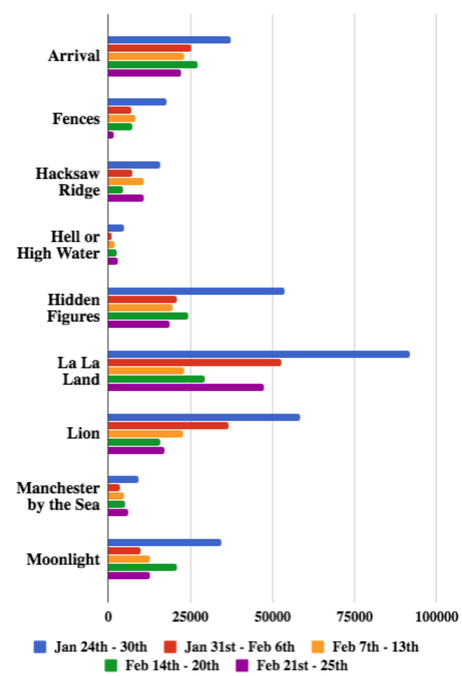

Fig. 3. Number of tweets posted about each movie throughout the weeks.

Also, based on Fig. 3 it is possible to discover that the most commented movies were “Arrival”, “Hidden Figures”, “La La Land”, and “Moonlight” – it is interesting to observe that “Arrival”, “La La Land”, and “Moonlight” are among the top 5 positions of the proposed 2017 Oscars ranking (Table VI).

3) Tweets Classification: The goal of this experiment was to compare the performance of the three classifiers being analyzed and choose the best one of them to classify the complete database.

After the classifiers were tested and their respective quality measures were calculated, Table VII was built. It is possible to see that the Naive Bayes classifier got the best results according to all of the three measures calculated, therefore it was decided that this algorithm was the best one to be applied over the complete preprocessed database.

This result was expected since the Naive Bayes is the only supervised learning classifier being tested, therefore it was expected that its performance was better than the others. The polarity function from the TextBlob library considers the terms of a sentence individually, ergo it was not able to classify each tweet as a whole. Finally, the Sentiment140 classifier, that uses distant supervising learning and has a training set based on emoticons, might not have been able to accordingly represent the database of this paper.

TABLE VII
ASSESSMENT OF THE CLASSIFIERS

Classifier	Accuracy	Precision	Recall
Naive Bayes	74.1%	69.8%	68.8%
TextBlob	63.9%	62.7%	64.2%
Sentiment140	26.2%	60.6%	42.3%

After the generation of the classification model based on the Naive Bayes algorithm, the test sets – that is, the individual files containing the tweets posted about each movie – were loaded into the classifier. This process was performed individually for each of the nine movies being analyzed.

Later, a summary of the tweets classified by the multinomial Naive Bayes classification was built, and it is shown on Table VIII. Each line of the table shows one of the nine movies being analyzed, followed by the number of tweets classified as “positive” by the classifier and the percentage of these tweets in relation to the movie’s tweets. The next columns show the same for the “negative” and “neutral” tweets.

4) Twitter and Oscar Indicators: There are several indicators that can be obtained from all the data and information that was gathered, as it can be seen on Fig. 4. These indicators can help comparing the sentiment of the tweets and the 2017 Oscars result, making it possible to know if the Twitter users sentiment could predict the Oscar winners.

Within the scope of the sentiment expressed by Twitter users, there are three indicators that were obtained: the number of tweets in the database for each movie (Ind. 1), the number of positive tweets among all of the instances classified as “positive” for each movie (Ind. 2), and the number of positive

TABLE VIII
SENTIMENT OF TWEETS ACCORDING TO NAIVE BAYES CLASSIFIER

Movie	Positive		Negative		Neutral	
Arrival	68,485	51%	27,002	20%	39,727	29%
Fences	20,199	48%	8,687	21%	12,796	31%
Hacksaw Ridge	31,998	66%	6,380	13%	10,362	21%
Hell or High Water	7,077	53%	2,064	15%	4,179	31%
Hidden Figures	77,314	56%	16,212	12%	43,625	32%
La La Land	104,507	43%	57,237	23%	82,469	34%
Lion	75,906	50%	36,978	25%	37,757	25%
Manch. by the Sea	14,879	52%	5,410	19%	8,312	29%
Moonl.	49,668	55%	17,130	19%	23,480	26%
Total	**50,033**	**50.6%**	**177,100**	**19.9%**	**262,707**	**29.5%**

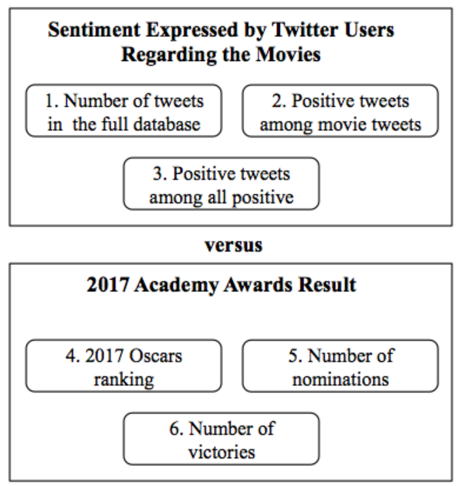

Fig. 4. Indicators obtained from the data and information gathered.

tweets among all of the instances classified as “positive” in the complete database (Ind. 3).

From the 2017 Academy Awards result, there were also three indicators obtained: the 2017 Oscars ranking created for this paper (Ind. 4 – Table VI), the number of nominations received by each movie (Ind. 5), and the number of victories gotten by each movie (Ind. 6).

The Tables IX and X show rankings based on the six indicators and they help on to provide a detailed analysis of the correlation between the two data scopes.

5) Comparing the Sentiment of the Tweets with the the 2017 Oscars Result: By analyzing Tables IX and X, it can be noted that “La La Land” has the biggest amount of tweets in the complete database, although it also has the smallest percentage of tweets classified as “positive” among its tweets. However, “La La Land” obtained the biggest amount of tweets classified as “positive” (Tabela VIII). This Twitter analysis corresponds to the fact that it is the movie that received most nominations

TABLE IX
RANKINGS BASED ON THE TWITTER INDICATORS

Movie	Ind. 1	Ind. 2	Ind. 3
Arrival	4th (15.2%)	6th (51%)	4th (15.2%)
Fences	7th (4.7%)	8th (48%)	7th (4.5%)
Hacksaw Ridge	6th (5.5%)	1st (66%)	6th (7.1%)
Hell or High Water	9th (1.5%)	4th (53%)	9th (1.6%)
Hidden Figures	3rd (15.4%)	2nd (56%)	2nd (17.2%)
La La Land	1st (27.4%)	9th (43%)	1st (23.2%)
Lion	2nd (16,9%)	7th (50%)	3rd (16.9%)
Manchester by the Sea	8th (3.2%)	5th (52%)	8th (3.3%)
Moonlight	5th (10.1%)	3rd (55%)	5th (11.0%)

TABLE X
RANKINGS BASED ON THE 2017 OSCARS INDICATORS

Movie	Ind. 4	Ind. 5	Ind. 6
La La Land	1st	1st (24%)	1st (40%)
Moonlight	2nd	2nd (14%)	2nd (20%)
Manchester by the Sea	3rd	4th (10%)	3rd (13%)
Hacksaw Ridge	4th	4th (10%)	3rd (13%)
Arrival	5th	2nd (14%)	5th (7%)
Fences	6th	7th (7%)	5th (7%)
Lion	7th	4th (10%)	7th (0%)
Hell or High Water	8th	7th (7%)	7th (0%)
Hidden Figures	9th	9th (5%)	7th (0%)

and wins and it has gotten the first place on the proposed Oscars ranking, which makes “La La Land” the winner of the ceremony, corresponding to the fact that this was the movie that has gotten the biggest number of nominations and wins.

On the other hand, the Tables also show that “Hell or High Water” can be considered one of the least prestigious among the movies being analyzed. The amount of tweets about this movie represents only 1.5% of the complete database and it is the movie with the smallest amount of positive tweets among all of the tweets classified as “positive” on the full base, implying that the Twitter audience showed little interest in this movie during the time period when the data was collected. This reflects on the fact that “Hell or High Water” also did not get any win on the 2017 Academy Awards.

Thereby, it can be noted by comparing Tables IX and X that there are some cases in which there is a conformity between the Oscars results and the sentiment expressed on the tweets.

However, although “Hidden Figures” is the third movie with the biggest amount of tweets in the complete database, having more than half of these tweets classified as “positive”, the movie got the smallest number of nominations among the movies analyzed and did not win any category, thus getting the last place on the proposed 2017 Oscars ranking. This is one of the unconformities that can be found by analyzing these tables.

A controversy also happens regarding the movie “Manchester by the Sea” which has gotten the third place on the proposed Oscars ranking, but it is the second movie with the smallest amount of tweets in the complete database. It also fills the eighth place on the ranking based on the Indicator 3.

In order to conduct a mathematical analysis of the correlation between the sentiment of the tweets and the 2017 Academy Awards result, Spearman’s ranking correlation coefficient was calculated between the rankings based on the Twitter indicators (Table IX) and the Oscar indicators (Table X). The results are displayed on Table XI.

TABLE XI
SPEARMAN’S RANKING CORRELATION COEFFICIENT BETWEEN TWITTER INDICATORS AND OSCAR INDICATORS

Indicators	Ind. 4	Ind. 5	Ind. 6
Ind. 1	0.15	0.43	0.12
Ind. 2	-0.23	-0.35	-0.12
Ind. 3	0.11	0.36	0.12

According to Table XI, statistically there is no strong evidence of an associations between the rankings. The biggest coefficients were found between the ranking of nominations (Ind. 5) and the rankings of the amount of tweets in the complete database (Ind. 1) and the positive tweets among all positive (Ind. 3). The latter has gotten a 0.43 coefficient, which indicates a moderate correlation between the rankings [21]. Consequently, it is possible to infer that the 2017 Oscars nominations affect on the Twitter discussion, causing a big amount of tweets to be published about the theme, but the opinion of the public does not necessarily reflects on the winners chosen by the Academy.

Consequently, it is possible to notice that although none of the coefficients found represent a high correlation between the rankings, a prediction of which movie will be the winner of the ceremony – in this case, “La La Land”– and which ones will be among the losers – like “Hell or High Water” – by classifying Twitter data and analyzing the results is very likely to be correct. In other words, this methodology allows the prediction of the extremities of the proposed Oscars ranking, even though the correlation between the other rankings is low.

V. CONCLUSION

The aim of this paper was to perform the sentiment analysis of tweets related to the movies nominated to the category of Best Picture of the 2017 Academy Awards. In order to do this, the steps of a Sentiment Analysis task focused on Twitter were executed: tweets collection, construction of a labeled database, tweets preprocessing, tweets classification, and validation of the results. Three different text classification approaches were applied on the labeled database: supervised learning – using the Naive Bayes algorithm –, distant supervision learning – using the Sentiment140 tool –, and polarity function – using the TextBlob library.

After testing the classifiers with the labeled database and evaluating the results by using different quality measures, the multinomial Naive Bayes algorithm was chosen to classify the complete database. It was possible to recognize that this classifier is a great choice for similar tasks, once it has obtained great accuracy, precision, and recall levels. From the results, it can be concluded that this methodology is useful to conduct observations about the Oscar nominated movies and by using this method, the prediction of which movie will be the winner of the ceremony and which ones will be among the losers is very likely to be correct.

It can also be concluded that a big amount of Twitter users prefer to post positive comments about movies instead of tweeting hateful comments about the movies they do not like. The users were very excited and active on the social network during the week when the 2017 Oscars nominees were announced, but the excitement decreased with time then increased moderately on the weeks closer to the ceremony.

However, only moderate and weak mathematical associations were found between the proposed 2017 Oscars ranking and the other rankings based on the classifier results. This means that the data collected must be more deeply interpreted in order to get conclusions that are more satisfactory instead of using only mathematical interpretations. Another explanation would be that the movies that please the Twitter audience not always will be chosen as the best ones by the Academy of Motion Picture Arts and Sciences – for example, the movie "Hidden Figures"; it was very acclaimed by Twitter users, but did not win any awards and only got three nominations.

In this paper, the title of each movie nominated for Best Picture of the 2017 Oscars was used as keyword when collecting the tweets. A deeper analysis might be made considering the other categories of the award show, making it possible to predict the sentiment of Twitter users regarding the best actors, actresses, directors, songs, etc. Besides, a more meticulous prediction might be performed by obtained a more resourceful database, also including comments published on specialized social networks – like IMDb –, critics' reviews, etc.

Another manner of comparing the results of the classifier and the Oscars result would be by analyzing certain dates, for instance, considering only the tweets publish on the week when the nominees were announced or only the ones posted on the week before the ceremony. A timeline could be built, demonstrating how the users are expressing themselves across the different dates.

A different approach would be considering only the hashtags present on each tweet and verifying if they contain a certain sentiment that reflects on the respective tweet.

The user's profile could also influence on the quality of the information. In a future work, it would be interesting to analyze these different profiles that exist on the social networks and consider only the tweets that were published by users with the desired profile.

Furthermore, the methodology used in this paper can also be applied to future award shows or other situations in which it is desired to obtain an overview of Twitter user opinions regarding certain topics. Developing this study before the Oscars result (or other award show), drawing conclusions about the ceremony, and then observing if the analysis corresponds to the ceremony's result would also be an appealing approach to this methodology.

Additionally, performing this study using a database composed by tweets posted after the Oscars would be an interesting way of observing the sentiment of Twitter users regarding the winners of the ceremony.

References

[1] D. Teixeira and I. Azevedo, "Análise de opiniões expressas nas redes sociais," *RISTI - Revista Ibérica de Sistemas e Tecnologias de Informação*, vol. 8, pp. 53–65, 2011.

[2] F. Benevenuto, J. M. Almeida, and A. S. Silva, "Explorando Redes Sociais Online: Da Coleta e Análise de Grandes Bases de Dados às Aplicações," in *Livro Texto de Minicursos - SBRC 2011*, Campo Grande, MT, Brazil, 2011, ch. 2, pp. 63–101.

[3] Twitter, "Company," Apr. 2017. [Online]. Available: https://about.twitter.com/company

[4] E. Bothos, D. Apostolou, and G. Mentzas, "Using Social Media to Predict Future Events with Agent-Based Markets," *IEEE Intelligent Systems*, vol. 25, no. 6, pp. 50–58, 2010.

[5] V. Wong, "How Oscar Nominations Affect the Box Office," Apr. 2013. [Online]. Available: https://www.bloomberg.com/news/articles/2013-01-10/how-oscar-nominations-affect-the-box-office

[6] J. Krauss, S. Nann, and D. Simon, "Predicting movie success and academy awards through sentiment and social network analysis," in *16th European Conference on Information Systems*, 2008, p. 12.

[7] A. Pak and P. Paroubek, "Twitter as a Corpus for Sentiment Analysis and Opinion Mining," in *Proceedings of the Seventh Conference on International Language Resources and Evaluation*, 2010, pp. 1320–1326.

[8] I. L. Stats, "Twitter Usage Statistics," Apr. 2017. [Online]. Available: http://www.internetlivestats.com/twitter-statistics/

[9] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Found. Trends Inf. Retr.*, vol. 2, no. 1-2, pp. 1–135, Jan. 2008.

[10] B. Liu, *Web Data Mining*, 2nd ed. Chicago, IL, USA: Springer, 2011.

[11] A. Cetinsoy, "Predicting the 2017 Oscar Winners with BigML," Nov. 2017. [Online]. Available: https://dzone.com/articles/predicting-the-2017-oscar-winners

[12] N. Felix, "Análise de sentimentos em textos curtos provenientes de redes sociais," Ph.D. dissertation, Universidade de São Paulo - São Carlos, São Paulo, SP, Brazil, 2016.

[13] L. B. Ribeiro, "Análise de sentimento em comentários sobre aplicativos para dispositivos móveis: Estudo do impacto do pré-processamento," 2015.

[14] R. Baeza-Yates and B. Ribeiro-Neto, *Recuperação de Informação: Conceitos e Tecnologia das Máquinas de Busca*, 2nd ed. Porto Alegre, RS, Brazil: Bookman Editora, 2013.

[15] V. F. Schmitt, "Uma Análise Comparativa De Técnicas De Aprendizagem De Máquina Para Prever a Popularidade De Postagens No Facebook," 2013.

[16] DeepDive, "Distant Supervision," Nov. 2017. [Online]. Available: http://stanford.io/2Br4o1E

[17] A. Go, R. Bhayani, and L. Huang, "Twitter Sentiment Classification using Distant Supervision," *Processing*, vol. 150, no. 12, pp. 1–6, 2009.

[18] S. Loria, P. Keen, M. Honnibal, R. Yankovsky, D. Karesh, E. Dampsey, W. Childs, J. Schnurr, A. Qalieh, L. Ragnarsson, J. Coe, A. Calvo, N. Kulshrestha, J. Eslava, and J. Albert, "TextBlob: Simplified Text Processing," Aug. 2017. [Online]. Available: http://textblob.readthedocs.io/en/dev/index.html

[19] T. De Smedt and W. Daelemans, "Vreselijk mooi! (terribly beautiful): A Subjectivity Lexicon for Dutch Adjectives," *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*, pp. 3568–3572, 2012.

[20] IMDb, "Big Five Oscar nominees," Nov. 2017. [Online]. Available: http://www.imdb.com/list/ls075321784/

[21] J. P. Guilford, "Fundamental statistics in psychology and education," *Science Education*, vol. 41, no. 3, p. 565, 1957.