Sentiment Analysis of President Trump's Tweets: From Winning the Election to the Fight against COVID-19

Preliminary communication, DOI 10.22522/cmr20210272, received on 17 September 2021
UDK: 001.82:159.942
004.738(049.1)

Mihael Kampić
Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb, Croatia.

Marina Bagić Babac (corresponding author)
Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb, Croatia
E-mail: marina.bagic@fer.hr

Abstract

Twitter, as one of the popular social networks today and big data generator, can affect and change the public discourse, so political candidates are using it extensively as the vehicle for attracting and keeping their followers. Since Donald Trump's 2016 presidency election, his Twitter account with millions of followers has become an important subject for various statistical analyses, mostly because of his controversy. Therefore, this paper uses sentiment analysis of a large set of his tweets to explore his influence, as well the set of affective and cognitive aspects of his messages. The results of this analysis indicate what kind of findings in political domain can be recognized from tweets, and how they can be interpreted.

Keywords: social media, sentiment analysis, sentiment lexicon, President Trump, Twitter, tweets

1. Introduction

Forecasting opinion trends from real-time social media is the long-standing goal of modern day big-data analytics (Li et al., 2019). Despite its importance, there has been no conclusive scientific evidence so far that social media activity can capture the opinion of the general population at large. In this paper, an exploratory analysis provides us with various information about the reactions of public on President Donald Trump's tweets.

With the increasing importance of Twitter in political discussions (Smith, 2019), several studies have shown that it is possible to even predict political elections from Twitter data. The social media platform Twitter is one of the popular platforms where people post their real time experiences and opinions on various day-to-day issues used by researchers for various types of analytics. Recently, political candidates are using it extensively as the vehicle for attracting and keeping followers and thus influencing election outcomes (Bagić Babac, Podobnik, 2018). The overwhelming uses of Twitter by political candidates makes tweets an important domain for predictive analytics (Liu, 2015).

The goal of this study was to perform a sentiment analysis of a large set of Donald Trump's tweets to explore his influence, as well as the set of affective and cognitive aspects of his messages. Our results support the argument that data from social media can be considered as a reliable predictor of events in political campaigns.

2. Related work

Since Donald Trump's 2016 presidency election, his Twitter account has become an important subject for various statistical analyses, mostly because of his controversy (Nithyanand, Schaffner, Gill, 2017). The following works are closely related to the topic of this paper, either in research agenda, or the methodological approach to the analysis.

2.1. Using social media to predict election results

Bovet, Morone and Makse (2016) developed analytic tools combining statistical physics of complex networks, percolation theory, natural language processing and machine learning classification to infer the opinion of Twitter users regarding the candidates of the 2016 US Presidential Election. Using a large-scale dataset of 73 million tweets collected from June 1 to September 1, 2016, they investigated the temporal social networks formed by the interactions among millions of Twitter users. They inferred the support of each user to the presidential candidates and showed that the resulting Twitter trends follow the New York Times National Polling Average, which represents an aggregate of hundreds of independent traditional polls, with high accuracy of 0.9. More importantly, the Twitter opinion trend forecasted the aggregated New York Times polls by 6 to 15 days, showing that Twitter could be an early warning signal of global opinion trends at the national level. These analytics unleashed the power of Twitter to predict social trends from elections, brands to political movements, and at a fraction of the cost of national polls.

For these analytics, the team developed a supervised learning approach to classify the tweets as supporting or opposing the political candidates. Contrary to previous studies, they did not try to classify tweets as expressing positive or negative sentiment. Instead, they classified the tweets as supporting or opposing one of the candidates. By doing that, they avoided the problem of correctly assigning the object of the sentiment of a tweet. Indeed, a tweet containing a mention of Donald Trump and expressing a negative sentiment might be expressing opposition to Donald Trump as well as support. In this case, the context of the tweet was extremely important.

Another study on the use of data from social media platforms such as Twitter for predicting political elections (Watts et al., 2016) have raised many questions as well as created the interest in using Twitter data for predictive analysis. The overarching objective of this paper is to study the capability of Twitter data as an ex-ante indicator of event outcomes by modeling the momentum of political campaigns. Three indicators with predictive capability are proposed as measures of momentum of political campaigns. An asset price model was adapted to model momentum of candidates. Empirical validation was provided based on Twitter data from the 2014 US midterm election and the 2016 Presidential primary elections. These observations on momentum can be compared against the behavior of asset prices leading to bubbles showing analogy. In this work, similarities are drawn between winning campaign momentum and asset price bubbles until they burst. It was proposed adopting an asset price bubble model to model campaigns of political candidates using Twitter data.

Then, the model could be used to predict the outcome of the election if the election were held at the end of the analysis period. To this team's knowledge, there were no previous attempts to apply financial asset price models to model political campaigns. Nor was there a previous attempt to define indicators to measure and compare momentum of campaigns. For all practical purposes, they were proposing a methodology to analyze Twitter data that had not been previously attempted. After that, they conducted empirical analysis with the objective to validate the indicators using twitter data related to US elections. However, they were unable to identify any other models for comparison or benchmarking. Therefore, they have used Primary voting results following the periods of analysis related to the 2016 US Presidential primaries as a supporting data for validation.

2.2. The role of sentiment during elections on social media

Emotions are an inseparable part of how people use social media (Lin et al., 2012). While a more cognitive view on social media has initially dominated the research looking into areas such as knowledge sharing, the topic of emotions and their role on social media is gaining increasing interest (Bao et al., 2021). As is typical to an emerging field, there is no synthesized view on what has been discovered so far and, more importantly, what has not been.

Hyvärinen and Beck (2018) provided an overview of research regarding expressing emotions on social media and their impact and makes recommendations for future research in the area. Considering differentiated emotion instead of measuring positive or negative sentiment, drawing from theories on emotion (Ekman, 1992), and distinguishing between sentiment and opinion could provide valuable insights in the field.

Traditionally, sentiment analysis has measured the positive and negative sentiment of a sentence or longer text (Bradley and Lang, 1994), but there are recent examples of using more fine-grained approaches based on emotion categories (Bagić Babac, Podobnik, 2016). There are two main methodological approaches. Lexicon based methods (Staiano, Guerini, 2014) utilize a dictionary of words and their sentiment values, most often positive and negative to assign a sentiment score to an input text, whereas machine learning approaches (Jurafsky, Martin, 2015) classify documents into sentiment categories based on training data. Some recent studies combine the two by using lexicon scores as input for a classifier. Based on the results of these insights (Mohammad, Turney, 2013), theories on emotion are infrequently used to support the research, key terms such as sentiment, emotion and opinion are not always defined precisely, and sentiment analysis is mostly limited to measuring positivity and negativity instead of considering differentiated emotions.

For the 2016 US Presidential election, many people expressed their likes or dislikes for a particular presidential candidate. Joyce and Deng (2017) calculated the sentiment expressed by these tweets, and then compare this sentiment with polling data to see how much correlation they share. For this research, a lexicon and Naive Bayes Machine Learning Algorithm were used to calculate the sentiment of political tweets collected one-hundred days before the election. Manually labeled tweets as well as automatically labeled tweets were used based on hashtag content/topic. The end results suggest that Twitter is becoming a more reliable platform in comparison to previous work. By focusing on tweets 43 days before the election (beginning with the first presidential debate), the students found a correlation as high as 94% to polling data using a moving average smoothing technique.

Political campaigns have exploited this vast array of information available on the above platforms to draw insights about user opinions and thus design their marketing campaigns. Huge investments by politicians in social media campaigns right before an election along with arguments and debates between their supporters and opponents only enhance the claim that views and opinions posted by users have a bearing on the results of an election. Various sentiment analysis algorithms can be used to identify the attitude of an election candidate or a political party.

Sentiment analysis on social media data has been seen by many as an effective tool to monitor user preferences and inclination. Popular text classification algorithms like Naive Bayes and Support Vector Machines are supervised learning algorithms require a training dataset to perform sentiment analysis. The accuracy of these algorithms is contingent upon the quantity as well as the quality (features and contextual relevance) of the labeled training data. Since most applications suffer from lack of training data, they resort to cross domain sentiment analysis which misses out on features relevant to the target data. This, in turn, takes a toll on the overall accuracy of text classification. Ramteke et al. (2016) proposed a two-stage framework which can be used to create a training data from the mined Twitter data without compromising on features and contextual relevance. They proposed a scalable machine learning model to predict the election results using their two-stage framework.

Heredia, Prusa and Khoshgoftaar (2018) aimed to determine the effectiveness of using location-based social media to predict the outcome of the 2016 presidential election. To this aim, they created a dataset consisting of approximately 3 million tweets ranging from September 22 to November 8 related to either Donald Trump or Hillary Clinton. Twenty-one states were chosen, with eleven categorized as swing states, five as Clinton favored and five as Trump favored. They incorporated two metrics in polling voter opinion for election outcomes: tweet volume and positive sentiment. Their data is labeled via a convolutional neural network trained on the sentiment140 dataset[1]. To determine whether Twitter is an indicator of election outcome, they compared their results to the election outcome per state and across the nation. They used two approaches for determining state victories: winner-take- all and shared elector count. The results showed that tweet sentiment mirrors the close races in the swing states; however, the differences in distribution of positive sentiment and volume between Clinton and Trump were not significant using their approach. Thus, at the end of the research the team concluded neither sentiment nor volume is an accurate predictor of election results while using their collection of data and labeling process.

3. Research Methodology

3.1. Data Collection and Preprocessing

Twitter may not be the most popular social network, but it has a significant role in area of politics, especially in the US (Smith, 2019). On Twitter, there are 330 million monthly active users and 145 million daily users. There are 500 million tweets sent each day, that is 6,000 tweets every second.

Based on the US accounts, 10% of users write 80% of tweets, which makes an account of person like current US president Donald Trump an interesting subject for various statistical analyses. Also, in 2019, Donald Trump was the 2nd most influential person on Twitter. In 2020, Donald Trump had 71.1 million followers on Twitter, which made his account 10th on the list of accounts with most followers.

The dataset used in this study can be found online, under the name The Trump Twitter Archive (Brown, 2019), and it contains all 44,500+ tweets from the @realDonaldTrump

[1] https://www.kaggle.com/kazanova/sentiment140

Twitter account from 2009. Since this paper is about sentiment analysis of Donald Trump's tweets, retweets are not of much use, so they have been disabled while generating export file. To continue, not all of data is used in this sentiment analysis, so the features that are of use are:

1. Tweet Id - every database needs to have primary keys
2. Text - the context of tweet
3. Created at (GMT) - date of tweet posting
4. Favorite count - number of people who liked the specific tweet

The delimiter used for splitting features in the file is ",". Number of tweets downloaded is 36,168, with the oldest tweet being from 4 May 2009, and the most recent one from 5 June 2020. That makes up for a total of 11 years, 1 month and 1 day.

The first step in dataset processing was to read the file that contains the tweets obtained from Trump Twitter Archive. After importing the dataset, its context (text) is extracted. During the preprocessing stage, certain noise in data that could wrongly affect the sentiment analysis, was removed. Before removing unwanted strings, the whole text was converted to lower case, and then unwanted strings were substituted with blank spaces, thus, removing them. Some of the things that were removed are links (http\\...), at symbols (@), punctuation marks (".", ",", "?","!", ":", ";", ...), numbers, tabs, leading and trailing whitespaces.

3.2. Data Processing using Sentiment Analysis

In this paper, both data preprocessing and processing is implemented using R programming language. More specifically, R libraries used for these purposes are: *readr, plyr, stringr, stringi, magrittr,dplyr, tm, RColorBrewer, plotly*. The package used for sentiment analysis is *sentimentr*.

After the dataset preprocessing, the next step was to load opinion lexicons. There are two opinion lexicons: positive and negative (Mohammad, 2013). The positive opinion lexicon contains a list of 2,006 positive opinion words (or sentiment words), while the negative opinion lexicon contains a list of 4,783 negative opinion words (or sentiment words).

However, it should be noted that the appearance of an opinion word in a sentence does not necessarily mean that the sentence expresses a positive or negative opinion.

It is also worth noting that there are many misspelled words in the lists. They are not mistakes; they are included as these misspelled words appear frequently in social media content.

4. Results of Sentiment Analysis

4.1. Analysis of the tweets before the onset of COVID-19

For the purposes of this study, sentiment analysis was performed by implementing an R function which calculates a sentiment score from a given tweet. This function takes text of tweets, with positive and negative words as arguments. The function gets tweets as a vector of sentences. First, each sentence is cleaned up, i.e., punctuation marks, control characters and digits are removed. The sentence is converted to lower case and split into words. Then, words from tweets were compared to positive and negative words from the positive and negative opinion lexicons, and the position of matched term is retrieved, or NA *(Not Available)* if not found. NA is not useful for this analysis, so all NA matches were ignored in this sentiment analysis.

The score of a tweet is a sum of positive matches subtracted with sum of negative matches. There are three possible scenarios. A tweet is positive its score is greater than zero, negative if the score is less than zero and neutral if the sentiment score equals zero. The obtained score distribution is shown in Fig 1.

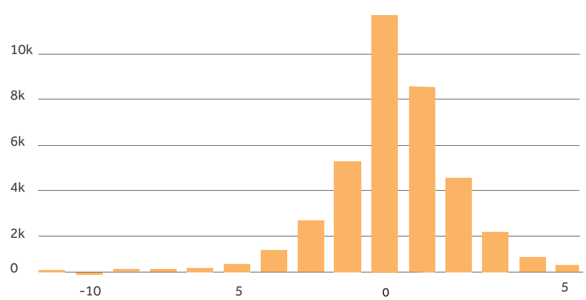

Figure 1. Score distribution of tweets

The most negative tweet had the score of -11 (Table 1), while the most positive tweet had the score of 12 (Table 2).

Table 1. Distribution of negative and neutral tweets

Score	Tweets
-11	1
-10	0
-9	1
-8	7
-7	21
-6	60
-5	156
-4	352
-3	968
-2	2333
-1	5034
0	11650

The most negative tweet was from August 20, 2018, and it says:

"Where's the Collusion? They made up a phony crime called Collusion, and when there was no Collusion, they say there was Obstruction (of a phony crime that never existed). If you FIGHT BACK or say anything bad about the Rigged Witch Hunt, they scream Obstruction!"

The most positive tweet is from April 8, 2018, and it says:

"Congratulations to Patrick Reed on his great and courageous MASTERS win! When Patrick had his amazing win at Doral 5 years ago, people saw his great talent, and a bright future ahead. Now he is the Masters Champion!"

Table 2. Distribution of positive tweets

Score	Tweets
12	1
11	0
10	0
9	1
8	5
7	28
6	84
5	261
4	680
3	1798
2	4280
1	8447

After performing sentiment analysis, tweets were grouped into three groups: positive tweets, negative tweets, and neutral tweets (tweet is negative if it has a score < 0, positive if score > 0, neutral if score = 0). Overall, there were 8933 negative tweets, 11650 neutral and 15585 tweets.

To gain better insight into the quality of these tweets, words from tweets were observed. To get the frequency of words in a collection of positive tweets, the following steps were taken.

First, a text corpus was created, that is a collection of text document over which text mining and natural language processing routines can be applied to derive interfaces. A special function interprets each element of the vector of positive tweets as a document. Next step is creating a term document matrix. Only words of lengths from 3 to 20 are included. Stop words like "the", "a" and "amp" are avoided, and numbers are not included. The matrix created is two dimensional. The total number of columns is equal to the number of positive tweets, where each number represents a tweet. Total number of rows is equal to number of unique words in all positive tweets. That means that every column represents one positive tweet and the words it contains. Cells are represented with 0 and 1 such that if a cell has a value 0, then that specific tweet does not contain that word, or if a cell's value is 1, then that specific tweet contains that word.

After creating the matrix, word counts are calculated (by rows) and sorted in a decreasing order. Data frame is created, and bar chart is plotted as shown in Fig. 2. Bar chart of positive tweets consists of 15 most mentioned words in tweets that had a positive sentiment score. The x- axis represents a word, while the y-axis represents the number of times that word has been mentioned in a collection of tweets.

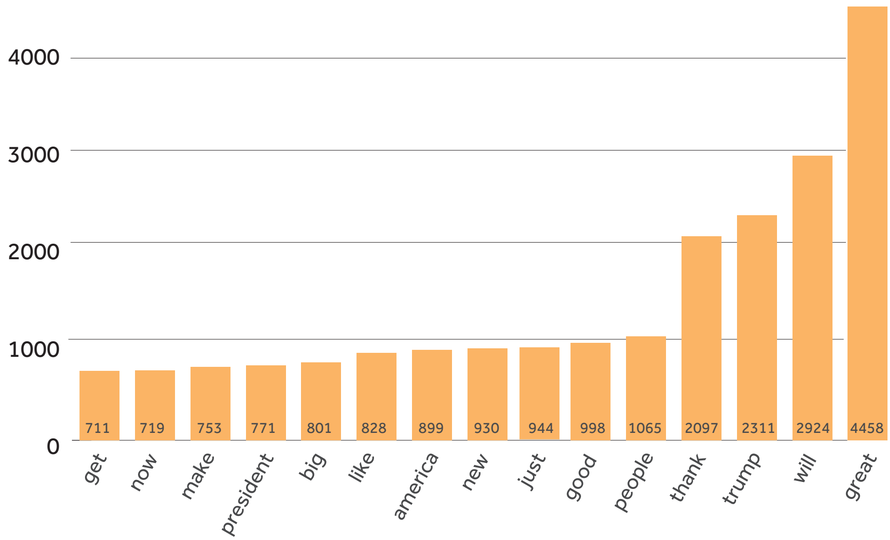

Figure 2. Distribution of the most mentioned words in positive tweets

In positive tweets, the most mentioned word is "great", which is expected outcome since that word is one of the most used in communication while expressing positive thoughts, and it is covered in the positive lexicon that was used. It's also worth noting that Donald Trump's surname is mentioned in a lot of positive tweets. The word "America" is also mentioned quite a lot, which shows that Trump usually links his country to thoughts of positive context.

Bar chart of negative tweets consists of 15 most mentioned words in tweets that had a negative sentiment score as shown in Fig. 3.

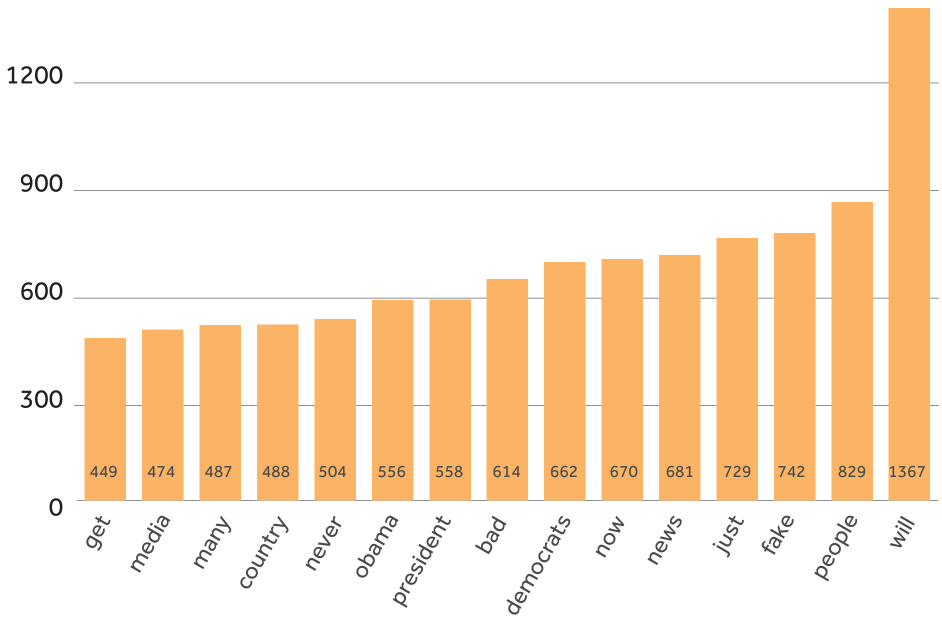

Figure 3. Distribution of the most mentioned words in negative tweets

In negative tweets, the lexicon of negative words had a lot more words, but despite that the most common word is "will" and "people". Other words that were common are "fake" and "news", which is a term Trump often uses to criticize the media, which is also commonly used word in negative tweets. Since the tweets covered the times when Barack Obama was the president of the United States, he is mentioned a lot in negative tweets, which means Trump used to criticize him a lot. Another word which is mentioned a lot is "democrats", which is logical since Trump is of different political orientation (republican).

Bar chart of neutral tweets consists of most mentioned words in tweets that had a sentiment score of 0 (Fig. 4).

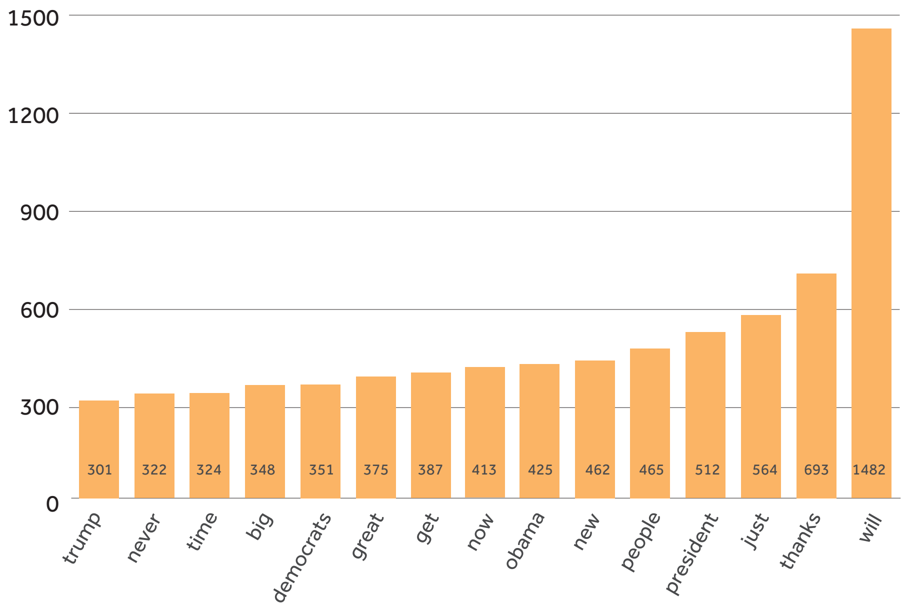

Figure 4. Distribution of the most mentioned words in neutral tweets

In neutral tweets, the most common words are "will", "thanks", "just" and "president". It is also worth noting that some words from most mentioned words in positive and negative tweets also appeared in neutral, since neutral tweets are tweets with equal amount of positive and negative words. Some of those words are: "Obama", "democrats", "great", "trump".

After analyzing the most mentioned words, it is safe to conclude that there is a slight correlation of some words to their sentiment.

The word "will" appeared a lot of times in all types of tweets since that word was not covered in the lexicons and it can be used in various context since will can be "good" and "bad". The same applies to "people", as the people can be linked to positive and negative meanings.

Next, from the time analysis of sentiment scores, average sentiment score of tweets by year can be observed from Fig. 5.

Figure 5. Time overview of annual sentiment scores

In the first two years of using Twitter, Trump had his best positive annual sentiment score. Then, in 2011 he reached his lowest sentiment score, and in the following couple of years his annual sentiment score was fluctuating.

Over the period of 11 years, 1 month and 1 day, his annual sentiment score was never negative, which means that even though Trump has his bad tweets, on average, they are of positive sentiment.

Most of President Trump's tweets fall into category of positive tweets (43,1%), which is different than most people would expect. There is a higher number of neutral tweets (32,2%) than negative (24,7%), which may be due to the chosen method of performing sentiment analysis with sentiment lexicons.

4.2. Analysis of the COVID-19 related tweets

In order to perform sentiment analysis related to COVID-19 pandemic, there are some parameters that need to be established. The oldest tweet used for this analysis was created January 24, 2020., while the most recent tweet was created May 28, 2020.

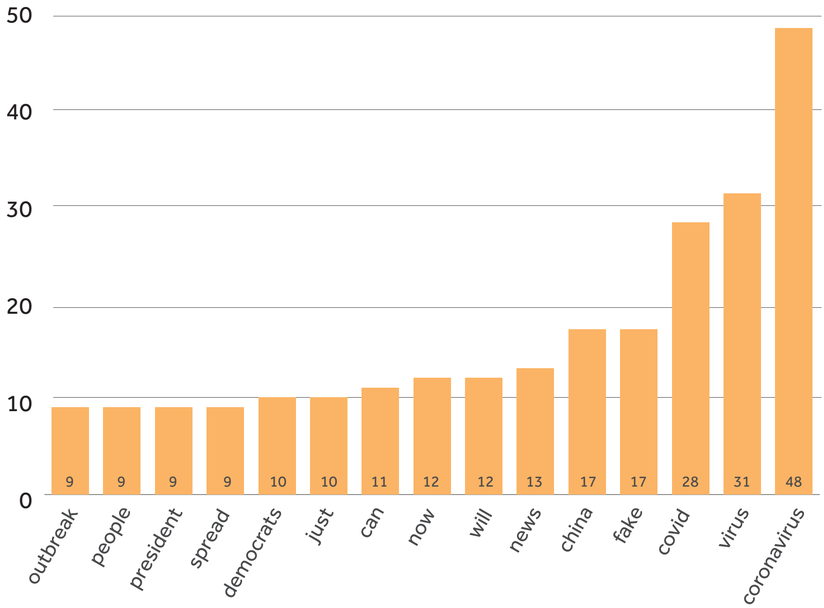

Figure 6. Distribution of words in COVID-19 negative tweets

The tweets were filtered by keywords: "corona", "covid" and "virus". That type of filtering left 407 tweets as a result. The results were following: 38,3% positive, 34,4% neutral, 27,3% negative. When it comes to tweets related to the pandemic, most of them are neutral, lesser amount is positive and the least amount are negative tweets

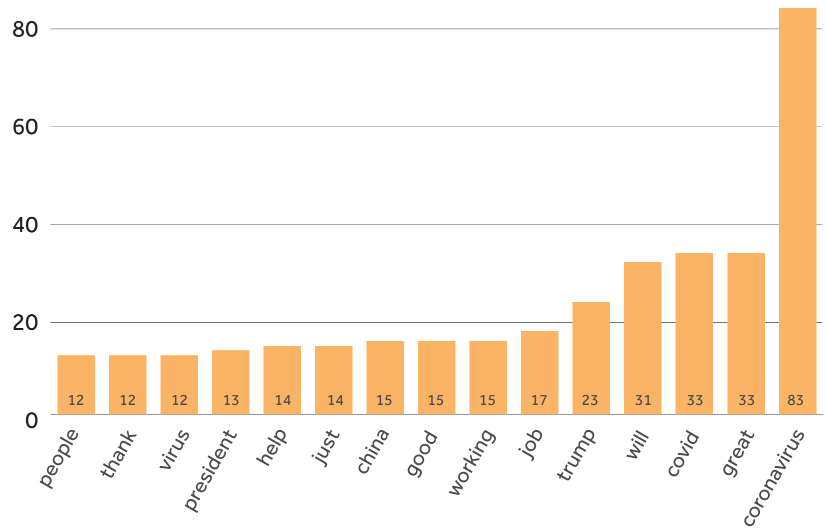

Figure 7. Distribution of words in COVID-19 positive tweets

As the results from Figs. 6 and 7 show, the most common words in both cases are "coronavirus", "covid" or "virus", as expected. Trump often linked the coronavirus with China, both in negative and positive context. He also claimed that the virus was fake, and it was not surprising that the democrats once again showed up in the negative context relating to the COVID-19 pandemic.

Words like "spread" and "outbreak" showed up in negative tweets which is expected because of the fear that pandemic caused. Words like "job" and "working" were brought up in positive tweets which correlates with the fact that he was thanking to everyone who did a good job suppressing the virus.

The most positive tweet related to COVID-19 is:

"Just had a long and very good conversation by phone with President Xi of China. He is strong, sharp, and powerfully focused on leading the counterattack on the Coronavirus. He feels they are doing very well, even building hospitals in a matter of only days. Nothing is easy, but…"

The most negative tweet related to COVID-19 is:

“Some wackos in China just released a statement blaming everybody other than China for the Virus which has now killed hundreds of thousands of people. Please explain to this dope that it was the “incompetence of China”, and nothing else, that did this mass Worldwide killing!”

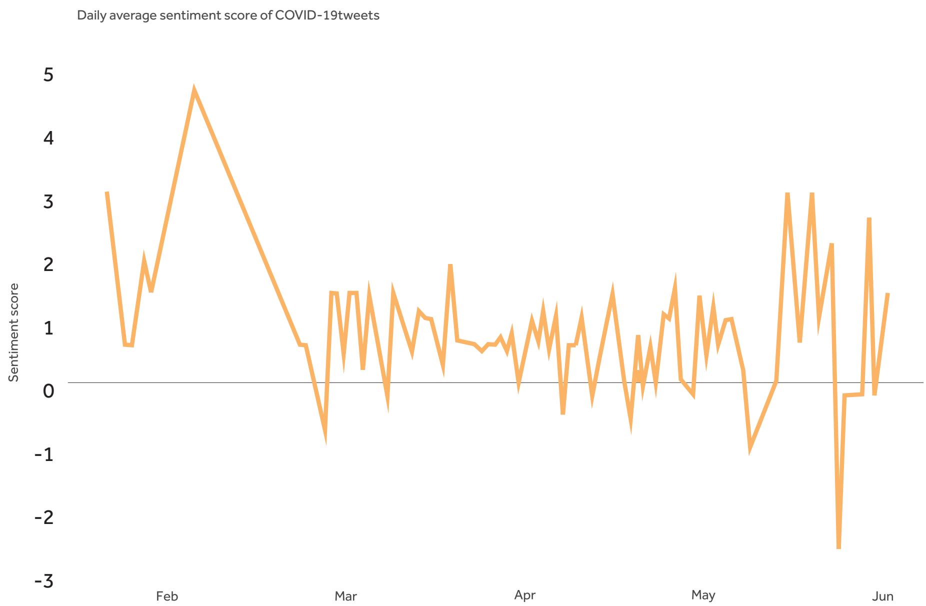

Figure 8. Time sentiment of COVID-19 tweets

Fig. 8 shows the daily average sentiment score of Trump’s tweets related to the COVID-19 topic. In first couple of months, Trump’s tweets had positive and neutral score, which is due to his opinion that the virus is fake and there is nothing to worry about. In the following few months, the coronavirus outbreak was declared a pandemic (March 11, 2020). Around that point, the sentiment score started to fluctuate, and in May the pandemic situation escalated beyond control in the USA. As the USA president, Trump was the responsible person, which resulted in his tweets being about throwing the blame on others and trying to give positive recognition to the people who fight the pandemic on the front lines.

5. Conclusion

Social media is changing the public discourse and has become a platform for generating and consuming information for the masses (Bovet et al., 2016). There are several social media platforms accessible to the world population. They have become obvious examples of big data generators.

Sentiment analysis represents the use of natural language processing to quantify attitudes about a certain topic (Guerini, Staiano, 2015). Marketers often use sentiment analysis to understand how the public feels toward their brand (Bagić Babac, Podobnik, 2016).

The social media platform Twitter is one of the popular platforms where people post their real time experiences and opinions on various day-to-day issues used by researchers for various types of analytics. Recently, political candidates are using it extensively as the vehicle for attracting and keeping followers and thus influencing election outcomes. The overwhelming uses of Twitter by political candidates makes tweets an important domain for predictive analytics.

This paper showed results of the performed sentiment analysis with lexicon-based approach, relying on collections of pre-compiled terms that are known to express a sentiment. According to the most frequently used positive (e.g., "america", "great") and negative (e.g., "fake", "news") words in President Trump tweets, it was concluded what topics and sentiment were of his greatest concern at a time. In 2011, he reached his lowest sentiment score, while over the course of years his annual sentiment score was fluctuating depending on the actual political context. From the results presented in this study, it is safe to conclude that over the course of years President Trump used his account for impacting others by linking various persons (Obama) or political groups (democrats) to specific sentiment meanings, which can be either positive or negative. Overall, it was concluded that most of President Trump's tweets fall into category of positive tweets, which is different than most people would expect.

However, certain limitations of the study should also be noted. First, Twitter language is known to be unstructured, informal and to contain various spelling mistakes, links, hashtags, emoticons, and usernames that limits the applicability of predefined lexicons (Bovet et al., 2016). Thus, more extensive data would provide more detailed and accurate insights into the attitudes of President Trump. On the other hand, lexicon-based approach is limited to the specific set of labelled words which may vary over different domains. Thus, a different approach that could be used for a more in-depth sentiment analysis is based on machine learning where a set of existing labeled documents, from which features are extracted, is used to classify the rest of the documents.

References

- Bagić Babac, M., Podobnik, V. (2018). What social media activities reveal about election results? The use of Facebook during the 2015 general election campaign in Croatia, *Information Technology & People,* 31 (2), 327-347. https://doi.org/10.1108/ITP-08-2016-0200
- Bagić Babac, M., Podobnik, V. (2016), A sentiment analysis of who participates, how and why, at social media sport websites: How differently men and women write about football, *Online Information Review*, 40 (6), 814-833. https://doi.org/10.1108/OIR-02-2016-0050
- Bao, S., Xu, S., Zhang, L., Yan, R., Su, Z., Han, D., Yu, Y. (2012). Mining Social Emotions from Affective Text, *IEEE Transactions on Knowledge and Data Engineering,* 24 (9), 1658-1670.
- Bazarova, N. N., Choi, Y. H., Sosik, V. S., Cosley, D., Whitlock, J. (2015). *Social Sharing of Emotions on Facebook: Channel Differences, Satisfaction, and Replies.* CSCW 2015, Vancouver, BC, Canada.
- Bovet, A., Morone F., Makse H. A. (2016), Predicting election trends with Twitter: Hillary Clinton versus Donald Trump, *ArXiv*, abs/1610.01587.
- Bradley, M. M., Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behaviour therapy and experimental psychiatry,* 25 (1), 49–59.
- Brown, B. (2019). *The Trump Twitter Archive.* Retrieved from: http://www.trumptwitterarchive.com/.
- Ekman, P. (1992). An argument for basic emotions, *Cognition & emotion,* 6 (3-4),169–200.
- Guerini, M., Staiano, J. (2015). Deep Feelings: A Massive Cross-Lingual Study on the Relation between Emotions and Virality, *Proceedings of the 24th International Conference on World Wide Web (WWW '15)*, 299-305.
- Heredia, B., Prusa, J., Khoshgoftaar, T. (2018). *Location-Based Twitter Sentiment Analysis for Predicting the U.S. 2016 Presidential Election,* Florida Artificial Intelligence Research Society Conference, North America.
- Hu, M., Liu, B. (2004). Mining and Summarizing Customer Reviews, *Proceedings of the ACM SIGKDD International Conference on Knowledge, Discovery, and Data Mining (KDD-2004)*, Seattle, Washington, USA.
- Hyvärinen, H., Beck, R. (2018). Emotions Trump Facts: The Role of Emotions in on Social Media: A Literature Review, *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2018, pp. 1797-1806.
- Joyce, B., Deng, J. (2017). *Sentiment analysis of tweets for the 2016 US presidential election,* 2017 IEEE MIT Undergraduate Research Technology Conference (URTC), Cambridge, MA, pp. 1-4.
- Jurafsky, D., Martin, J. H. (2015). *Speech and Language Processing. Chapter 21: Lexicons for Sentiment and Affect Extraction*
- Katz, E., Lazarsfeld, P. F. (1955). *Personal Influence: The Part Played by People in the Flow of Mass Communication.* Glencoe, IL: Free Press.
- Li, Z. Fan, Y., Jiang, B., Lei, T., Liu, W. (2019). A survey on sentiment analysis and opinion mining for social multimedia, *Multimedia Tools and Applications,* 78 (6), 6939-6967.
- Lin, C., He, Y., Everson, R. Ruger, S. (2012). Weakly Supervised Joint Sentiment-Topic Detection from Text, *IEEE Transactions on Knowledge and Data Engineering,* 24 (6), 1134-1145.
- Lin, R., Utz, S. (2015). *The emotional responses of browsing Facebook: Happiness, envy, and the role of tie strength, Computers in Human Behavior,* 52, 29–38.
- Liu, B. (2015). *Sentiment Analysis: mining opinions, sentiments, and emotions.* Cambridge University Press.
- Liu, B., Cheng, J., Hu, M. (2005). Opinion Observer: Analyzing and Comparing Opinions on the Web, *Proceedings of the 14th International World Wide Web conference (WWW-2005)*, Chiba, Japan.
- Mohammad, S. M. (2016). Sentiment Analysis: Detecting Valence, Emotions, and Other Affectual States from Text, *Emotion Measurement.* H. Meiselman (Ed.), Elsevier.
- Mohammad, S., Turney, P. (2013). Crowdsourcing a Word-Emotion Lexicon. *Computational Intelligence*, 29 (3), 436-465.
- Nithyanand, R., Schaffner, B., Gill, P. (2017). *Online Political Discourse in the Trump Era*; University of Massachusetts at Amherst, arXiv:1711.05303
- Parkinson, B. (1996). Emotions are social. *British Journal of Psychology,* 87, 663-683.
- Ramteke, J., Shah, S., Godhia, D., Shaikh, A. (2016). *Election result prediction using Twitter sentiment analysis,* 2016 International Conference on Inventive Computation Technologies (ICICT), Coimbatore, pp. 1-5.

- Rimé, B., Mesquita, B., Boca, S., Philippot, P. (1991). Beyond the emotional event: Six studies on the social sharing of emotion. *Cognition & Emotion,* 5 (5–6), 435–465.
- Sanford, K. (2007). Hard and soft emotion during conflict: Investigating married couples and other relationships, *Personal Relationships,* 14 (1), 65–90.
- Severin, W., Tankard, J. (1997). *Communication theories: Origins, methods, and uses in the mass media.* Longman.
- Smith, K. (2019). *Twitter user statistics; 60 Incredible and Interesting Twitter Stats and Statistics.* Retrieved from: https://www.brandwatch.com/blog/twitter-stats-and-statistics/ , 20.11.2019.
- Staiano, J., Guerini, M. (2014). DepecheMood: a Lexicon for Emotion Analysis from Crowd-Annotated News, *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics,* Baltimore, Maryland, USA, June 23-25, pp. 427–433.
- Van Kleef, G. A., De Dreu, C. K. W., Manstead A. S. R. (2004). The interpersonal effects of emotions in negotiations: A motivated information processing approach. *Journal of Personality & Social Psychology,* 87, pp. 510-528.
- Wang, E. (2017). *Trumps Greatest tweets - Dremio*, 25.1.2017. Retrieved from: https://www.dremio.com/trump-twitter-sentiment-analysis/, A sentiment analysis of the greatest tweets from the tweeter-in-chief, 21.11.2019.
- Watts, D., George, K. M., Kumar, T. K. A., Arora, Z. (2016). *Tweet sentiment as proxy for political campaign momentum*, 2016 IEEE International Conference on Big Data (Big Data), pp. 2475-2484, doi: 10.1109/BigData.2016.7840885.
- Zech, E., Rimé, B., Nils, F. (2004). Social sharing of emotion, emotional recovery, and interpersonal aspects. *The Regulation of Emotion,* 157–185.

Mihael Kampić

Mihael Kampić received the B.S. degree in Computer Science from the University of Zagreb, Faculty of Electrical Engineering and Computing, Croatia, where he is currently pursuing the M.S. degree in Computer Engineering. His research interests are in the field of artificial intelligence.

Marina Bagić Babac

Marina Bagić Babac is an Assistant Professor at Department of Applied Computing at the University of Zagreb, Faculty of Electrical Engineering and Computing, Croatia. She received the B.Sc., M.Sc. and Ph.D. degrees in Electrical Engineering from the University of Zagreb, Faculty of Electrical Engineering and Computing, Croatia, and M.Sc. in Journalism from the University of Zagreb, Faculty of Political Science. She is currently a collaborator on several international projects related to linguistic and linked data. She serves as a program committee member of a few international scientific conferences and as a reviewer in numerous international journals. Her research interests include natural language processing, linked data, social media, and social network analysis.