

Determining Reasons of Political Rating Changes Based on Twitter Data

Taras Rudnyk¹, Oleg Chertov¹

¹ National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", 37, Peremohy ave., Kyiv, 03056, Ukraine

Abstract

This paper presents the results of a study on determining the rating of politicians based on a dataset collected from Twitter, comparing it with opinion polls. It is shown in the example of the analysis of the rating of the President of Ukraine Volodymyr Zelenskyy for the period from January 2019 to March 2021 that the differences in the rating can reveal events that influenced its change. Based on the specific reasons for the drop in rating, you can provide recommendations on how to stop the drop. On the other hand, the search for rating growth reasons can be used to determine the ways of increasing the respective politician's rating. To avoid misleading information and to verify the accuracy, detected Twitter events were compared to Google Trends and their consistency was confirmed.

Keywords

political rating, sociological polls, Twitter, natural language processing, statistics, Google Trend, Ukraine, Volodymyr Zelenskyy.

1. Introduction

Politicians are carefully considering the ratings. A lot of money is spent on sociological polls. Their implementation requires the involvement of various specialists and institutes that conduct thorough research, spending a lot of time on it. But times are changing. The massive diffusion of social networks provides opportunities for researching electoral preferences of different categories of potential voters. An automated algorithm could significantly reduce the time, money and number of people involved in sociological polls.

Such polls often make it difficult to understand which events have affected the rise or fall of the rating. If the poll is conducted once a month, then during this period there are usually many events. Which of these events had a decisive influence on the opinion of a particular respondent can only be known if the authors of the poll have provided a corresponding specific question on such an event or a group of relevant events.

By researching social media data, it could be aggregated for different periods, it can be a month, a week, a day or even an hour. Statistical counting of the number of messages on the social network will provide important information, indicating which topics are most discussed. This responsiveness and flexibility allow us to highlight key events and recommend how to respond to them to improve society's response.

In this paper, we present the challenges and solutions in the following structure. Section 2 contains the literature review on topics like our study. Section 3 explains the approach to determining the political rating and the reasons for its changes. Section 4 presents the experiment, results, and discussion. The outcome is compared with sociological polls and Google Trends. Section 5 contains conclusions and future work.

XXI International Scientific and Practical Conference "Information Technologies and Security" (ITS-2021), December 9, 2021, Kyiv, Ukraine
EMAIL: tarasrudnyk@gmail.com (T. Rudnyk); chertov@i.ua (O. Chertov)
ORCID: 0000-0001-9492-0374 (T. Rudnyk); 0000-0003-0087-1028 (O. Chertov)



© 2021 Copyright for this paper by its authors.
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

2. Literature review

Perhaps, Walter Lippmann was the first to theoretically substantiate the influence of the classical mass media (press, radio, cinema, or, inferentially, television) on the political preferences of citizens back in 1922 [1]. Both classical and modern studies show that the media both influence the voters' preferences during elections [2] and produce more long-term effects: they influence the formation of a stable vision about parties [3], the formation of political coalitions after elections [4], etc.

The idea of using social networks to calculate political ratings also is not new. Interestingly, the first studies of this kind (see, for example, [5]) showed a low correlation between a politician's involvement in online activity and her/his rating. This result was obtained when analyzing the impact of social media on the US presidential election in 2012. But already in the next election in 2016, almost all experts associated the victory of Donald Trump, in particular, with his great activity on Twitter [6, 7]. Burnap et al. [8] sought to predict in advance the results of the general election in the UK in 2015, based on the recognition that "more tweets - more votes". Tumasjan et al. [9], based on the analysis of the elections, made almost the same conclusion, but they analyzed not only the number of tweets but also carried out their content-analysis of over 100,000 messages containing a reference to either a political party or a politician. Anuta et al. [10] conducted a sentiment analysis to see if social media could pave the way for less biased results than regular polls. Their findings suggest that, although numerical shifts are common in both approaches, using only social media for predictions may lead to less accurate predictions. Cameron et al. [11] explored whether a candidate's online presence could affect his chances of being elected. Using two regression models, they concluded that there is statistical, albeit small, significance between the number of people who follow or be friends with a politician on social media and the election results.

Researchers often use such popular social network as Facebook. Stephen R. Neely in [12] considers politically motivated unfriending or unfollowing on Facebook in the lead-up to the 2020 USA Presidential election. But due to several restrictions, it is much harder to collect data from this social network than from Twitter. It is Twitter that provides access to some data about its users and their actions, which potentially allows drawing reasonable conclusions about their electoral preferences.

The vast majority of articles that explore the relationship between the popularity of a politician or some political force and their social activity or the activity of their supporters/opponents in online social networks are based on an analysis of a snapshot of relevant messages. The first article in which such an analysis is done based on data collected over four years was published only last year [13]. Therefore, studies that operate with real data collected over a sufficiently long period (from six months), which allows you to accumulate and highlight the factors that characterize the electoral prospects of a certain politician, are of particular interest.

In social networks, users can act alone and even unite into groups. In [14] the approach to detect groups of phony accounts on Facebook was introduced. Chronological analysis of user messages allowed us to detect those who tried to influence other group members. Analyzing the electoral prospects of a particular political leader, it is advisable to study, first of all, the dynamics of change in that part of voters who unite around their leader in a social network group [15]. There are three main types of human bias that are manifested in social networks [16]: a tendency to support the opinion of an authority figure, filtering out only those events that confirm previous beliefs or values, moreover, events that contradict a person's opinion only increase her/his confidence in her/his rightness. It is obvious that a certain political force or a specific political leader, explicitly or implicitly, forms various support groups in social networks, through which they exert their influence both on their supporters and on the general mass of voters.

Objectives: This study sets in a certain sense the inverse problem and it intends (i) to confirm the possibility, based on the activity of some users of a social network during a sufficiently long period (from six months), to determine how popular this or that political force is in the electoral sense, and (ii) to find out and analyze: is it then possible to identify the events that led to a change in the corresponding political rating?

3. Approach to determining the political rating and the reasons for its changes

To determine the political rating of Ukrainian President Volodymyr Zelenskyy, the Twitter dataset from our previous research [15] was used. We have made a few changes compared to our previous article that has improved the calculated rating by 4%. The number of subscribers was no longer considered to find opinion leaders. It has been found that many accounts have so-called “dead” subscribers who were once registered but have not been active for a long time. These accounts can be old bots or just people who have stopped using Twitter. Most of these subscribers were seen on pages of politicians who have been in politics for a long time but did not win the latest election. For example, the page of the former Prime Minister of Ukraine – Arseniy Yatsenyuk. At the same time, some bloggers have relatively few subscribers, but almost all of them are active – like, retweet and comment on tweets. For example, the page of Sergei Sternenko.

Another improvement in the formula was the greater importance of retweets compared to likes. Therefore, we multiply them by an empirically selected coefficient equal to 2. The same multiplier for retweets was used by researchers of Donald Trump’s activity in [17]. The following is the final formula for subscriber activity calculation:

$$\text{followers}_{\text{activity}} = \text{average}_{\text{likes}} + 2 * \text{average}_{\text{retweets}} \quad (1)$$

The next important change in the formula is the use of the sum of the natural logarithms of each component instead of the product. This change allowed us to consider all the terms separately and apply the coefficient equal to 4 so that the influence of the account was not dominant compared to the other two terms. The coefficient was chosen empirically to obtain the best result. Since the tweet score can be less than one, instead of the usual logarithm we use the logarithm of 1 + tweet score.

$$\text{rating} = \sum_{n=1}^n \log(1 + \text{tweet}_{\text{score}}) + \log(\text{user}_{\text{activity}}) + \frac{\log(\text{account}_{\text{influence}})}{4} \quad (2)$$

To detect dates of opinion changes the data is arranged into shorter periods. Initially to weeks. Once the algorithm detects the week with anomaly rating changes – rapid growth or decline of the chart, then the weeks are split by days. For each period rating changes, which may be not only one day but several days in a row, statistics of the most common words were collected. For each word, except for stop words (a set of commonly used words in any language), a number of occurrences in popular news got calculated. The words which occurred in the news the most were called keywords and got stored as potentially important in terms of political rating impact. To avoid false keywords selected in the previous step they could be double-checked in Google trends. Once the dates of tweets and Google Trends fit then specific news that affected the rating received. On some dates, the political rating may be affected not only by one event. Sometimes ratings changed after several positive or negative news. Therefore, it is important to collect all popular news.

4. Experiment, Results and Discussion

4.1. Volodymyr Zelenskyy’s political rating compared to a sociological poll

The new approach allowed us to achieve results 4% better than with the original formula. The total deviation of the rating is 16%. Calculated ratings from Twitter data and sociological poll results are presented in Figure 1. Black bars represent the results of the algorithm, and the dashed grey line shows sociological polls’ result².

² https://ratinggroup.ua/files/ratinggroup/reg_files/rg_ukraine_covid_cati_ix_wave_022021_press.pdf

Volodymyr Zelenskyy support difference by date aggregated monthly

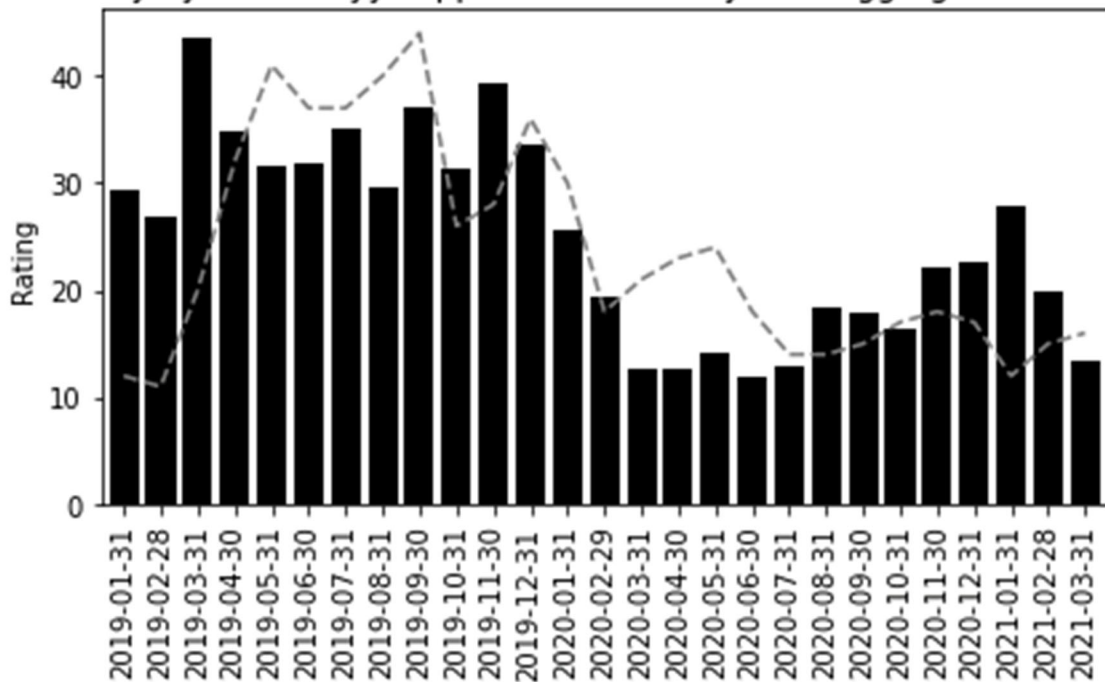


Figure 1: Volodymyr Zelenskyy support difference by date aggregated monthly

4.2. Identification of the reasons for the rating fall

To recognize the reasons for the fall in rating, data aggregation was carried out not by months, but by weeks. The first detected drop in the rating is a consequence of the beginning of the COVID-19 epidemic in Ukraine. Many people became ill, and on March 25th in the year 2020, the Cabinet of Ministers of Ukraine imposed a 30-day state of emergency across the country due to the spread of coronavirus disease. Examples of tweets with negative scores for observed dates are presented in Table 1. Calculated by proposed model rating changes for this period are presented in Figure 2.

Table 1
Negative tweets related to COVID-19 at the end of March in Ukraine

Created at	Text	Score
2020-03-21 19:39:44	A state of emergency has been declared in Kharkiv Oblast due to the COVID-19 pandemic.	-3
2020-03-27 10:52:57	To see on the screen the inaction of the authorities to prepare for the coronavirus is extremely saddening and indignant.	-5
2020-04-05 17:14:49	We are filing a lawsuit against the absolutely illegal decision of the Cabinet of Ministers, by which he actually introduced a state of emergency in the country, bypassing the President and the Verkhovna Rada. No "good intentions" can be the basis for violating the Constitution of Ukraine.	-2

Volodymyr Zelenskyy support difference by date aggregated weekly

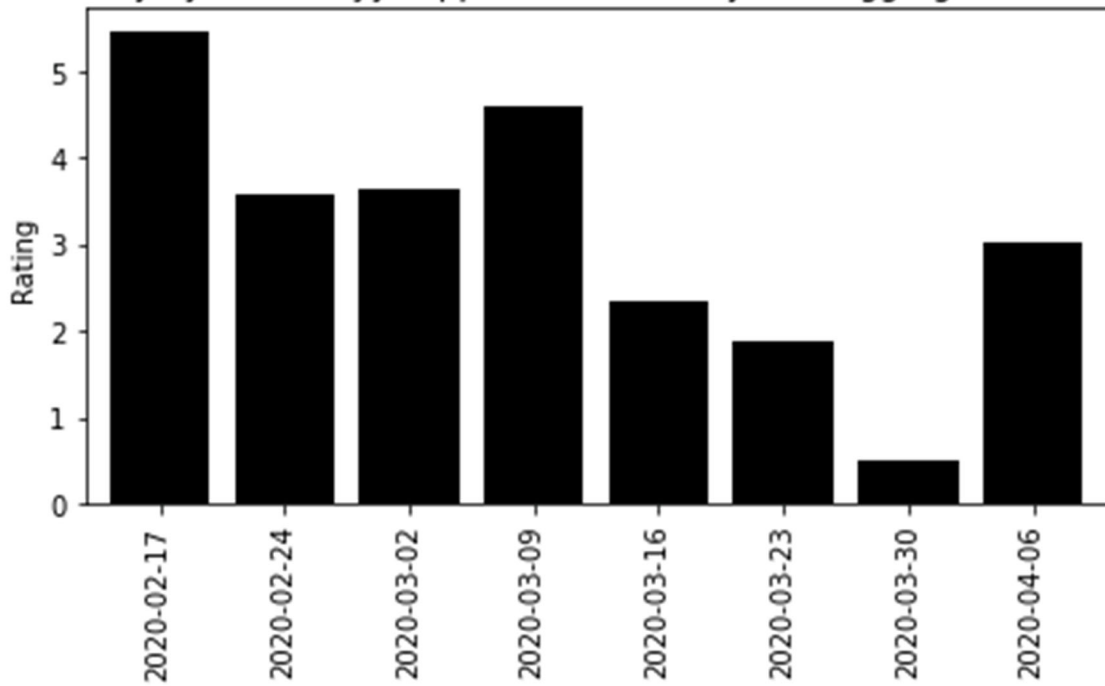


Figure 2: Volodymyr Zelenskyy rating from 02.03.2020 to 20.04.2020 calculated by the proposed model

Another rating loss was detected when on July 14, 2020, the Council legalized the gambling business. Rating changes for this period are presented in Figure 3.

Volodymyr Zelenskyy support difference by date aggregated weekly

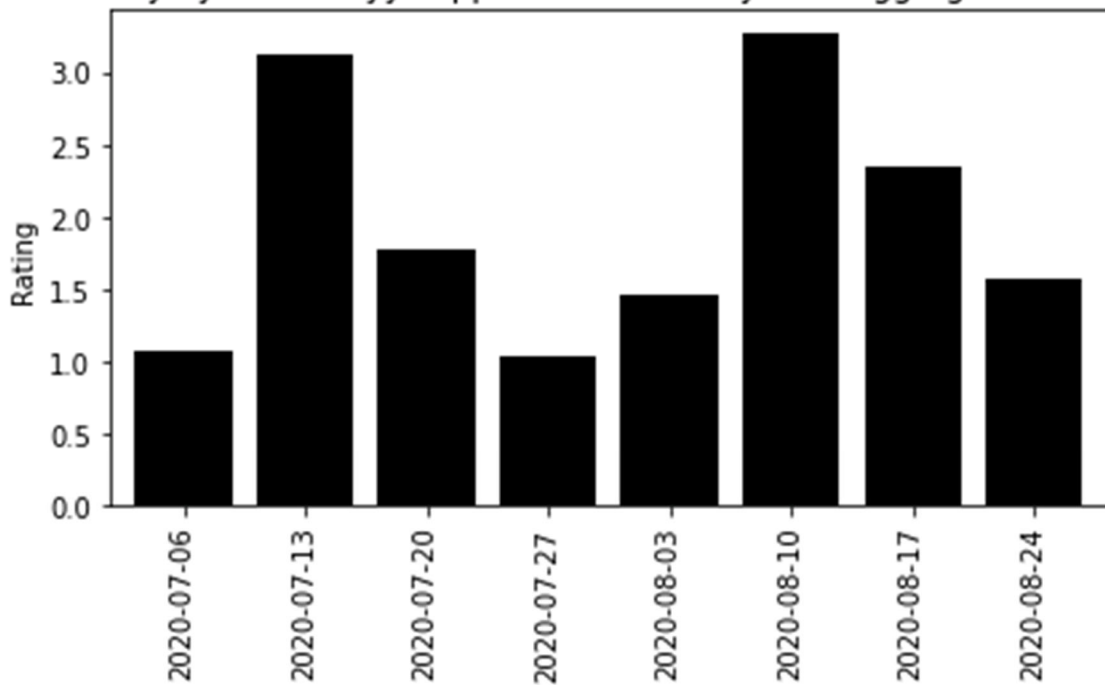


Figure 3: Volodymyr Zelenskyy rating from 22.06.2020 to 10.08.2020 calculated by the proposed model

4.3. Identification of the reasons for the rating growth

To recognize the reasons for the growth of rating, data aggregation was carried out not by months or weeks, but by days. The first detected growth of rating started on the 25th of January year 2021 and exposed interesting reasons on how algorithms could be improved. Volodymyr Zelenskyy was born on the 25th of January and received on that date and a few days after a lot of good congratulatory words. Our algorithms accidentally detected such behavior as the rating grows. To improve the proposed approach, the major dates from personal politician life such as birthdays, marriage or family celebrations should be removed.

The next rating growth was detected on the 2nd of February year 2021. On that date, The President of Ukraine Volodymyr Zelenskyy put into effect the decision of the National Security and Defense Council on the application of sanctions against the People’s Deputy Taras Kozak and the TV channels 112 Ukraine, NewsOne and ZIK, which were blocked. The deputy and his TV channels carried out anti-Ukrainian activities.

The last rating growth was detected on the 16th of February year 2021. This one is interesting because is a result of multiple news on the same date which is alone not so powerful as the previous one but gives a significant growth as an aggregated result. On that date:

- Volodymyr Zelenskyy had an official visit to the United Arab Emirates and agreed on agreements and memoranda worth more than three billion dollars, cooperation in various fields, from defense to agriculture, foreign direct investment in Ukraine and readiness to increase interstate trade several times. Examples of tweets related to this news are presented in the first two rows of Table 2;
- Volodymyr Zelenskyy announced the reduction of the powers of the Kyiv District Administrative Court. An example of a tweet related to the announcement is presented in the third row of Table 2;
- The speaker of parliament Dmitry Razumkov said that deputies will start to be left without mandates for “button-pressing”. An example of a tweet related to the speaker’s words is presented in the fourth row of Table 2.

Table 2
Examples of tweets with a positive score

Created at	Text	Score
2021-0-15 13:47:18	During the official visit of the President of Ukraine Volodymyr Zelenskyy to the United Arab Emirates, the Ukrainian delegation signed several bilateral documents.	4
2021-02-14 18:44:28	Olena Zelenska suggested intensifying cultural cooperation with the UAE. Among the initiatives are weeks of Ukrainian cinema and days of folk art in the Emirates.	3
2021-02-13 16:35:01	Zelenskyy announced the reduction of the powers of the Kyiv District Administrative Court.	1
2021-02-13 10:29:44	Razumkov said that the deputies will lose their mandates due to button-pressing	4

Rating changes for the period from 24.01.2021 to 01.03.2021 are presented in Figure 4.

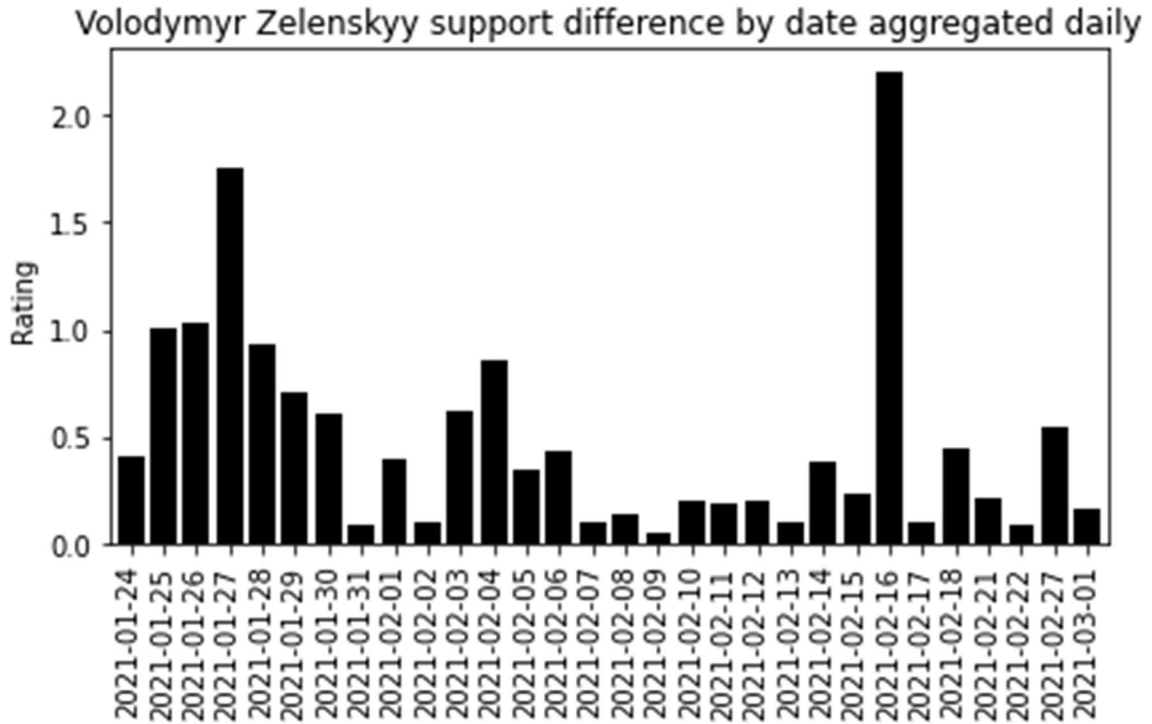


Figure 4: Volodymyr Zelensky rating from 24.01.2021 to 01.03.2021 calculated by the proposed model

When several news items fall upon the same date, it is important to understand how much each of them affected the result. To do this, we calculated the statistics of the frequency of news data. The results in Figure 5 show that the news about the United Arab Emirates was much more resonant than the other two.

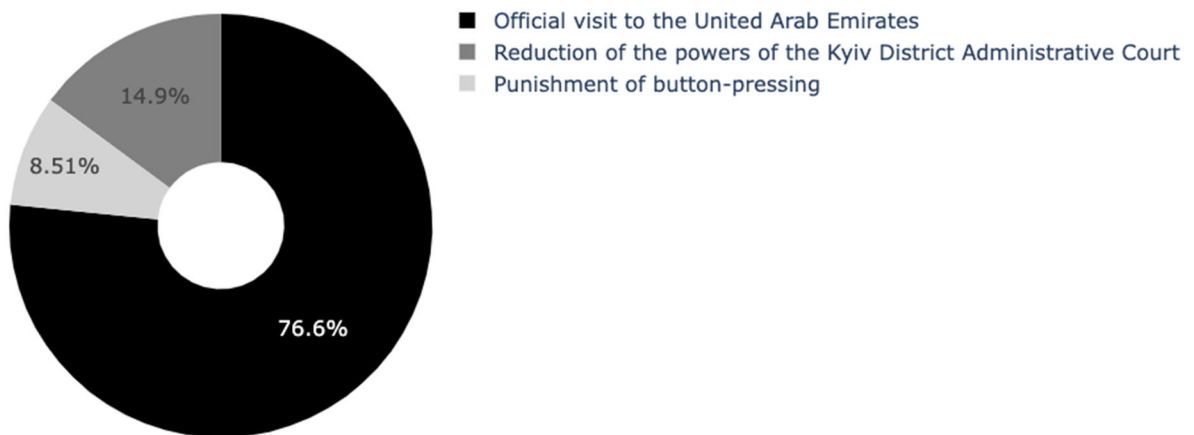


Figure 5: News importance for rating changes on 16.02.2021

Google Trends could be used to verify that the right news is selected. This tool will allow checking which news has been popular in each period and whether it coincides with the results obtained based on data from Twitter. Figures 6-8 show that news detected by an algorithm using Twitter data fell on the same dates as they appeared in google search in Ukraine. Some news could start one or two days earlier than got discussed on Twitter. The administrative court and button-pressing are highly

discussable topics in Ukraine, even if they were not the absolute maximum on the 14th of February they were still on the local maximum.

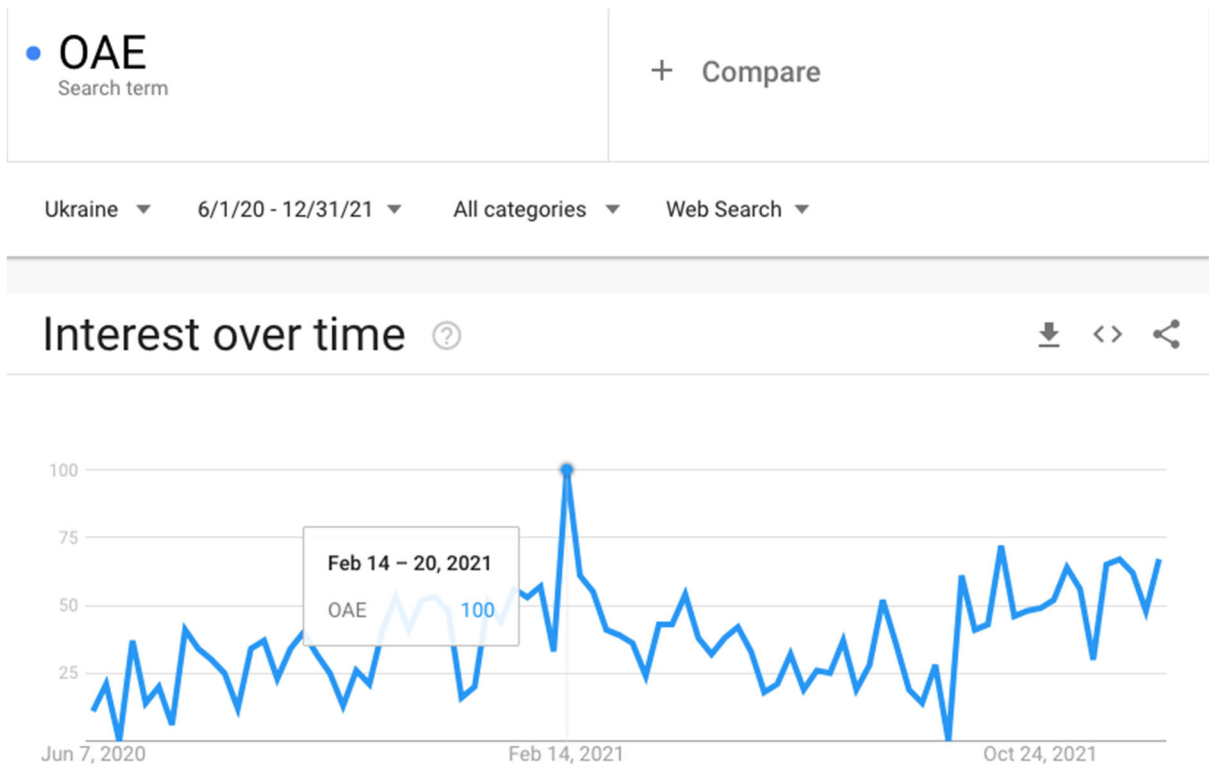


Figure 6: News about the United Arab Emirates

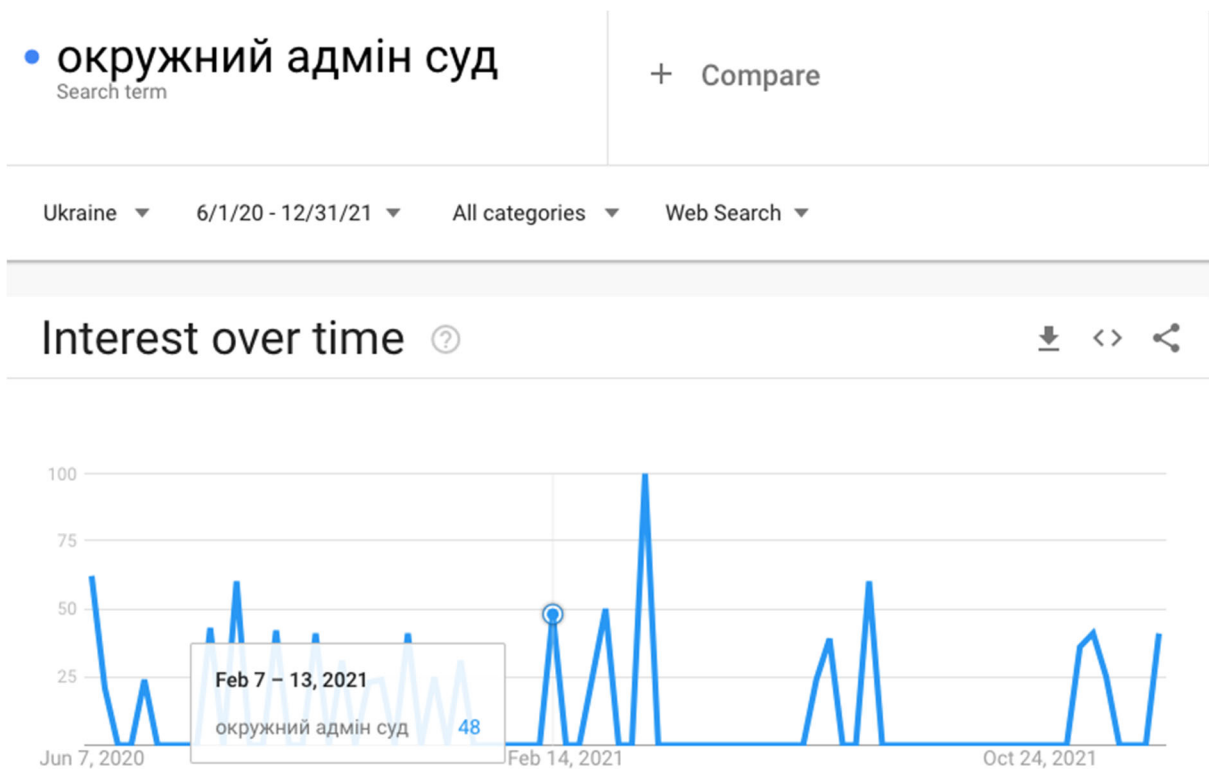


Figure 7: News about Administrative court

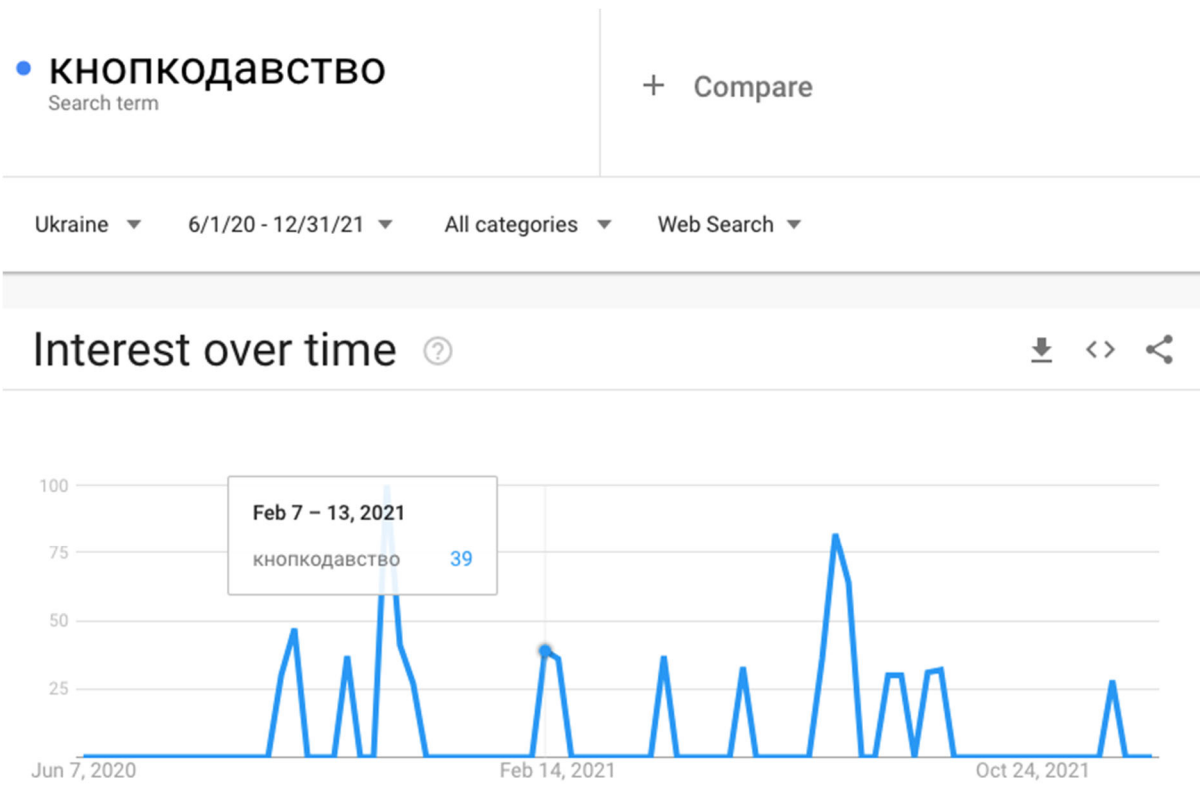


Figure 8: News about button-pressing

5. Conclusion and Future work

The conducted research proposed several algorithms to determine the rating of politicians, detect dates when news affected ratings the most and identify specific news which influenced grows or fall of the rating. Experimental results conducted on Ukrainian President Volodymyr Zelenskyy's page on Twitter show that the proposed approach allows not only to detect of ratings and their changes but detect news that influenced such changes the most.

The result is slightly different from sociological polls. There are several explanations for this:

- Twitter is not very popular in Ukraine;
- Not all segments of the population participating in the elections use this social network. For example, there are very few elderly people;
- Twitter is used by people who do not yet have the right to vote — minors.

In future work proposed model will be evaluated on a calculation of political rating for French elections candidates – Emmanuel Macron and Marine Le Pen. The conducted research confirms that it is possible to identify the events that led to a change in the corresponding political rating. In future work, we plan to develop a system of recommendations for politicians or commercial brands based on the identified key events on how to react to news or informational attacks in order to avoid or decrease the loss of rating.

6. References

- [1] W. Lippmann, Public opinion, Transaction Publishers, New Brunswick, 1998.
- [2] T. Faas, C. Mackenrodt, R. Schmitt-Beck, Polls that mattered: effects of media polls on voters' coalition expectations and party preferences in the 2005 German parliamentary election, International Journal of Public Opinion Research 20 (2008) 299-325. doi:10.1093/ijpor.edn034.

- [3] J.-M. Eberl, C. Plescia, Coalitions in the news: How saliency and tone in news coverage influence voters' preferences and expectations about coalitions, *Electoral Studies* 50 (2018) 30-39. doi:10.1016/j.electstud.2018.07.004.
- [4] C. Plescia, J. Aichholzer, On the nature of voters' coalition preferences, *The Journal of Elections, Public Opinion & Parties* 27 (2017) 254-273. doi:10.1080/17457289.2016.1270286.
- [5] S. Hong, D. Nadler, Which candidates do the public discuss online in an election campaign? The use of social media by 2012 presidential candidates and its impact on candidate salience, *Government Information Quarterly* 29 (2012) 455-461. doi:10.1016/j.giq.2012.06.004.
- [6] G. Enli, Twitter as arena for the authentic outsider: exploring the social media campaigns of Trump and Clinton in the 2016 US presidential election, *Eur. J. Commun.* 32 (2017) 50-61. doi:10.1177/0267323116682802.
- [7] D. Kreiss, S. C. McGregor, Technology firms shape political communication: The work of Microsoft, Facebook, Twitter, and Google with campaigns during the 2016 US presidential cycle, *Political Communication* 35 (2018) 155-177. doi:10.1080/10584609.2017.1364814.
- [8] P. Burnap, R. Gibson, L. Sloan, R. Southern, M. Williams, 140 characters to victory? Using twitter to predict the uk 2015 general election, *Electoral Studies* 41 (2016) 230-233. doi:10.1016/j.electstud.2015.11.017.
- [9] A. Tumasjan, T. O. Sprenger, P. G. Sandner, T. M. Welpe, Election forecasts with Twitter: How 140 characters reflect the political landscape, *Soc. Sci. Comput.* 29 (2011) 402-418. doi.org:10.1177/0894439310386557.
- [10] D. Anuta, J. Churchin, J. Luo, Election bias: Comparing polls and Twitter in the 2016 U.S. election, arXiv:1701.06232 (2017). doi:10.48550/arXiv:1701.06232.
- [11] M. P. Cameron, P. Barrett, B.S., Can social media predict election results? Evidence from New Zealand, *Journal of Political Marketing* 15 (2016) 416-432. doi:10.1080/15377857.2014.959690.
- [12] S. R. Neely, Politically motivated avoidance in social networks: A study of Facebook and the 2020 presidential election, *Social Media + Society* 7 (2021). doi:10.1177/20563051211055438.
- [13] A. Sosnkowski, C. J. Fung, S. Ramkumar, An analysis of Twitter users' long term political view migration using cross-account data mining, *Online Soc. Networks and Media* 26 (2021). doi: 10.1016/J.OSNEM.2021.100177.
- [14] O. Chertov, T. Rudnyk, O. Palchenko: Search of phony accounts on Facebook. Ukrainian case, in *Proceedings of the International Conference on Military Communications and Information Systems, ICMCIS-2018, IEEE, Warsaw, Poland, 2018*. pp. 22-23. doi:10.1109/ICMCIS.2018.8398725.
- [15] O. Chertov, T. Rudnyk, Mathematical model for determining political rating based on social networks, *Communication and Society* (2021). To appear.
- [16] M. S. Alvim, B. Amorim, S. Knight, S. Quintero, F. Valencia, A Multi-agent Model for Polarization Under Confirmation Bias in Social Networks, in: K. Peters, T.A.C. Willemse (Eds.), *Formal Techniques for Distributed Objects, Components, and Systems*, volume 12719 of *Lecture Notes in Computer Science*, vol. Springer, Cham 2021, pp. 22-41. doi:10.1007/978-3-030-78089-0_2.
- [17] P. Concha, L. Pilar, Political influencers. a study of Donald Trump's personal brand on Twitter and its impact on the media and users, *Communication and Society* 32 (2018), 57-75. doi:10.15581/003.32.1.57-75.