Sentiment Analysis of Posts and Comments in the Accounts of Russian Politicians on the Social Network

Kirill Svetlov, Konstantin Platonov
Saint Petersburg State University
Saint Petersburg, Russia
{k.svetlov, k.platonov}@spbu.ru

Abstract—**Russian politicians are increasingly using social networks publishing a lot of texts. One of the important issues in the context of the analysis of political online communication is the choice of negative and positive topics in publications as well as the reaction of the audience. In order to analyze the main patterns of this process we have collected the data from the social network Vkontakte. Our sample covers the period from 1 January, 2017 to 25 April, 2019, in total 46293 posts and 2197063 comments in 23 politician's accounts. To build the classifier we used two text corpora: Rubtsova's corpus and RuSentiment corpus. The algorithm of sentiment analysis was implemented on the basis of bidirectional recurrent neural network. Using Rubtsova's corpus we provided the accuracy of 91% and using RuSentiment we provided the accuracy of 84% (accuracy is calculated as the proportion of correctly identified cases from the test sample). We found that the markup of data significantly differs when different corpora were used. The most adequate results in the analysis of posts and comments, in our opinion, are obtained by using an ensemble of models based on the both corpora. As a result of classification, we identified a number of patterns. Thus, the number of likes and views of posts is higher for the posts classified as positive, and the number of reposts is higher for the posts classified as negative. We also found that the number of comments is higher for the posts with a negative sentiment, and the average sentiment of comments on positive posts is more positive than the average sentiment of comments on negative posts.**

I. Introduction

The Internet has fundamentally transformed the mechanisms for discussing political issues, and democratic participation [1]. Nowadays the key messages are delivered to audiences online, and politicians are discovering new platforms every year. During 2007-2008 Barack Obama's team made a breakthrough using a variety of online technologies for political purposes: the election campaign involved a lot of digital channels: websites, email newsletters, Facebook, Twitter, MySpace and Youtube [2]. In Russia, a kind of an innovator in online presence was Dmitry Medvedev, a President who actively used Twitter and did not refuse the social networks even after the incident, when his account with 2.5 million audience was hacked [3]. Gradually, blog adoption in politics became typical of not only Federal politicians but also regional leaders [4]. The concept of political marketing existed since the 70s [5] and now, in the era of Internet presence, it is very topical. Today, it is the marketing approach that allows consultants and press services to determine the key values of target audiences [6] to form the appropriate image [7] and brand of the leader or party [8]. In recent years, social media marketing (SMM) is increasingly being used in political communication, which includes the promotion, positioning and formation of loyalty to the leader in a social network, develop meaningful strategies for account management. Publishing content in the accounts of well-known politicians remains a part of the communication strategy, and sometimes the entire teams of PR specialists work on it. Accounts in social networks could be considered as media outlets actively used for agenda-building and influence on the audience [9]. One notable challenge for strategic communication in this context is the use of positive and negative content. On the one hand, a positive agenda is demanded by a large part of the audience, some Russian media, for example, Lenta gave users the opportunity to simply turn off the bad news with the toggle switch [10]. In recent years, Russian pro-government politicians have mainly focused on a positive agenda in order to prevent downgrade of ratings of the President [11] and the Deputy Prime Ministers [12]. The promotion of the positive agenda is implemented not only through the media, but also through state corporations [13]. On the other hand, negativity can also be effective, as such content is often used in political communication for demonstrative statements and provocations. But what online content strategy politicians have to choose to be as effective as possible in terms of engagement? In this article, we use data collected form the most popular accounts of Russian politicians on the most visited Russian social network "Vkontakte" [14] and implement machine learning algorithms and sentiment analysis to determine the impact of sentiment on the mechanisms of feedback from the audience. "Vkontakte" is the most popular social network in Russia. It is similar in functionality to Facebook, there are various communities, public communication between users is based on posts and comments, likes and reposts are also used. Since March 2017, "Vkontakte" posts also have views metric.

II. Related work

In the context of political communication, data from social networks can be used in solving the problems of opinion mining as a significant predictor of public opinion, including the growth of social tension [15] and prediction of election results [16], [17]. Sentiment analysis of a policy-related content in social networks is being actively used to monitor trends [18], [19], analysis of the political agenda, and assessment of the level of support for candidates and parties [20]. Wang & Can et

ISSN 2305-7254

al. showed how it is possible to start an online wave of positive or negative reactions through the provocation [21]. Working with Twitter data, Stieglitz & Dang-Xuan showed that emotionally charged publications get higher retweetability [22]. Experiments also show interdependency between intensity of publication activity and various events, such as debates [23]. Sentiment analysis of political online communication is also used to model the emotional background of the discussions1 [24], fact-checking [25], analysis of propaganda trends [26]. Early versions of the sentiment analysis appeared in the late 70's – early 80's [27]. In general, all automatic algorithms can be divided into, firstly, lexical-based, machine learning-based and mixed, and, secondly, ontology based and non-ontology based [28]. There are many sentiment classification algorithms, however, even advanced algorithms tend to reach accuracy about 70% on social media data [29], [30], [31]. Researchers use numerous methods to conduct sentiment analysis, for example, various modifications of the bag-of-words [32], [33], neural networks [34], machine learning [35], conditional random fields and support vector machines [36], deep learning [37]. The automatic sentiment classification for the Russian language is still a challenging task because of low number of open source and ready solutions, as well as the small community and low number and quality of thematic text corpora.

We submit the following research questions:

1) How an accuracy of sentiment classification can be increased?
2) Are there any significant differences in feedback levels depending on the sentiment of the content? (Which posts get more likes, comments, reposts?)
3) How is the positive/negative sentiment in comments related to the positive/negative sentiment in posts?

III. Data and methodology

A. Text processing

For sentiment classification we used the Russian language corpus of short texts RuTweetCorp (Rubtsova's corpus) [41] and the corpus of the posts RuSentiment [29]. RuTweetCorp corpus contains texts posted by Twitter users during the period from late November 2013 to late February 2014. There were 226834 records in total, 114911 annotated as positive and 111923 annotated as negative. RuSentiment corpus contains 30521 annotated posts on social network Vkontakte, divided into 5 classes: 6646 posts marked as positive (code "positive"), 3912 marked as negative (code "negative"), 3467 marked as direct speech (code "speech"), 4440 unrecognized (code "skip"), 12720 neutral (code "neutral"). We used only the first two classes of posts from this case in order to match the classes in Rubtsova's corpus. The first step in model building was to develop procedure for text preprocessing. We used standardized procedure [43] that covered all the text data in this study. This consist of the following steps.

1) We replaced the Russian letter "ё" with "е";
2) The particles "not" and "neither" were converted to the prefix "NOT" to the next word;
3) Links to web resources were excluded from the text;
4) User's mentions were excluded from the text;
5) Processing of emoticons and emojis was performed. The most popular of them were replaced with the tokens "POSITIVESMILE" or "NEGATIVESMILE";
6) We deleted all non-letter characters, including punctuation and numbers;
7) Normalization of words was performed with MyStem [44];
8) Any number of consecutive spaces was replaced with a single space;
9) Repeated consecutive words were replaced with a single words;

Taking emoticons in account allows to increase the accuracy of text classification [39]. We conducted a special procedure for processing emoticons and emojis. First of all, emoticons composed of typographic characters were replaced by these tokens. As positive emoticons we considered the following {:) , ;) , =) , :−) , ;−) ,)) ,))) , ...}, we considered the following negative {:(, ;(, =(, :−(, ;−(, ((, (((, ...}. Further, a similar procedure was applied to the following manually marked emojis, which are used in VKontakte:

- **Positive emojis**: thumbs_up, red_heart, folded_hands, smiling_face_with_smiling_eyes, grinning_face_with_big_eyes, clapping_hands, beaming_face_with_smiling_eyes, flexed_biceps, grinning_face_with_smiling_eyes, winking_face, winking_face_with_tongue, rose, smiling_face_with_heart-eyes, grinning_face, kiss_mark, tulip, oncoming_fist, hugging_face, OK_hand, victory_hand, face_blowing_a_kiss, hibiscus, fire, cherry_blossom, sparkles, handshake, party_popper, two_hearts, raised_fist, sparkling_heart, raised_hand, bouquet, smiling_face, slightly_smiling_face, blue_heart
- **Negative emojis**: rolling_on_the_floor_laughing, grinning_squinting_face, pouting_face, SOS_button, smirking_face, thumbs_down, fearful_face, thinking_face, pile_of_poo, middle_finger, crying_face, loudly_crying_face, person_facepalming, nauseated_face, pensive_face, grinning_face_with_sweat, unamused_face, see-no-evil_monkey, face_screaming_in_fear

It should be mentioned that in the process of checking the effectiveness of various methods of text preprocessing, we also considered the step of filtering the texts with Russian-language stop-words list included in the "nltk.corpus" package, however, their exclusion reduced the accuracy of the models on the considered corpora. So, this procedure was removed from the list of steps for text preprocessing.

B. Sentiment analysis

Traditional word-based sentiment analysis techniques are often not suitable for classifying policy-related content [38]. In order to construct the sentiment classifier based on the Rubtsova's corpus, we adopted an approach using deep learning models such as convolutional and recurrent neural networks.

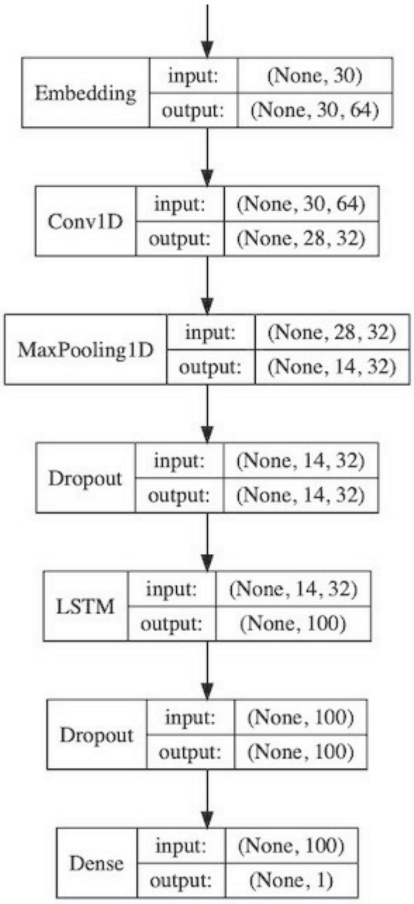

Fig. 1. CNN + LSTM neural network architecture

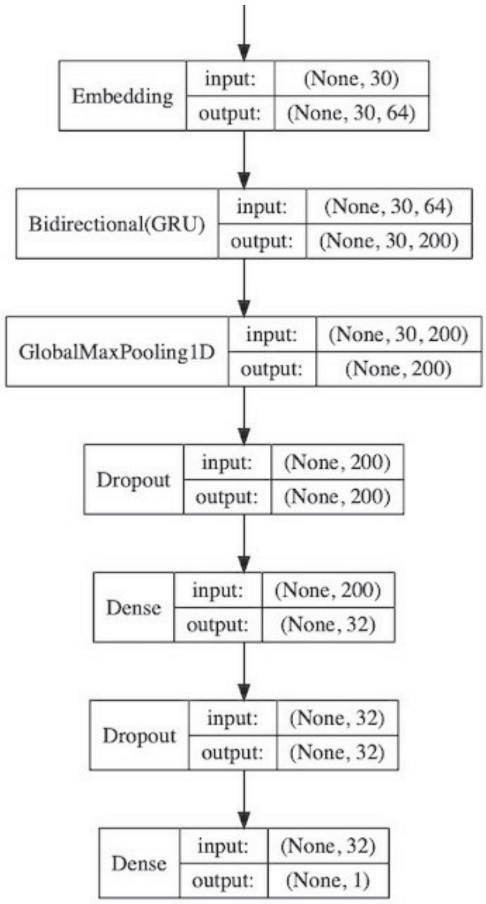

Fig. 2. BiGRU neural network architecture

As the first alternative, we used the neural network [40], in which the LSTM (Long Short-Term Memory) layer follows the one-dimensional convolution layer (Fig. (1)). As a second alternative, we used a neural network with a bidirectional GRU (Bidirectional Gated Recurrent Unit, BiGRU) layer (Fig. (2)) followed by dense layers. This architecture, in particular, was implemented in the library DeepPavlov [42]. In both neural networks, the first layer is the word embeddings vector representation layer which maps the first 30 words from the preprocessed text to a real-valued vectors from R^{64}. The size of the dictionary, which includes the most common words from the Rubtsova's corpus, was 30 thousand words. In order to avoid overtfitting, we use dropout layers with parameter 0.2 and L2-regularization with parameter 10^{-5} in dense layers. We used the following data separation for the neural networks training : 70% of the data used in the training sample, 10% of the data included in the validation sample and 20% – in the test sample. The selection of the number of training epochs was based on accuracy and losses metrics on the validation set. Fig.(3) and Fig. (4): to achieve optimal scores (both loss and accuracy) and to avoid the effect of overfitting CNN+LSTM network had completed 2 epochs of training, and BiGRU network had completed 3 epochs. The performance results of these models on the test set are represented at tables (I) and (II). Since the value of the F_1 metric for the BiGRU model was higher, then we decided to use it instead of CNN+LSTM.

We also used BiGRU architecture of a neural network with the RuSentiment corpus. Since this corpus is smaller, compared with the Rubtsova's one, we used a dictionary containing 15 thousand words, which are the most frequent in the corpus. The table (III) shows that this model is able to recognize only 54.1% of negative replies. Also, because the classes it was trained on were not balanced, model's predictions should be expected to be biased in the direction of positive sentiment.

TABLE I. RESULTS FOR CNN + LSTM MODEL ON THE TEST SET FROM RUBTSOVA'S CORPUS

Class	precision	recall	f1-score	support
Negative	0.90238	0.91206	0.90719	22196
Positive	0.91405	0.90458	0.90929	22950
macro avg	0.90822	0.90832	0.90824	45146
weighted avg	0.90831	0.90825	0.90826	45146

TABLE II. RESULTS FOR BIGRU MODEL ON THE TEST SET FROM RUBTSOVA'S CORPUS

Class	precision	recall	f1-score	support
Negative	0.92094	0.89223	0.90636	22196
Positive	0.89882	0.92593	0.91217	22950
macro avg	0.90988	0.90908	0.90927	45146
weighted avg	0.90970	0.90936	0.90932	45146

IV. RESULTS

We investigate the characteristics of the posts published by Russian politicians on the social network VKontakte for

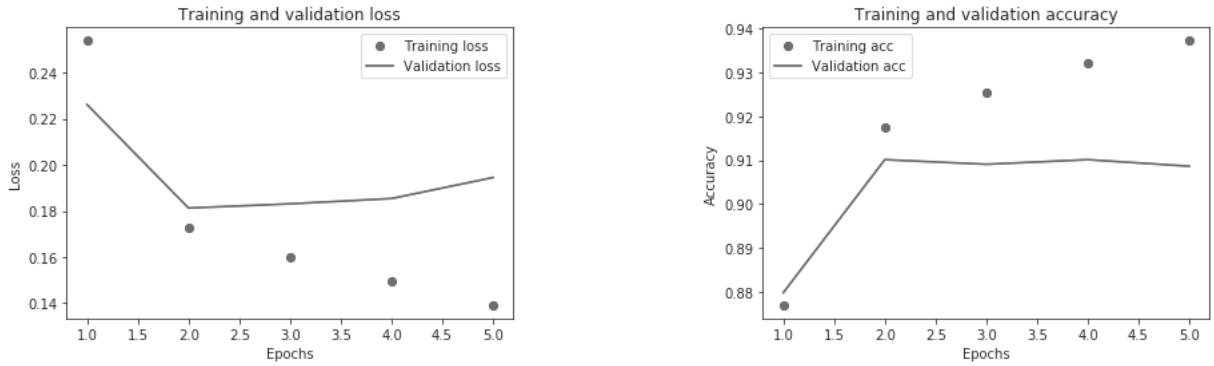

Fig. 3. Training of CNN + LSTM model. Comparing results on training and validation sets from Rubtsova's corpus

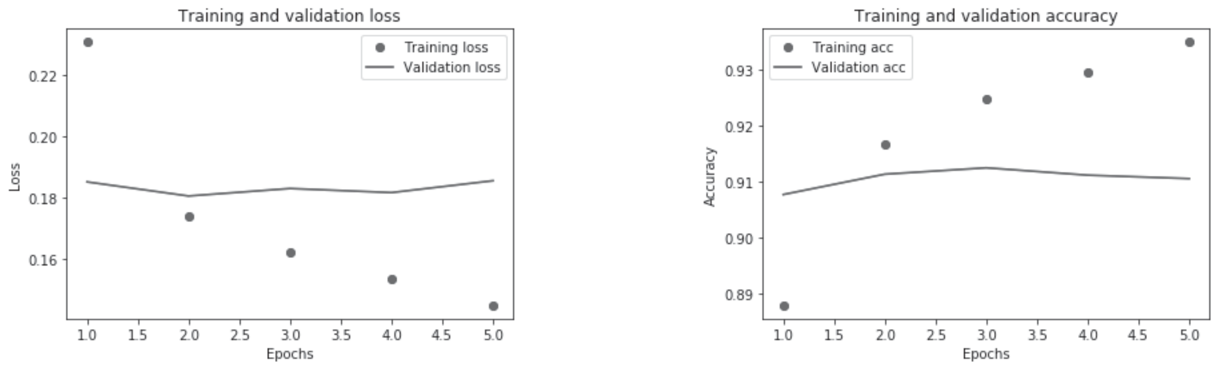

Fig. 4. Training of BiGRU model. Comparing results on training and validation sets from Rubtsova's corpus

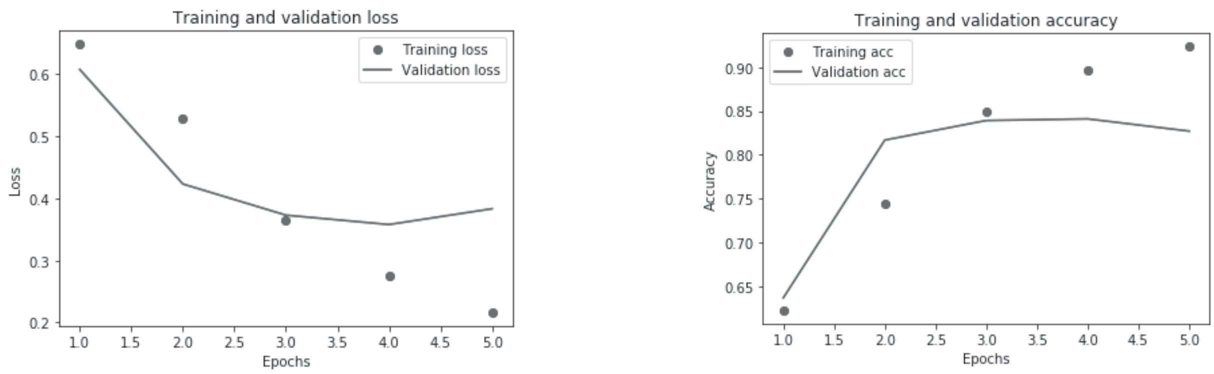

Fig. 5. Training of BiGRU model. Comparing results for training and validation sets from RuSentiment corpus

TABLE III. RESULTS FOR BIGRU MODEL FOR THE TEST SET FROM RUSENTIMENT CORPUS

Class	precision	recall	f1-score	support
Negative	0.89474	0.54140	0.67460	785
Positive	0.77968	0.96224	0.86139	1324
macro avg	0.83721	0.75182	0.76800	2109
weighted avg	0.82251	0.80560	0.79187	2109

the period from January 1, 2017 to April 25, 2019. There are 46293 posts of 23 politicians and 2197063 of users' replies to them. We decided to conduct sentiment analysis for each of the models separately (Rubtsova's corpus and RuSentiment corpus), as well as using both models. By using both models, we mean using the ensemble technique, in which the probabilities of assigning text to a particular class are determined in accordance with the probabilities for each individual model taken with given weights (in our case $(0.5, 0.5)$). In order to calculate the average sentiment of the politicians posts, we use the following technique: we assign a number +1 to each positive post, -1 to each negative post and 0 to unrecognised (skipped) posts. Unrecognised texts are those consisting of tokens that are not contained in the dictionary of the most frequent words in the corpus. Then we calculate the averages for all posts and comments for the given politicians. The engagement rate is calculated for each post as the sum of the number of likes and comments per thousand views of this post:

$$Eng.Rate = \frac{Likes + Commetns}{Views} \cdot 1000 \ .$$

Statistics on the average sentiment as well as on the number of views and the engagement rate are presented in the table (IV). The average sentiment of posts takes values from the range $[-1, 1]$. If among the texts there is an identical number of positive and negative posts, then the average sentiment will be 0. If most of the posts are positive, then this indicator will be greater than 0. And on the contrary, in a situation when majority of the posts are negative, the average sentiment will be less than 0. As can be seen from the table (IV), most politicians have a positive average sentiment of posts. At the same time, for most of the politicians this indicator calculated on the basis of Rubtsova's corpus is much higher than that obtained on the RuSentiment corpus. The use of an ensemble of models leads to mid-range results.

The table (V, VI, VII) summarizes the data on classification of the politicians' posts in accordance with the model built on Rubtsova's corpus, the model built on the RuSentiment corpus and the model which represents the ensemble of the two mentioned above.

V. Conclusion

Despite the difference in overall assessments of positive, negative and unrecognized texts obtained as a result of the application of various models, the following patterns were identified:

- the number of likes from users is higher for posts classified as positive
- the number of comments is higher for posts with a negative sentiment, and this difference is mainly provided by comments that also have a negative sentiment. The average tone of comments on positive posts is more positive than the average tone of comments on negative posts.
- the number of reposts is higher for posts classified as negative.
- the number of views is higher for positive posts.
- according to the sentiment classifier built on the Rubtsova's corpus, the level of engagement rate is higher for negative posts. Inversely, using the classifier built on the RuSentiment corpus, we found the opposite relation. In general, we can say that there is no clear relationship between the tone of posts and engagement rate.

We found that the most adequate results of sentiment classification can be achieved by using an ensemble of the models. Also we showed that a special emoji encoding scheme, taking into account the frequent use of sarcasm, also increased the quality of classification. Another procedure towards improving the quality was inclusion of stop-words. Significant differences in the results obtained from different corpora suggest that more specific corpora are needed for a more reliable classification, high precision and recall. A balance of classes, the specificity of the content are main requirements for such corpora. In terms of strategy, we showed that in general it is possible to influence on the magnitude and sentiment of the feedback from the audience on social media using negative or positive content.

Further, we plan to analyze the dynamic patterns associated with publication activity and sentiment to determine the general trends and identify the characteristics of particular authors. We also plan to describe in detail the topics represented by positive and negative sentiment in posts and comments. Another issue is to understand why there were such significant variations between the results of the models trained on different corpora, and which texts were assigned to opposite classes by the different models.

Acknowledgment

This work was funded by RFBR and EISR, project number 19-011-31651. We would like to thank our colleagues from the Center for Sociological and Internet Research of Saint Petersburg State University for guidance and support.

References

[1] J.C. Tedesco, "Changing the channel: Use of the Internet for communicating about politics" *Handbook of political communication research*, Routledge, 2004, pp. 525-550.

[2] J. Katz, M. Barris, A. Jain, *The social media president: Barack Obama and the politics of digital engagement*, Springer, 2013.

[3] BBC News, Russian PM's Twitter hacked, posting 'I resign'. Web: https://www.bbc.com/news/technology-28786683 .

[4] F. Toepfl, "Blogging for the sake of the president: The online diaries of Russian governors", *Europe-Asia Studies*, vol. 64.8, 2012, pp. 1435-1459.

[5] A. Shama, "The marketing of political candidates", *Journal of the academy of marketing science*, vol. 4.4, 1976, pp. 764-777.

[6] B. McNair, *An introduction to political communication*, Routledge, 2017.

[7] M. Lalancette, R. Vincent, "The power of political image: Justin Trudeau, Instagram, and celebrity politics", *American Behavioral Scientist*, vol. 63.7, 2019, pp. 888-924.

[8] Rutter, R. N., Hanretty, C., & Lettice, F. (2018). Political brands: can parties be distinguished by their online brand personality. Journal of Political Marketing, 17(3), 193-212.

[9] J.T. Feezell, " Agenda setting through social media: The importance of incidental news exposure and social filtering in the digital era", *Political Research Quarterly*, vol. 71.2, 2018, pp. 482-494.

[10] Lenta, Good news feed. Web: https://lenta.ru/news/2018/10/15/dobro/.

[11] TVRain, Kremlin in search of positive agenda for Presidential Address. (In Russian) Web: https://tvrain.ru/news/v_kremle_zahoteli_najti_pozitivnuju_povestku_dlja_poslanija_putina-480117/ .

[12] Foundation "Politics in Saint Petersburg". Web: https://fpp.spb.ru/fpp-review-viceagenda-2019-01-04 .

[13] Reuters, Exclusive: Kremlin tells companies to deliver good news. Web: https://www.reuters.com/article/us-russia-kremlin-pr/ . exclusive-kremlin-tells-companies-to-deliver-good-news-idUSKBN1DE0MP

[14] Vkontakte social network. Web: https://vk.com .

[15] P. Burnap, "Detecting tension in online communities with computational Twitter analysis", *Technological Forecasting and Social Change*, vol. 95, 2015, pp. 96-108.

[16] A. Tumasjan, T. O. Sprenger, P. G Sandner, I. M Welpe, "Predicting elections with twitter: What 140 characters reveal about political sentiment", *Fourth international AAAI conference on weblogs and social media*, 2010.

[17] A. Bermingham, A.Smeaton, "On using Twitter to monitor political sentiment and predict election results", *Proceedings of the Workshop on Sentiment Analysis where AI meets Psychology*, SAAIP 2011, 2011, pp. 2-10.

[18] H. Wang, D. Can, A. Kazemzadeh, F. Bar, S. Narayanan, " A system for real-time twitter sentiment analysis of 2012 us presidential election cycle", *In Proceedings of the ACL 2012 system demonstrations*, 2012, July, pp. 115-120.

TABLE IV. STATISTICS ON POLITICIANS'S POSTS

	Name	Post count	Avg. sent. Rubtsova's	Avg. sent. RuSentiment	Avg. sent. ensemble	Avg. thousand views	Avg. eng. rate
1	Dmitry Medvedev	798	0.456140	0.155388	0.393484	181.622846	11.662548
2	Alexander Beglov	28	0.642857	0.357143	0.714286	107.953321	29.401820
3	Alexei Navalny	1811	0.455549	0.015461	0.264495	102.382516	32.315528
4	Ramzan Kadyrov	2247	0.671117	-0.038273	0.409435	58.082514	25.994744
5	Rustam Minnikhanov	2308	0.418977	0.639948	0.649913	33.551016	27.844797
6	Sergei Sobyanin	2179	0.643414	0.383662	0.615879	28.000022	21.625136
7	Andrei Vorobyov	1132	0.556537	0.522085	0.658127	19.484549	25.495139
8	Vladimir Zhirinovsky	2328	0.272337	-0.198024	0.048110	19.450439	18.963259
9	Nail Magdeev	37	0.756757	0.540541	0.756757	15.787838	16.642964
10	Gennady Zyuganov	433	0.341801	-0.055427	0.124711	13.587981	40.389430
11	Igor Orlov	104	0.778846	0.432692	0.759615	11.654760	11.721684
12	Oleg Kuvshinnikov	474	0.814346	0.691983	0.873418	9.114207	20.526467
13	Veniamin Kondratiev	631	0.559429	0.175911	0.454834	8.100658	22.595314
14	Ilya Yashin	670	0.383582	-0.085075	0.162687	7.671209	45.627586
15	Vitaly Milonov	796	0.163317	0.243719	0.233668	7.548525	19.542449
16	Grigory Yavlinsky	329	0.373860	-0.294833	-0.057751	6.515948	43.406006
17	Anatoly Lokot'	539	0.541744	0.307978	0.564007	6.113163	22.806604
18	Sergey Mironov	1358	0.509573	0.024300	0.266568	5.965062	42.358556
19	Igor Koshin	508	0.690945	0.438976	0.651575	3.970091	26.252655
20	Vladimir Medinsky	566	0.733216	0.602473	0.775618	3.402200	27.821765
21	Dmitry Berdnikov	608	0.736842	0.578947	0.779605	1.998965	23.828365
22	Mikhail Delyagin	23984	0.369830	-0.484031	-0.193170	1.298584	26.397373
23	Yunus-Bek Yevkurov	2425	0.735670	0.466804	0.689485	0.791460	35.689487

TABLE V. PARAMETERS CALCULATED USING THE MODEL BASED ON RUBTSOVA'S CORPUS

	Negative	Positive	Skip
Count	11097	31803	3393
Avg. likes per thousand views	23.3007	23.6702	22.5529
Avg. comments per thousand views	4.82512	4.18799	5.34748
Avg. reposts per thousand views	3.35964	2.72232	2.29247
Avg. thousand views	13.0054	16.3057	25.6833
Avg. Engagement Rate	27.2452	26.9801	25.8803
Comments sentiment	0.134148	0.210379	0.257837
Avg. positive comments per thousand views	2.81579	2.49562	3.0288
Avg. negative comments per thousand views	2.20619	1.86972	2.19351
Avg. skipped comments per thousand views	0.833537	0.698572	0.921454

TABLE VI. PARAMETERS CALCULATED USING THE MODEL BASED ON RUSENTIMENT CORPUS

	Negative	Positive	Skip
Count	24564	18241	3488
Avg. likes per thousand views	22.8638	24.5242	22.6266
Avg. comments per thousand views	4.8092	3.69643	5.19736
Avg. reposts per thousand views	3.57032	1.97699	2.25258
Avg. thousand views	11.3292	20.9597	25.6677
Avg. Engagement Rate	26.8703	27.2913	25.9022
Comments sentiment	0.065667	0.258147	0.287601
Avg. positive comments per thousand views	2.4118	2.05598	2.71743
Avg. negative comments per thousand views	2.70584	1.73681	2.51048
Avg. skipped comments per thousand views	0.888601	0.612741	0.971881

TABLE VII. PARAMETERS CALCULATED USING THE ESSEMBLE OF MODELS

	Negative	Positive	Skip
Count	18694	24213	3386
Avg. likes per thousand views	22.9861	24.0298	22.5478
Avg. comments per thousand views	4.81585	3.96692	5.3588
Avg. reposts per thousand views	3.72727	2.23764	2.29472
Avg. thousand views	10.0768	19.6081	25.6831
Avg. Engagement Rate	27.0285	27.0638	25.8812
Comments sentiment	0.0818128	0.259005	0.305826
Avg. positive comments per thousand views	2.58167	2.25659	2.99126
Avg. negative comments per thousand views	2.56862	1.85776	2.37994
Avg. skipped comments per thousand views	0.89374	0.622672	0.90714

[19] A. Ceron, C. Luigi, M. I. Stefano, "Using sentiment analysis to monitor electoral campaigns: Method matters - evidence from the United States and Italy", *Social Science Computer Review*, vol. 33.1, 2015, pp. 3-20.

[20] E. Martínez-Cámara, M. T. Martín-Valdivia , L. A. Urena-López, A. R. Montejo-Ráez, "Sentiment analysis in Twitter", *Natural Language Engineering*, vol. 20.1, 2014, pp. 1-28.

[21] H. Wang, D. Can , A. Kazemzadeh, F. Bar, S. Narayanan, "A system for real-time twitter sentiment analysis of 2012 us presidential election cycle", *Proceedings of the ACL 2012 system demonstrations*, 2012, July, pp. 115-120.

[22] S. Stieglitz, L. Dang-Xuan, "Political communication and influence through microblogging–An empirical analysis of sentiment in Twitter messages and retweet behavior", *Hawaii International Conference on System Science*, IEEE, 2012, pp. 3500-3509.

[23] Y. Mejova , P. Srinivasan , B. Boynton, "Gop primary season on twitter: popular political sentiment in social media", *Proceedings of the sixth ACM international conference on Web search and data mining*, ACM, 2013, pp. 517-526.

[24] J. Bollen, H. Mao, A. Pepe, "Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena", *Fifth International AAAI Conference on Weblogs and Social Media*, 2011.

[25] D. Gayo-Avello et al., "Predicting information credibility in time-sensitive social media",*Internet Research*, 2013.

[26] K. Watanabe, "Conspiracist propaganda: How Russia promotes anti-establishment sentiment online", *ECPR General Conference*, Hamburg, 2018.

[27] B. Pang, L. Lee, "Opinion mining and sentiment analysis", *Foundations and Trends® in Information Retrieval*, vol. 2.1–2, 2008, pp. 1-135.

[28] K. Ravi, V. Ravi, "A survey on opinion mining and sentiment analysis: tasks, approaches and applications", *Knowledge-Based Systems*, vol. 89, 2015, pp. 14-46.

[29] A. Rogers, A. Romanov, A. Rumshisky, S. Volkova, M. Gronas, A. Gribov, "RuSentiment: An Enriched Sentiment Analysis Dataset for Social Media in Russian", *Proceedings of the 27th International Conference on Computational Linguistics*, 2018, pp. 755-763.

[30] A. Bakliwal, J. Foster, J. van der Puil, R. O'Brien , L. Tounsi, M. Hughes, "Sentiment analysis of political tweets: Towards an accurate classifier", *Association for Computational Linguistics*, 2013.

[31] E. Martínez-Cámara, M. T. Martín-Valdivia, L. A. Urena-López, A. R. Montejo-Ráez, "Sentiment analysis in Twitter", *Natural Language Engineering*, vol. 20(1), 2014, pp.1-28.

[32] M. Haselmayer, M. Jenny, "Sentiment analysis of political communication: combining a dictionary approach with crowdcoding", *Quality & quantity*, vol. 51.6, 2017, pp. 2623-2646.

[33] E. Rudkowsky, "More than bags of words: Sentiment analysis with

word embeddings", *Communication Methods and Measures*, vol 12.2-3, 2018, pp. 140-157.

[34] S. V. Georgakopoulos, S. K. Tasoulis, A. G. Vrahatis, V. P. Plagianakos, "Convolutional Neural Networks for Twitter Text Toxicity Analysis", *In INNS Big Data and Deep Learning conference*, 2019, , pp. 370-379.

[35] V. S. Shirsat, R. S. Jagdale, S. N. Deshmukh, "Sentence Level Sentiment Identification and Calculation from News Articles Using Machine Learning Techniques", *Computing, Communication and Signal Processing*, 2019, pp. 371-376.

[36] H. Xia, Y. Yang, X. Pan, Z. Zhang, W. An, "Sentiment analysis for online reviews using conditional random fields and support vector machines", *Electronic Commerce Research*, 2019, pp. 1-18.

[37] R. Bose, R.K. Dey, S. Roy, D. Sarddar, "Analyzing Political Sentiment Using Twitter Data", *Information and Communication Technology for Intelligent Systems. Smart Innovation, Systems and Technologies*, vol, 107, 2019.

[38] T. Mullen, R. Malouf, "A Preliminary Investigation into Sentiment Analysis of Informal Political Discourse", *AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs*, 2006, pp. 159-162.

[39] A. Hogenboom, D. Bal, F. Frasincar, M. Bal, F. de Jong, U. Kaymak, "Exploiting emoticons in sentiment analysis", *In Proceedings of the 28th annual ACM symposium on applied computing*, 2013, pp. 703-710.

[40] S. Nikolenko , A. Kadurin, E. Arkhangelsky, *Deep learning*, Publishing house, Piter, 2017.

[41] Y.V. Rubtsova, "Constructing a corpus for sentiment classification training", *Software products and systems*, vol. 1(109), 2015, pp. 72-78.

[42] M. Burtsev et al. "DeepPavlov: open-source library for dialogue systems", *Proceedings of ACL 2018, System Demonstrations*, 2018, pp. 122-127.

[43] S. Smetanin, Sentiment Analysis of Tweets in Russian using Convolutional Neural Network with Word2Vec Word Embeddings. Web: https://github.com/sismetanin/sentiment-analysis-of-tweets-in-russian .

[44] I. Segalovich, "A fast morphological algorithm with unknown word guessing induced by a dictionary for a web search engine", *MLMTA*, 2003, pp. 273-280.