

Political Advertising Dataset: the use case of the Polish 2020 Presidential Elections

Łukasz Augustyniak and Krzysztof Rajda and Tomasz Kajdanowicz

Wrocław University of Science and Technology

Department of Computational Intelligence, Wrocław, Poland

lukasz.augustyniak@pwr.edu.pl krzysztof.rajda@pwr.edu.pl

tomasz.kajdanowicz@pwr.edu.pl

Michał Bernaczyk

Wrocław University

Faculty of Law, Administration and Economics, Wrocław, Poland

michal.bernaczyk@uwr.edu.pl

Abstract

Political campaigns are full of political ads posted by candidates on social media. Political advertisements constitute a basic form of campaigning, subjected to various social requirements. We present the first publicly open dataset for detecting specific text chunks and categories of political advertising in the Polish language. It contains 1,705 human-annotated tweets tagged with nine categories, which constitute campaigning under Polish electoral law. We achieved a 0.65 inter-annotator agreement (Cohen’s kappa score). An additional annotator resolved the mismatches between the first two annotators improving the consistency and complexity of the annotation process. We used the newly created dataset to train a well established neural tagger (achieving a 70% percent points F1 score). We also present a possible direction of use cases for such datasets and models with an initial analysis of the Polish 2020 Presidential Elections on Twitter.

1 Introduction

The emergence of social media has changed how political campaigns take place around the world (Kearney, 2013). Political actors (parties, action committees, candidates) utilize social media platforms, Twitter, Facebook, or Instagram, to communicate with and engage voters (Skogerbø and Krumsvik, 2015). Hence, researchers must analyze these campaigns for several reasons, including enforcement of the laws on campaign contribution limits, implementation of freedom and fairness in electoral campaigns or protection against slander, hate-speech, or foreign interference. Unlike U.S. federal or state law, European jurisdictions present a rather lukewarm attitude to unrestrained online campaigning. Freedom of expression in Europe has more limits (Rosenfeld, 2003), and that can be seen in various electoral codes in Europe (e.g. in France, Poland) or statutes imposing mandatory systems of notice-and-takedown (e.g. the German Network Enforcement Act, alias the Facebook Act). In Poland, *agitation* (an act of campaigning) may commonly be designated ‘political advertisement’, corporate jargon originating from such as Twitter’s or Facebook’s terms of service. Primarily, however, it has a normative definition in article 105 of the Electoral Code. It covers any committees’ or voters’ public acts of inducement or encouragement to vote for a candidate or in a certain way, regardless of form. Election promises may appear in such activities, but do not constitute a necessary component. A verbal expression on Twitter falls into this category. There exist some natural language resources for the analysis of political content in social media. These include collections related to elections in countries such as Spain (Taulé et al., 2018), France (Lai, 2019), and Italy (Lai et al., 2018). Vamvas and Sennrich (2020) created an X-stance dataset that consists of German, French, and Italian text, allowing for a cross-lingual evaluation of stance detection. While the datasets on political campaigning are fundamental for studies on social media manipulation (Aral and Eckles, 2019), there are no Polish language datasets related to either political advertising or stance detection problems. We want to fill this gap and expand natural language resources for the analysis of political content in the Polish language. Our contributions are as follows: (1) a novel, publicly open dataset for detecting specific text chunks and categories of political advertising in the Polish language,

(2) a publicly available neural-based model for tagging social media content with political advertising that achieves a 70% F1 score, and (3) an initial analysis of the political advertising during the Polish 2020 Presidential Election campaign.

2 Political Advertising Dataset Annotation

We created nine categories of political advertising (incl. election promises) based on manually extracted claims of candidates (see sample of examples in Table 1) and a taxonomy proposed by Vamvas and Sennrich (2020). We gathered political advertising topics (incl. election promises) provided by presidential candidates and committee websites, Facebook fanpages, and other websites (such as news agencies). We tried to fit them into the Vamvas and Sennrich (2020) taxonomy. However, we spotted that some categories should be corrected or omitted, such as the economy. It was particularly hard to divide whether political advertising should fall into the category of welfare or economy. The final version of categories was created after a couple of iterations with our annotator team, and we chose categories for which we got a consensus and repeatable annotations. By repeatable, we understand that the same or another annotator will consistently choose the same categories for repeated examples or very similar examples. To the best of our knowledge, there do not exist any commonly used political advertising categories. Moreover, we are aware that they can evolve in future election campaigns. We shall be ready to update the dataset according to political advertising types, categories, and any concept drift in the data.

We extracted all Polish tweets related to the election based on appearances in the tweet of (1) specific hashtags such as *wybory* (*elections*), *wybory2020* (*elections2020*), *wyboryprezydenckie* (*presidential-elections*), *wyboryprezydenckie2020* (*presidentialelections2020*), and (2) unigram and bigram collocations¹ generated using examples of political advertising from all categories (as in Table 1). We name these sets of hashtags and collocations together as *search keywords or search terms*. We gathered almost 220,000 tweets covering approximately ten weeks, between February 5, 2020 and April 11, 2020 using the search terms mentioned. We assigned sentiment orientation for each tweet using an external Brand24 API². Then we sampled tweets using a two stage procedure: (1) we divided all tweets into three sets based on sentiment orientation (positive, neutral, negative) to have representations of various attitudes, and (2) for each sentiment category we randomly selected $\log_2(|E_k|)$ examples, where E is a set of all tweets for a particular *search keyword* k . We used almost 200 different *search keywords* and finally reached 1,705 tweets for the annotation process. The dataset was annotated by two expert native speaker annotators (linguists by training). The annotation procedure was similar to Named Entity Tagging or Part-of-Speech Tagging, so the annotators marked each chunk of non-overlapping text that represents a particular category. They could annotate a single word or multi-word text spans. Table 1 presents an example of political advertising with corresponding categories. These kinds of examples (they could be named as seed examples) were presented to the annotators as a starting point for annotating tweets. However, the annotators could also mark other chunks of text related to political advertising when the chunk is semantically similar to examples or it clearly contains examples of a political category but not present in the seed set. We achieved a 0.48 Cohen’s kappa score for exact matches of annotations (even a one-word mismatch was treated as an error) and a 0.65 kappa coefficient counting partial matches such as *reduce coil usage* and *reduce coil* as correct agreement. We disambiguated and improved the annotation via an additional pass by the third annotator. He resolved the mismatches between the first two annotators and made the dataset more consistent and comprehensive. According to McHugh (2012), the 0.65 kappa coefficient lies between a moderate and strong level of agreement. We must remember that we are annotating social media content. Hence there will exist a lot of ambiguity, slang language, and very short or chunked sentences without context that could influence the outcome. Then, we trained a Convolutional Neural Network model using a spaCy Named Entity classifier (Honnibal, 2018), achieving a 70% F1 score for 5-fold cross-validation. We used fastText vectors for Polish (Grave et al., 2019) and default spaCy model hyperparameters. The dataset and trained model are publicly available in the GitHub repository³. Table 3 presents per category precision, recall, F1 score, as well as the number of

¹Collocation extracted using NLTK library (<https://www.nltk.org/>).

²<https://brand24.com/>

³<https://github.com/laugustyniak/misinformation>

Healthcare	Welfare	Education	Immigration	Infrastructure and Environment	Defense and Security	Foreign Policy	Society	Political and Legal System
drug reform, national oncology strategy	tax-free allowance, 500+ programme, minimum wage	e-learning, remote learning, dentist in every school	immigrant, refugees	climate change, reduce coal usage, water crisis	national security, alliance with the USA, NATO alliance	lack of foreign policy, EU collaboration, Weimar Triangle	death penalty, separation of church and state, LGBT	presidential veto, independence of the courts

Table 1: Political advertising – categories with examples.

examples for each category in the whole dataset. As we can see, there are 2507 spans annotated. Interestingly, 631 tweets have been annotated with two or more spans. Finally, 235 tweets do not contain any annotation span, and they represent 13.8% of the whole dataset.

3 Polish 2020 Presidential Election - Use Case

We present the analysis of 250,000 tweets related to the Polish 2020 Presidential Elections gathered between February 2, 2020 and April 23, 2020. The data acquisition and sentiment assignment procedures were similar to those described in Section 2. The dataset and model we propose enabled us to analyze sentiment polarity across all election promise categories. Figure 3 shows the overall average sentiment categories. The sentiment has been assigned on a -1 (negative) to 1 (positive) scale. None of the categories were positive on average; hence for readability we show only the zoomed part of the graph with a scale from -0.5 (moderate negative) to 0 (neutral) sentiment. All categories contain a much more negative attitude on the scale of an absolute sentiment analysis score. As we can imagine, most tweets gathered by us are from potential voters, and there are many more negative than positive messages. Most of the sentiment analysis tools available right now perform only general sentiment detection, saying only how many positive or negative tweets they have identified and analyzed. However, our dataset and model enable us to go deeper into the analysis of attitudes towards particular political advertising categories or even more granular towards specific election promises.

Category	P	R	F1	#
Healthcare	0.75	0.72	0.74	586
Welfare	0.76	0.60	0.67	526
Defense	0.92	0.71	0.80	57
Legal	0.79	0.59	0.68	352
Education	0.81	0.57	0.67	163
Infrastructure	0.77	0.59	0.67	284
Society	0.82	0.67	0.74	386
Foreign Policy	0.86	0.43	0.57	93
Immigration	0.89	0.67	0.76	60
	0.82	0.62	0.70	2507

Table 2: Precision, Recall, and F1 score for each of the promises categories. The last column presents the number of examples in the dataset.

4 Conclusions and Future Work

A new dataset and model enables us to analyze the Polish political scene, counter political misinformation in social media, and evaluate the political advertising of candidates. We plan to work on more datasets and models to fight fake news, classify political agitation content, and widen natural language solutions in regards to elections and political content in Polish social media. The dataset annotation will be very challenging due to many potential concept drifts between different election types such as presidential, parliamentary, European Union, and others. We use political advertising model to generate presidential candidates' vector representations. We can compare candidates with each other and say who is similar to whom.

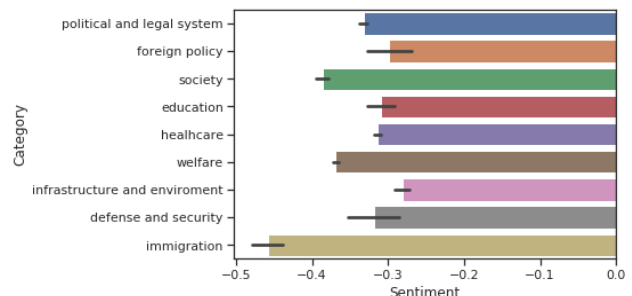


Figure 1: Sentiment orientation for each of political advertising category. The bars mean standard deviations.

Acknowledgment

The work was partially supported by the National Science Centre, Poland grant No. 2016/21/N/ST6/02366 and by the Faculty of Computer Science and Management, Wrocław University of Science and Technology statutory funds.

References

- Sinan Aral and Dean Eckles. 2019. Protecting elections from social media manipulation. *Science*, 365(6456):858–861.
- Hareesh Bahuleyan and Olga Vechtomova. 2018. UWaterloo at SemEval-2017 Task 8: Detecting Stance towards Rumours with Topic Independent Features. Technical report.
- Petra Saskia Bayerl and Karsten Ingmar Paul. 2011. What determines inter-coder agreement in manual annotations? Ameta-analytic investigation. *Computational Linguistics*, 37(4):699–725, 12.
- Michał Bernaczyk, Tomasz Gąsior, Jan Misiuna, and Maciej Serowanec. 2020. *Znaczenie nowych technologii dla jakości systemu politycznego*. WYDAWNICTWO NAUKOWE UNIwersYTETU MIKOŁAJA KOPERNIKA, Toruń.
- Michael L. Best and Amanda Meng. 2015. Twitter Democracy: Policy versus identity politics in three emerging African democracies. In *ACM International Conference Proceeding Series*, volume 15. Association for Computing Machinery, 5.
- Pete Burnap, Rachel Gibson, Luke Sloan, Rosalynd Southern, and Matthew Williams. 2016. 140 characters to victory?: Using Twitter to predict the UK 2015 General Election. *Electoral Studies*, 41:230–233, 3.
- Andrea Ceron and Giovanna d’Adda. 2016. E-campaigning on Twitter: The effectiveness of distributive promises and negative campaign in the 2013 Italian election. *New Media and Society*, 18(9):1935–1955.
- Raviv Cohen and Derek Ruths. 2013. Classifying political orientation on Twitter: It’s not easy! Technical report.
- Daniel Gayo-Avello. 2012. ”I Wanted to Predict Elections with Twitter and all I got was this Lousy Paper” – A Balanced Survey on Election Prediction using Twitter Data. Technical report.
- Yevgenii Golovchenko, Cody Buntain, Gregory Eady, Leon Yin, Megan Brown, and Joshua Tucker. 2020. Cross-Platform State Propaganda: Russian Trolls on Twitter and YouTube During the 2016 US Presidential Election. *International Journal of Press/Politics*.
- Robert Gorwa. 2017. Computational Propaganda in Poland: False Amplifiers and the Digital Public Sphere. Technical Report 4.
- Edouard Grave, Piotr Bojanowski, Prakhar Gupta, Armand Joulin, and Tomas Mikolov. 2019. Learning word vectors for 157 languages. *LREC 2018 - 11th International Conference on Language Resources and Evaluation*, pages 3483–3487, 2.
- Brian Heredia, Joseph Prusa, and Taghi Khoshgoftaar. 2017. Exploring the Effectiveness of Twitter at Polling the United States 2016 Presidential Election. In *Proceedings - 2017 IEEE 3rd International Conference on Collaboration and Internet Computing, CIC 2017*, volume 2017-Janua, pages 283–290. Institute of Electrical and Electronics Engineers Inc., 12.
- Matthew Honnibal. 2018. spaCy Models Documentation, <https://spacy.io/models>.
- Jim Isaak and Mina J. Hanna. 2018. User Data Privacy: Facebook, Cambridge Analytica, and Privacy Protection. *Computer*, 51(8):56–59, 8.
- Michael W Kearney. 2013. Political Discussion on Facebook: an Analysis of Interpersonal Goals and Disagreement. Technical report, 12.
- Elena Kochkina, Maria Liakata, and Isabelle Augenstein. 2018. Turing at SemEval-2017 Task 8: Sequential Approach to Rumour Stance Classification with Branch-LSTM. Technical report.
- Mirko Lai, Viviana Patti, Giancarlo Ruffo, and Paolo Rosso. 2018. Stance evolution and twitter interactions in an italian political debate. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10859 LNCS:15–27.

- M Lai. 2019. On Language and Structure in Polarized Communities, <https://riunet.upv.es/handle/10251/119116>. 4.
- Mary L. McHugh. 2012. Interrater reliability: The kappa statistic. *Biochemia Medica*, 22(3):276–282.
- Nugroho Dwi Prasetyo and Claudia Hauff. 2015. Twitter-based election prediction in the developing world. In *HT 2015 - Proceedings of the 26th ACM Conference on Hypertext and Social Media*, pages 149–158. Association for Computing Machinery, Inc, 8.
- Michel Rosenfeld. 2003. Hate Speech in Constitutional Jurisprudence: A Comparative Analysis. *Cardozo Law Review*, 24, 1.
- Sippo Rossi, Matti Rossi, Bikesh Upreti, and Yong Liu. 2020. Detecting Political Bots on Twitter during the 2019 Finnish Parliamentary Election. *Proceedings of the 53rd Hawaii International Conference on System Sciences*, 1.
- Eli Skogerbø and Arne H. Krumsvik. 2015. Newspapers, facebook and twitter: Intermedial agenda setting in local election campaigns. *Journalism Practice*, 9(3):350–366, 5.
- Mariona Taulé, Francisco Rangel, M. Antònia Martí, and Paolo Rosso. 2018. Overview of the task on multimodal stance detection in Tweets on catalan #1Oct referendum. Technical report.
- Andranik Tumasjan, Timm O. Sprenger, Philipp G. Sandner, and Isabell M. Welp. 2010. Predicting elections with Twitter: What 140 characters reveal about political sentiment. Technical report.
- Jannis Vamvas and Rico Sennrich. 2020. X-Stance: A Multilingual Multi-Target Dataset for Stance Detection. 3.
- Ussama Yaqub, Soon Ae Chun, Vijayalakshmi Atluri, and Jaideep Vaidya. 2017. Analysis of political discourse on twitter in the context of the 2016 US presidential elections. *Government Information Quarterly*, 34(4):613–626.