Propaganda Detection in Russian Telegram Posts in the Scope of the Russian Invasion of Ukraine

Natalia Vanetik [1] natalyav@sce.ac.il
Marina Litvak [1] marinal@sce.ac.il
Egor Reviakin [1] igorre@ac.sce.ac.il
Margarita Tyamanova [1] margati@ac.sce.ac.il

[1]Shamoon College of Engineering, Beer-Sheva, Israel

Abstract

The emergence of social media has made it more difficult to recognize and analyze misinformation efforts. Popular messaging software Telegram (Durov, 2013) has developed into a medium for disseminating political messages and misinformation, particularly in light of the conflict in Ukraine (Wikipedia contributors, 2023). In this paper, we introduce a sizable corpus of Telegram posts containing pro-Russian propaganda and benign political texts. We evaluate the corpus by applying natural language processing (NLP) techniques to the task of text classification in this corpus. Our findings indicate that, with an overall accuracy of over 96% for confirmed sources as propagandists and oppositions and 92% for unconfirmed sources, our method can successfully identify and categorize pro-Russian propaganda posts. We highlight the consequences of our research for comprehending political communications and propaganda on social media.

1 Introduction

Because of social networks' rising use in daily life, we increasingly rely on other people's opinions when making both big and minor decisions, such as whether to vote for a new government or buy new products online. It is not surprising that by spreading propaganda, social media became a weapon of choice for manipulating public opinion. Social media is rife with fake content and propaganda, which needs to be identified and blocked or removed. Recent years have seen a major increase in the issue of information authenticity on social media, leading to significant research community efforts to address fake news (Pariser, 2011), clickbait (Chen et al., 2015b), fake reviews (Akoglu et al., 2013), rumors (Hamidian and Diab, 2016), and other types of misinformation. In this paper, we deal with Russian state-sponsored propaganda disseminated in Telegram. Telegram is one of the most widely used venues for information sharing in Russia, especially after blocking META Platforms. Therefore, Telegram draws much attention from organized groups that spread similar views through its channels and (most probably) funded by either state or related organizational sources. To influence the public to favor the war, the Russian government implemented new regulations that gave it control over traditional media channels (Geissler et al., 2022). The fundamentals of propaganda communication: persuasion using symbols, emotions, stereotypes, and pre-existing frameworks with the intention of swaying perceptions and influencing cognition and behavior in order to further the propagandist's agenda (Alieva et al., 2022). Our work focuses on specific pro-Russian propaganda during the conflict between Russia and Ukraine. Several researchers have documented Russian propaganda during previous conflicts (Golovchenko, 2020; Geissler et al., 2022).

This paper has two contributions: (1) it introduces a new dataset of posts about the Russia-Ukraine war in Russian, collected from Telegram channels and annotated with binary propaganda-related labels; (2) it reports the results of our case study on this dataset, where we examine a supervised method for propaganda detection.

2 Related Work

Propaganda is the spread of information to influence public opinion or behavior, and it is a growing concern in today's digital era. With the vast amount of digital media available, it can be challenging to differentiate between genuine information and propaganda.

In recent years, there has been a growing interest in using machine learning techniques for propaganda detection. Numerous studies have attempted to classify texts' propagandistic content (Rashkin et al., 2017). For instance, Martino et al. (2019) allows analyzing texts at a finer level by identifying all passages that contain propaganda tactics and

Proceedings of Recent Advances in Natural Language Processing, pages 1162–1170
Varna, Sep 4–6, 2023
https://doi.org/10.26615/978-954-452-092-2_123

their types. A corpus of news articles was created and manually annotated at the fragment level with eighteen propaganda techniques. Authors Yoosuf and Yang (2019) used the Fragment Level Classification (FLC) task dataset consisting of news articles from various sources, each annotated with labels representing one out of 18 predefined techniques. The goal of the task introduced in Yoosuf and Yang (2019) was to predict the propaganda techniques associated with each text fragment in the articles. The authors fine-tuned a BERT model on the FLC task dataset using a multitask learning approach, where the model is simultaneously trained to perform both fragment-level and article-level classification. Another paper, Khanday et al. (2021), proposed a supervised learning approach using Support Vector Machine (SVM) to classify news articles as propaganda or non-propaganda. Despite demonstrating fairly good accuracy, the aforementioned studies are mostly limited to English.

The recent political developments have increased the number of Russian-language studies. Topic modeling is one of the methods that have been successfully applied in the field of NLP. In this article, Yakunin et al. (2020) suggests a method for identifying texts that contain propaganda by leveraging a text corpus's topic model. With the suggested method, analysis is attempted at a much higher level of abstraction (themes and the relationships between texts and subjects rather than individual words in a phrase). Other researchers in Park et al. (2022) analyzed agenda creation, framing, and priming—three tactics that underlie information manipulation using both established and newly developed NLP models on VoynaSlov (38M+ posts from Twitter and VKontakte in Russian), revealing variance across media outlet control, social media platform and time. A structured topic model (STM) and a contextualized neural topic model were both used. Another researcher used news stories and Telegram news channels in Ukrainian, Russian, Romanian, French, and English to examine how the media influenced and reflected public opinion during the first month of the war between Ukraine and Russia (Solopova et al., 2023). The existing literature on propaganda detection offers a wide variety of methods, datasets, and perspectives that can be used to develop effective and responsible propaganda detection systems.

To the best of our knowledge, our dataset is a large dataset of political posts with substantial differences between pro-war and anti-war Telegram posts about the Russia-Ukraine war.

3 Case Study

3.1 The Dataset

Telegram channels are widely used in Russia because they are simple, usually focus on short text posts, and do not need special personal verification. Everyone can create a channel anonymously and start posting any type of information without any validation or fact-checking. In addition, the CEO of Telegram, Pavel Durov, advertised Telegram as an independent and the most protected messenger in the world marketplace (Durov, 2014).

We used Telegram API (Telegram, 2021) to extract texts from Telegram (Durov, 2013) channels representing Russian government official sources and opposition political sources into our dataset (described below in Figure 1). We have selected a period for downloading texts from the 24th of February 2022 to the 24th of February 2023, as the first year of the Russia's full invasion of Ukraine. We relied on the EU sanction list (European External Action Service, Accessed on 14 May 2023) to assign texts to a propaganda or benign category. For example, "Channel One Russia" has been added to the sanction list as a government company (council of the EU, 2014), and "SolovievLive," the personal telegram-channel of Vladimir Roudolfovitch Soloviev (council of the EU, 2023), has been added to the list as an individual propagandist who works at the government channels. Benign political text sources have been selected from the channels declared to be Foreign Agents by the Ministry of Justice of the Russian Federation (according to the Russian Foreign Agents law (The Federal Assembly of the Russian Federation, 2022)), and as such, are unlikely to contain pro-Russian propaganda. The Russian Foreign Agents Law (The Federal Assembly of the Russian Federation, 2022) is described as a freedom-restricting law by the International Center for Not-for-Profit Law (International Center for Not-for-Profit Law (ICNL), 2021), an international non-governmental organization that works to promote and protect the right to freedom of association, assembly, and expression for civil society organizations and individuals around the world.

The list of sources for two classes in our dataset is listed in Figure 1 – we provide the name of the

Telegram channel, its translation, and the channel's ID. Figure 2 contains two representative examples of propaganda and benign political texts along with their English translation.

Texts have been downloaded from Telegram channels with two filters: seed words for each class and the post length greater than or equal to 80. According to the article about how text characteristics impact user engagement in social media, posts greater than or equal to 80 characters are "easy to read", and they get a better user engagement (Gkikas et al., 2022). Seed words for propaganda sources have been chosen from the articles about the Russian-Ukrainian war (Umland, 2022), (Ganchev et al., 2022). The opposition's seed words are neutral synonyms of the propaganda's words. These seed words are listed in Figure 3. We used seed words for searching and downloading posts only related to the war, except for advertising posts and other subjects in Telegram channels. The dataset available on GitHub (2023).

3.1.1 Data Analysis

Tables 1-2 contain basic statistics of the data, including the number of documents in every class for each set, in total, and the average number of words in a document. The positive class (propaganda) contains 6038 texts and the negative class (benign) contains 5282 texts.

Text length analysis (in characters) shows propaganda texts tend to be longer, while benign texts in general are shorter and their length distribution is different (no big differences between thresholds). A comparison of these distributions appears in Figure 4.

During our research, we underline, for example, that the word 'НА' (meaning 'on') is prominent in propaganda texts because of the Russian expression 'on Ukraine' used in Russia contrary to the expression 'in Ukraine' used in Ukraine.

The variance threshold (Kohavi and John, 1997) serves as a straightforward method for feature selection, wherein features failing to meet a certain threshold for variance are eliminated. Specifically, it eliminates features with zero variance, meaning those that have identical values across all samples, as the default criterion. Figure 5 shows the most important words extracted with this method for two classes in our data - benign texts and propaganda texts - for different values of the threshold. We can see that for a variance threshold of 0.7 or above no words are found for the benign class, implying that this class contains only lower-variance features (meaning that the values of word features across the class do not vary much or are very similar). However, given a smaller threshold, the phrase "foreign agent" is selected for the benign class.

3.2 Data Representation

Besides expanding our training set, a universal solution might be developed if we find a "typical" writing style or dissemination of propaganda in general across different domains.

The following techniques were employed for the text representation:

1. Term frequency-inverse document frequency (tf-idf), which increases proportionally to the number of times a word appears in the document and is offset by the number of documents in the corpus that contain the word. Terms represent vector dimensions, while their tf-idf scores represent vector values. Every text item is treated as a separate document and the whole dataset as a corpus for computing tf-idf weights.

2. Word n-grams consisting of n consecutive words seen in the text, where n is a parameter. Each text is represented by a vector with N-grams as dimensions and their counts as values. In our evaluation, we used the values $n = 1, 2, 3$.

3. BERT sentence embeddings using one of the pre-trained BERT models:
   - a multilingual model (Sanh et al., 2019)
   - Russian-language BERT model (Arkhipov et al., 2019).

3.3 Classification Pipeline

Our classification pipeline consists of a few steps.

1. Representing texts with tf-idf vectors, word n-grams with $n = 1, 2, 3$, or pre-trained BERT sentence vectors.

2. Training and application of the following classifiers:
   - Traditional ML models (see Section 3.4) are applied to all of the above data representations.
   - Fine-tuned pre-trained BERT models are applied to raw texts residing in

Official Russian sources			Benign political texts sources		
channel name	translated name	ID	channel name	translated name	ID
Первый канал. Новости	Channel One Russia	1390	Медуза LIVE	Meduza	1313
Минобороны России	The MoD of Russia	991	Медиазона	Mediazone	864
ВЕСТИ	Vesti	1602	Телеканал Дождь	TV channel Rain	1269
Кремль. Новости	Kremlin. News	96	Важные истории	Important stories	571
СОЛОВЬЁВ	SolovievLive	1984	The Insider	The Insider	1240

Figure 1. Telegram channels used for data extraction.

Benign political text	Propaganda text
В политическом смысле это одно из самых крупных поражений России в этой войне. Однако с военной точки зрения, все несколько сложнее, поскольку российские войска займут более выгодные позиции, которые легче снабжать. Но и тут есть свои «но».	Уничтожен штаб укронацистов на базе фк металлист в харьковской области в спорткомплексе в высоком базировались украинские боевики. Российские военные вычислили штаб горе вояк и успешно провели денацификацию.
Translation	
Politically, this is one of the most Russia's major defeats in this war. However from a military point of view, everything is somewhat more complicated, since Russian troops will occupy more than advantageous positions that are easier to supply. But also there are "buts" here.	The headquarters of the Ukronazis at the base was destroyed fc metalworker in the kharkiv region in sports complex in a high-based Ukrainian fighters. Russian military calculated the headquarters of the mountain warrior and successfully carried out denazification.

Figure 2. Representative examples from our dataset.

Benign political texts	Propaganda
пророссийский, война, в украине, наступление, атака, взрыв, обстрел, трагедия, минобороны, генштаб, отступление, погибший, дискредитация, спецоперация, беспилотник, санкции, z, пропагандист, военкор, нато	спецоперация, специальная военная операция, денацификация, на украине, z, демилитаризация, хаймарсы, нато, украинские боевики, националист, нацист, укронацист, киевский режим, недружественные, санкции, военкор, дискредитация, беспилотник, укропы
Translation	
pro-Russian, war, in Ukraine, offensive, attack, explosion, shelling, tragedy, Ministry of Defense general staff, retreat, lost, discredit, special operation, drone, sanctions, z, propagandist, military commissar, NATO	special operation, special military operation, denazification, in Ukraine, z, demilitarization, Hymars, NATO, Ukrainian militants, nationalist, nazi, ukronazi, Kyiv regime, unfriendly, sanctions, commander, discredit, drone, dill (derogatory nickname for Ukrainians)

Figure 3. Seed words used for data filtering.

training documents	validation documents	test documents	total documents	min words	max words	avg wc	unique words
8214	913	2193	11320	6	631	97.34	81152

Table 1. Data statistic.

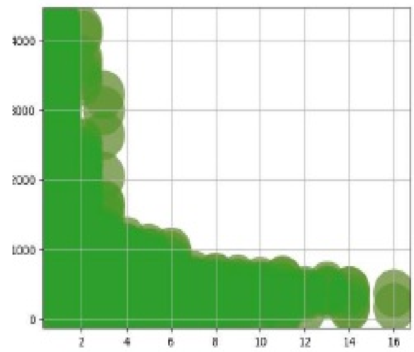
training documents		validation documents		test documents	
propaganda	benign texts	propaganda	benign texts	propaganda	benign texts
4385	3829	491	422	1162	1031

Table 2. Class balance.

the training data and then classifying the test data. We use a multilingual BERT model (Sanh et al., 2019), and a pre-trained model by DeepPavlov AI (Arkhipov et al., 2019) that is pre-trained on Russian News and four parts of Wikipedia: Bulgarian, Czech, Polish, and Russian.

3.4 Traditional ML Classifiers

We have applied three traditional classifiers – Random Forest (RF) (Pal, 2005), Logistic Regression (LR) (Wright, 1995), and Extreme Gradient Boost-



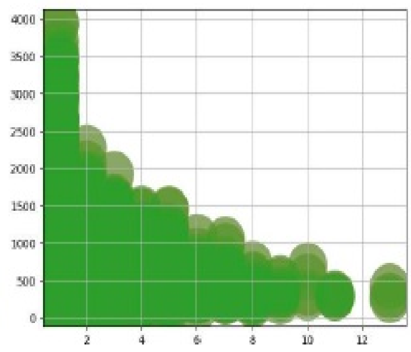


Figure 4. Texts lengths (in characters) for propaganda (left) and benign texts (right) distribution.

threshold	propaganda texts	benign texts
1	всу, народной, области, республики	–
0.9	всу, народной, области, район, республики	–
0.8	всу, направлении, народной, области, район, республики	–
0.7	всу, донецкой, направлении, народной, области, пункт, район, республики	–
0.6	военной, всу, донецкой, направлении, народной, области, пункт, район, республики, россии	россии
0.5	военной, всу, донецкой, направлении, народной, населенных, области, пункт, район, республики, россии, рф, украинских	агента, выполняющим, иностранного, области, россии, функции
0.4	военной, всу, донецкой, направлении, народной, населенных, области, пункт, район, республики, россии, рф, специальной, сша, украинских	агента, выполняющим, иностранного, области, россии, украины, функции
	Translation	
1	AFU (Armed Forces of Ukraine), national, region, republic	–
0.9	AFU (Armed Forces of Ukraine), national, region, republic	–
0.8	AFU (Armed Forces of Ukraine), direction, national, region, district, republic	–
0.7	AFU (Armed Forces of Ukraine), Donetsk, direction, people's, region, point, district, republic	–
0.6	military, AFU (Armed Forces of Ukraine), Donetsk, direction, national, region, point, district, republic, Russia	Russia
0.5	military, AFU (Armed Forces of Ukraine), Donetsk, direction, national, populated, region, point, district, republic, Russia, RF (Russian Federation), Ukrainian	agent, performing, foreign, region, Russia, functions
0.4	military, AFU (Armed Forces of Ukraine), Donetsk, direction, national, populated, region, point, district, republic, Russia, RF (Russian Federation), special, united states, Ukrainian	agent, performing, foreign, region, Russia, Ukraine, functions

Figure 5. Most important features (words) extracted with variance threshold method using NLTK Russian stopwords.

ing (XGB) (Chen et al., 2015a) to all text representations described in Section 3.2.

3.5 Results

Table 3 demonstrates the results for the traditional classifiers and text representations. The text representations use either word vectors with tf-idf (aka Vector Space Model) or n-grams with count weights (for $n = 1, 2, 3$). All the systems are significantly better than the majority rule. Also, the Logistic Regression (LR) classifier with unigrams outperforms the other classifiers and representations. In general, LR shows better performance than other classifiers (RF and XGB) for all text representations used in this experiment.

Table 4 shows classification results for two fine-tuned BERT models – a DeepPavlov model known to perform well on Russian Question Answering task (Zaytsev et al., 2018), and Russian sentiment analysis tasks (Chernykh et al., 2021), and a multilingual BERT model (Sanh et al., 2019) for comparison. Both models were trained for 15 epochs with batch size 16, a learning rate of 2e$-$5. Train-

N	propaganda texts	benign texts
10	всу, россии, рф, vestiru24, военной, области, украинских, украины, минобороны	агента, функции, выполняющим, иностранного, россии, области, украины, войны
20	всу, россии, рф, vestiru24, военной, области, украинских, минобороны, украине, специальной, направлении, операции, нато, сша, районе, спецоперации, россия	агента, функции, выполняющим, иностранного, россии, области, украины, войны, всу, минобороны, российским, человек, лицом, юридическим, читайте
	Translation	
10	AFU (Armed Forces of Ukraine), Russia, RF (Russian Federation), vestiru24, military, region, Ukrainian, Ukraine, ministry of defense	agent, functions, performing, foreign, Russia, region, Ukraine, war
20	AFU (Armed Forces of Ukraine), Russia, RF (Russian Federation), vestiru24, military, region, Ukrainian, ministry of defense, Ukraine, special, direction, operation, NATO, USA, district, special operation, Russia	agent, functions, performing, foreign, Russia, region, Ukraine, war, AFU (Armed Forces of Ukraine), ministry of defense, Russian, human, entity, legal, read

Figure 6. Top N words per class, ranked by their tf-idf weights (different morphological forms of the same words omitted).

representation	classifier	P	R	F1	acc
ML BERT SE	RF	0.8288	0.8182	0.8206	0.8240
	LR	0.8800	0.8808	0.8803	0.8810
	XGB	0.8377	0.8342	0.8354	0.8372
DeepPavlov SE	RF	0.8415	0.8320	0.8343	0.8372
	LR	0.8906	0.8911	0.8909	0.8915
	XGB	0.8483	0.8466	0.8473	0.8486
tf-idf	RF	0.9390	0.9328	0.9349	0.9357
	LR	0.9431	0.9349	0.9376	0.9384
	XGB	0.9133	0.8986	0.9023	0.9042
unigrams	RF	0.9289	0.9212	0.9237	0.9248
	LR	**0.9481**	**0.9448**	**0.9461**	**0.9466**
	XGB	0.9086	0.8928	0.8966	0.8988
bigrams	RF	0.8965	0.8681	0.8728	0.8769
	LR	0.9203	0.9080	0.9113	0.9129
	XGB	0.8620	0.8326	0.8365	0.8422
trigrams	RF	0.8982	0.8700	0.8747	0.8787
	LR	0.9203	0.9080	0.9113	0.9129
	XGB	0.8633	0.8340	0.8379	0.8436

Table 3. Traditional classifier baselines applied to sentence embeddings, n-grams, and tf-idf text representations.

Bert model	benign class			propaganda class			
	P	R	F1	P	R	F1	acc (macro avg)
DeepPavlov	0.9452	0.9762	0.9605	0.9791	0.9518	0.9653	0.9630
ML BERT	0.9457	0.9682	0.9569	0.9724	0.9527	0.9624	0.9598

Table 4. Fine-tuned BERT results.

Bert model	benign class			propaganda class			
	P	R	F1	P	R	F1	acc (macro avg)
DeepPavlov	0.9649	0.8678	0.9138	0.8907	0.9715	0.9293	0.9223
BERT ML	0.9466	0.8757	0.9098	0.8950	0.9555	0.9242	0.9176

Table 5. Fine-tuned BERT results for dataset ”without seed words”.

ing accuracy and training loss for the top model (DeepPavlov) were 0.9606 and 0.0003, and training time per epoch was approximately 270 seconds. We can see that this model achieves slightly better results than the multilingual BERT and that both fine-tuned models outperform all of the traditional classifiers mentioned in Table 3, although by a small margin.

Moreover, to check our results, we experimented with a dataset without using seed words for searching and downloading texts from Telegram. We extracted new posts from the channels that not used in the training dataset, but sometimes channels from the training dataset reposted posts from these channels. So we can decide on the type of one channel or another. Figure 7 presents Telegram channels

pro-Russian sources		
channel name	translated name	ID
Герои спецоперации Z	Heroes of the special military operation Z	1547226852
АРХАНГЕЛ СПЕЦНАЗА Z	ARCHANGEL SWAT Z	1583313036
Сладков +	Sladkov +	1164348791

benign political texts sources		
channel name	translated name	ID
Михаил Ходорковский	Mikhail Khodorkovsky	1105250846
Проект	Proekt	1190104199
Агенство. Новости	Agency. News	1583655041

Figure 7. Telegram channels used for dataset "without seed words".

for additional datasets. The class balance for the dataset is pro-Russian sources - 562 documents and benign political texts sources - 507. Results of the experiment with dataset "without seed words" and "new channels" in Table 5. In addition, we deployed the model with a Telegram bot API (Modrzyk, 2018). Users can paste a news post about the Russian-Ukrainian war in Russian, and the bot will respond with a special label and score (probability of label). The bot is available at Telegram-bot (2023).

4 Conclusions and Future Work

We are optimistic that our work will help people recognize texts that may not be objective and focus only on producing emotional feelings rather than a rational response. However, our models are trained on political texts that address the conflict in Ukraine and, therefore, cannot recognize propaganda in other domains. In addition, improving the model's ability to handle scenarios such as propaganda statements in stylistically complex texts is essential to develop a more widely trainable model.

We continue improving our model and will soon add a "neutral" class for correct classification. Besides expanding our training set, a universal solution might be developed if we find a "typical" writing style or dissemination of propaganda across different domains.

References

Leman Akoglu, Rishi Chandy, and Christos Faloutsos. 2013. Opinion fraud detection in online reviews by network effects. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 7, pages 2–11.

Iuliia Alieva, JD Moffitt, and Kathleen M Carley. 2022. How disinformation operations against Russian opposition leader Alexei Navalny influence the international audience on Twitter. *Social Network Analysis and Mining*, 12(1):80.

Mikhail Arkhipov, Maria Trofimova, Yurii Kuratov, and Alexey Sorokin. 2019. Tuning multilingual transformers for language-specific named entity recognition. In *Proceedings of the 7th Workshop on Balto-Slavic Natural Language Processing*, pages 89–93.

Tianqi Chen, Tong He, Michael Benesty, Vadim Khotilovich, Yuan Tang, Hyunsu Cho, Kailong Chen, et al. 2015a. Xgboost: extreme gradient boosting. *R package version 0.4-2*, 1(4):1–4.

Yimin Chen, Niall J Conroy, and Victoria L Rubin. 2015b. Misleading online content: recognizing clickbait as" false news". In *Proceedings of the 2015 ACM on workshop on multimodal deception detection*, pages 15–19.

Pavel Chernykh, Anna Mikhailovskaya, Zufar Miftahutdinov, and Dmitry Ustalov. 2021. DeepPavlov Goes to Russia: Baselines and Best Practices for Russian Language Processing. In *International Conference on Computational Linguistics and Intelligent Text Processing*, pages 176–187. Springer.

Pavel Durov. 2013. Telegram. https://telegram.org/. [Online; accessed 14 May 2023].

Pavel Durov. 2014. Why telegram? https://telegram.org/blog/why-telegram. [Online; accessed 14 May 2023].

The council of the EU. 2014. EU restrictive measures against Russia over Ukraine (since 2014). https://www.consilium.europa.eu/en/policies/sanctions/restrictive-measures-against-russia-over-ukraine. [Online; accessed 14 May 2023].

The council of the EU. 2023. Council Implementing Regulation (EU) 2023/571 of 13 March 2023 implementing Regulation (EU) No 269/2014 concerning restrictive measures in respect of actions undermining or threatening the territorial integrity, sovereignty and independence of Ukraine. https://eur-lex.europa.eu/

legal-content/EN/TXT/?uri=uriserv:OJ.LI.2023.075.01.0001.01.ENG. [Online; accessed 14 May 2023].

European External Action Service. Accessed on 14 May 2023. Consolidated list of sanctions. https://www.eeas.europa.eu/eeas/european-union-sanctions_en. [Online; accessed 14 May 2023].

Gancho Ganchev et al. 2022. The Ukrainian War and Economic Liberalism. *Bulgarian Journal of International Economics and Politics*, 2(1):3–11.

Dominique Geissler, Dominik Bär, Nicolas Pröllochs, and Stefan Feuerriegel. 2022. Russian propaganda on social media during the 2022 invasion of Ukraine. *arXiv preprint arXiv:2211.04154*.

GitHub. 2023. Russian propaganda and benign texts dataset. https://github.com/Sharik25/propaganda_dataset. [Online; accessed 14 May 2023].

Dimitris C Gkikas, Katerina Tzafilkou, Prokopis K Theodoridis, Aristogiannis Garmpis, and Marios C Gkikas. 2022. How do text characteristics impact user engagement in social media posts: Modeling content readability, length, and hashtags number in Facebook. *International Journal of Information Management Data Insights*, 2(1):100067.

Yevgeniy Golovchenko. 2020. Measuring the scope of pro-Kremlin disinformation on Twitter. *Humanities and Social Sciences Communications*, 7(1):1–11.

Sardar Hamidian and Mona Diab. 2016. Rumor identification and belief investigation on Twitter. In *Proceedings of the 7th Workshop on computational approaches to subjectivity, sentiment and social media analysis*, pages 3–8.

International Center for Not-for-Profit Law (ICNL). 2021. International Center for Not-for-Profit Law. https://www.icnl.org/resources/civic-freedom-monitor/russia. [Online; accessed 14 May 2023].

Akib Mohi Ud Din Khanday, Qamar Rayees Khan, and Syed Tanzeel Rabani. 2021. Svmbpi: support vector machine-based propaganda identification. In *Cognitive Informatics and Soft Computing: Proceeding of CISC 2020*, pages 445–455. Springer.

Ron Kohavi and George H John. 1997. Wrappers for feature subset selection. *Artificial intelligence*, 97(1-2):273–324.

Giovanni Da San Martino, Seunghak Yu, Alberto Barrón-Cedeño, Rostislav Petrov, and Preslav Nakov. 2019. Fine-grained analysis of propaganda in news articles. *arXiv preprint arXiv:1910.02517*.

Nicolas Modrzyk. 2018. *Building telegram bots: develop bots in 12 programming languages using the telegram bot API*. Apress.

Mahesh Pal. 2005. Random forest classifier for remote sensing classification. *International journal of remote sensing*, 26(1):217–222.

Eli Pariser. 2011. *The Filter Bubble: What the Internet Is Hiding from You*. Penguin UK.

Chan Young Park, Julia Mendelsohn, Anjalie Field, and Yulia Tsvetkov. 2022. Challenges and Opportunities in Information Manipulation Detection: An Examination of Wartime Russian Media. *Findings of the Association for Computational Linguistics: EMNLP 2022*, pages 5209–5235.

Hannah Rashkin, Eunsol Choi, Jin Yea Jang, Svitlana Volkova, and Yejin Choi. 2017. Truth of varying shades: Analyzing language in fake news and political fact-checking. In *Proceedings of the 2017 conference on empirical methods in natural language processing*, pages 2931–2937.

Victor Sanh, Lysandre Debut, Julien Chaumond, and Thomas Wolf. 2019. Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter. *arXiv preprint arXiv:1910.01108*.

Veronika Solopova, Oana-Iuliana Popescu, Christoph Benzmüller, and Tim Landgraf. 2023. Automated multilingual detection of Pro-Kremlin propaganda in newspapers and Telegram posts. *Datenbank-Spektrum*, pages 1–10.

Telegram. 2021. Telegram API. https://core.telegram.org/api. [Online; accessed 14 May 2023].

Telegram-bot. 2023. Telegram-bot. https://t.me/lovely_grandchild_bot. [Online; accessed 14 May 2023].

The Federal Assembly of the Russian Federation. 2022. Federal law of July 14, 2022 No.255-FZ On Control over the Activities of Persons under Foreign Influence. http://publication.pravo.gov.ru/Document/View/0001202207140041. [Online; accessed 14 May 2023].

Andreas Umland. 2022. Two scenarios for ukraine's “demilitarization” by russia: Why it is unlikely that ukrainians will disarm. *Available at SSRN*.

Wikipedia contributors. 2023. Conflict in Ukraine. [Online; accessed 14 May 2023].

Raymond E Wright. 1995. Logistic regression.

Kirill Yakunin, Ionescu George Mihail, Murzakhmetov Sanzhar, Mussabayev Rustam, Filatova Olga, and Mukhamediev Ravil. 2020. Propaganda identification using topic modelling. *Procedia Computer Science*, 178:205–212.

Shehel Yoosuf and Yin Yang. 2019. Fine-grained propaganda detection with fine-tuned BERT. In *Proceedings of the second workshop on natural language processing for internet freedom: censorship, disinformation, and propaganda*, pages 87–91.

Andrey Zaytsev, Yuri Kuratov, Dmitry Ustalov, and Andrey Kutuzov. 2018. DeepPavlov at RusEval 2018: Russian Question Answering and Machine Reading Evaluation. In *First Workshop on Evaluation of Human Language Technologies for Slavic Languages*, pages 1–6.