Method for Political Propaganda Detection in Internet Content Using Recurrent Neural Network Models Ensemble

Iurii Krak[1,2], Volodymyr Didur[3], Maryna Molchanova[3,*], Olexander Mazurets[3], Olena Sobko[3], Olha Zalutska[3] and Olexander Barmak[3]

[1] *Taras Shevchenko National University of Kyiv, Ukraine*

[2] *Glushkov Institute of Cybernetics of NAS of Ukraine, Kyiv, Ukraine*

[3] *Khmelnytskyi National University, Khmelnytskyi, Ukraine*

Abstract

The automation of propaganda detection processes in Internet content using natural language processing is extremely relevant in modern conditions and can provide fast and well-timed targeted detection of hostile manipulative influence in largescale amounts of Internet content. The paper proposes a method of automated propaganda detection that operates in the Ukrainian language. The method for detecting political propaganda in Internet content using ensemble of recurrent neural network models is intended to identify and analyze potentially propagandistic or manipulative content spread on the Internet. The input data of the method is an ensemble of trained models of recurrent neural networks with tokenizers and a text message for analysis. The output data are the level and percentage of propaganda presence for each neural network model of ensemble and in general.

To examine the effectiveness of developed method for detecting political propaganda in Internet content, which includes the ensemble use of recurrent neural network models of the BiLSTM and GRU architectures, a software implementation of the method was created. The software implementation allows training neural network models and using them to detect political propaganda in textual Internet content. The training data set in Ukrainian was prepared.

The applied efficiency research of propaganda detection by an ensemble of classifiers based on the BiLSTM and GRU recurrent neural network architectures was conducted. The proposed approach is capable of detecting political propaganda by an ensemble of RNN models with Accuracy 0.97, Precision 0.973, Recall 0.981, and F1 0.976 in the bagging mode, and Accuracy 0.95, Precision 0.977, Recall 0.987, and F1 0.981 in the stacking mode. The developed method has a limitation: it works with text posts from 200 to 6300 symbol long. For shorter and longer texts, performance degradation is observed.

Keywords

propaganda detection, recurrent neural networks, ensemble of neural networks, natural language processing

1. Introduction

Propaganda is an integral component of information manipulation and includes various forms, methods and means of influencing people in order to change their psychological attitudes in the desired direction, so its timely detection is an urgent task of information technologies. Such manipulations are often used to change the psychological climate in society, mobilize support or discredit opponents [1].

In connection with the growth of the consumption of textual Internet content, the threat of the propagandistic destructive manipulative influence of textual political media is growing. Propaganda, which is distributed on the Internet, represents a large-scale threat to the national security of the country [2], the untimely solution of which can lead to devastating consequences [3]. Therefore, the automation of the processes of detecting propaganda in textual Internet content by means of natural language processing is extremely relevant in modern conditions [4], and is capable of providing quick and timely targeted detection of hostile manipulative influence in large-scale volumes of Internet content.

2. Related Works

Modern scientific publications highlight the relevance of the problem of automated detection of propaganda in textual Internet content. Research areas dedicated to the intellectualization of propaganda detection processes, which allows avoiding a number of technological problems in monitoring media sources [5], and the problem of separating manifestations of propaganda techniques from other manipulative influences [6, 7] are especially relevant at the moment. It is noted that the elements of the propaganda model include the subject, content, forms and methods [8], as well as means or channels of information transmission [9].

The subject of propaganda is a social group that seeks to influence the audience. The content of propaganda is determined by the subject's social interests and their relation to the interests of society in general. Forms and methods of propaganda are chosen depending on the goals and the audience to be influenced. Media include print media, radio, television, etc. The object of propaganda is the audience or social groups that are the target of influence. Social interests of the subject of propaganda influence its content and choice of forms, methods and means of information transmission [10].

Detecting propaganda using NLP in text is a challenging task due to propaganda's use of subtle manipulation techniques and context dependencies. To solve this problem, the authors of [11] investigated the effectiveness of modern large language models, such as GPT-3 and GPT-4, for detecting propaganda. Experiments were performed using the SemEval-2020 task 11 dataset, which contains news articles tagged with 14 propaganda techniques. The performance of the models was determined by evaluating metrics such as F1 score, precision, and recall, comparing the results to the current state-of-the-art approach using RoBERTa. The obtained results show that GPT-4 achieves results comparable to the current state-of-the-art technology [12].

Statistical analysis of tests [14], multimodal visual-textual object graph attention network [15] are noted as sufficiently promising and effective means of semantic analysis of textual content [13] that can be used to detect propaganda. Also, at the present stage, the use of transformer-based neural network [16, 17], neural network models of complex architecture, such as RoBERTa [18], GPT [19, 20] and recurrent neural network [21], is a relevant direction for automated detection of propaganda. in turn focusing on the detection of our components or propaganda techniques, such as racial propaganda [22] and fake news [23].

At the same time, the authors [24] note that the existing methods of identifying propaganda are primarily focused on identifying the linguistic features of its content. However, these methods usually miss the information presented in the external news environment from which the propaganda news originated and spread. It is noted that methods for detecting propaganda in different languages may differ, depending on the type of language inflection [25]. The authors of [26] analyze how mass media influenced and reflected public opinion during the first month of the Russian invasion using articles and news channels in Telegram in Ukrainian, Russian, Romanian, French, and English. Two methods of multilingual automated identification of pro-Kremlin propaganda based on transformers (BERT) and linguistic features (SVM) were proposed and compared.

The purpose of the article is to create a method for political propaganda detection in internet content using recurrent neural network models ensemble, which will work with the Ukrainian language, as well as its approval.

As part of the research, the task was also completed: preparation of an educational Ukrainian-language data set; development of software that implements the created method; training of an ensemble of neural network classifiers; conducting a study of the effectiveness of the method using the developed software.

The main contribution of the article is the development of a workable method of automated detection of political propaganda in Ukrainian-language texts.

3. Method for Political Propaganda Detection

Considering the insufficient amount of Ukrainian-language data, there is a need to create an own labeled data set that will be used for training neural networks.

2.1. Dataset Preparation

For training models of recurrent neural networks, a data set with the volume of more than 25,000 posts was formed, which were marked according to belonging to the categories "Propaganda" and "Non-propaganda". The lists of propaganda and verified sources were formed according to the official channels of the President and Verkhovna Rada of Ukraine, as well as according to the data of other analytical international authoritative studies and analytical summaries.

To normalize the input data, records with a length of less than 200 and more than 6300 characters were discarded. As a result of data filtering, a set consisting of 21,222 items was obtained, where 10,737 records belong to the "propaganda post" class and 10,485 records to the "non-propaganda post" class.

To normalize the input data, entries shorter than 200 characters and longer than 6300 characters were discarded. As a result of data filtering, a set consisting of 21,222 items was obtained, where 10,737 records belong to the "propaganda post" class and 10,485 records belong to the "non-propaganda post" class. The graph of data distribution by length in characters is shown in Figure 1.
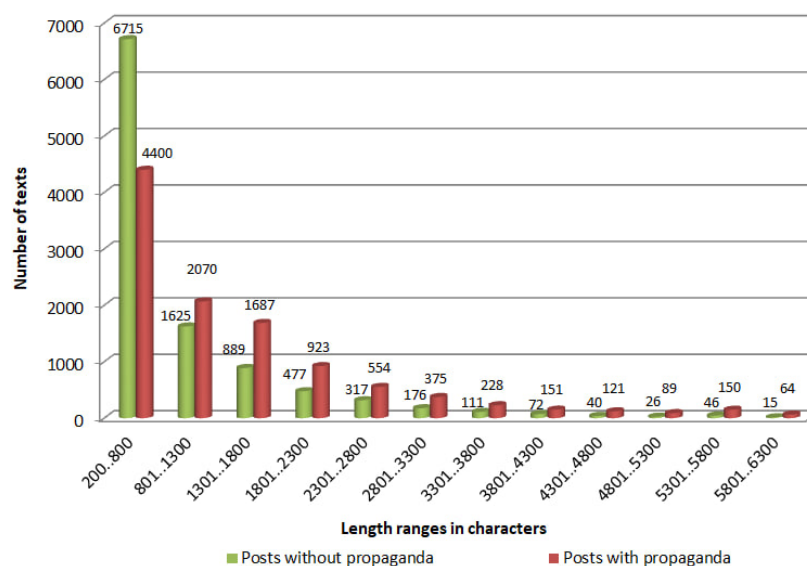

Figure 1: Distribution of dataset elements by the number of characters.

As can be seen from Figure 1, the number of records that do not contain propaganda and are in the length range of 200..800 characters makes up more than half of the set, and this may negatively

affect the quality of classification in the future. At the same time, the set of propaganda texts is more evenly distributed. All multilingual fragments were automatically translated into Ukrainian.

The described data set will be used to train neural network models within the framework of the developed method of detecting political propaganda.

3.2. Scheme of Method for Political Propaganda Detection

The scheme of the method of detecting political propaganda is shown in Figure 2. The input data of the method is an ensemble of trained models of recurrent neural networks with tokenizers, and a text post for analysis. In step 1, the ensemble of RNN models and their tokenizers are selected and loaded.

Next, step 2 pre-processes the user post for analysis, which includes converting the text to lowercase, removing stop chants and punctuation, etc.
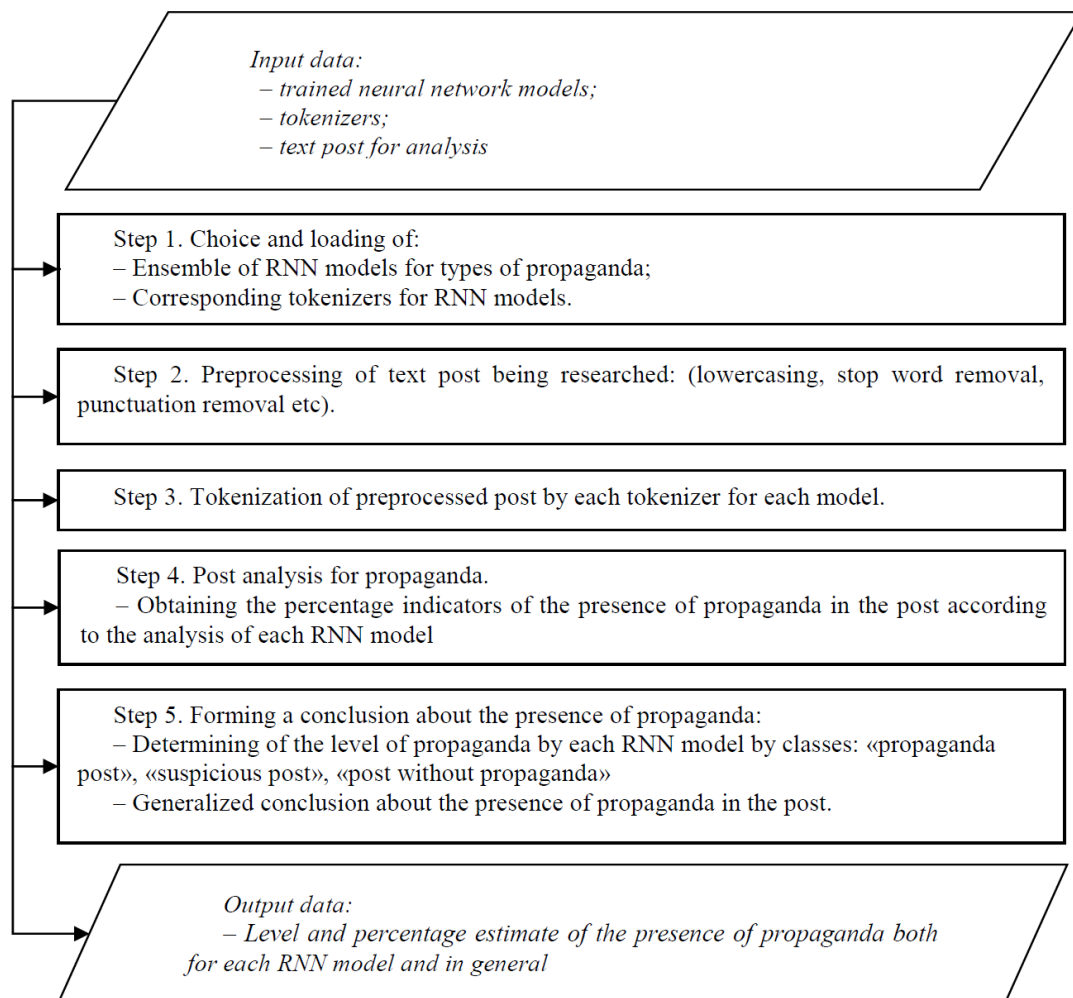

Figure 2: Scheme of method for political propaganda detection in internet content.

In step 3, the pre-processed text is converted into numerical sequences that will be fed to neural networks for further binary classification. Step 4 is the analysis of the post for the presence of propaganda, which includes obtaining the percentage indicators of the presence of propaganda in the post as analyzed by each RNN model.

At step 5, a conclusion is formed regarding the presence of propaganda. It is proposed to use two ensemble approaches - binary (stacking) and discrete (bagging). For the binary approach to determining the level of propaganda for ensemble neural networks, binary scores are obtained, where the score 0 is no propaganda, 1 is propaganda. In the discrete approach, the evaluation of

neural networks is taken as a discrete value from 0 to 1, where 1 is the maximum manifestation of propaganda, and 0 is its absence.

In the case of stacking, a binary score is obtained, and a conclusion about the class of the post is formed according to the rules: "propaganda post", if more than 50% of the models received a binary score of 1; "post without propaganda" if more than 50% of the models received a binary score of 0; "suspicious post" if neural network models have parity voting results (about half with scores of 0 and half with scores of 1).

To determine the level of propaganda in the case of a discrete evaluation, the limits of three classes are determined by experts: the upper limit of the "post without propaganda" class and the lower limit of the "propaganda post" class.

After that, total discrete assessment of the post's belonging to specified classes is calculated (1):

$$Eval = k_1 \cdot RNN_1 + k_2 \cdot RNN_2 + .. + + k_n \cdot RNN_n, \qquad (1)$$

where $k_1, k_2, .., k_n$ – influence coefficients of discrete estimates obtained by neural networks $RNN_1, RNN_2, .., RNN_n$ in accordance.

The influence coefficients of discrete neural network evaluations $k_1, k_2, .., k_n$ are chosen empirically, depending on the focus of the process on detecting the propaganda of the relevant species.

According to the above material, the result of the proposed method is the level and percentage estimate of the presence of propaganda as per each RNN-model of the ensemble, as well as the generalized level and percentage estimate of the presence of propaganda in the researched post.

4. Experiments

An ensemble of two neural network models was formed to conduct an experiment on the effectiveness of the developed method of detecting propaganda in Internet content. In particular, recurrent neural networks of BiLSTM and GRU architectures were used [27]. The selection of different neural network models is due to their specific capabilities for analyzing text sequences [28].

The architectures of BiLSTM and GRU neural networks for detecting propaganda in Internet content are shown in Figure 3.
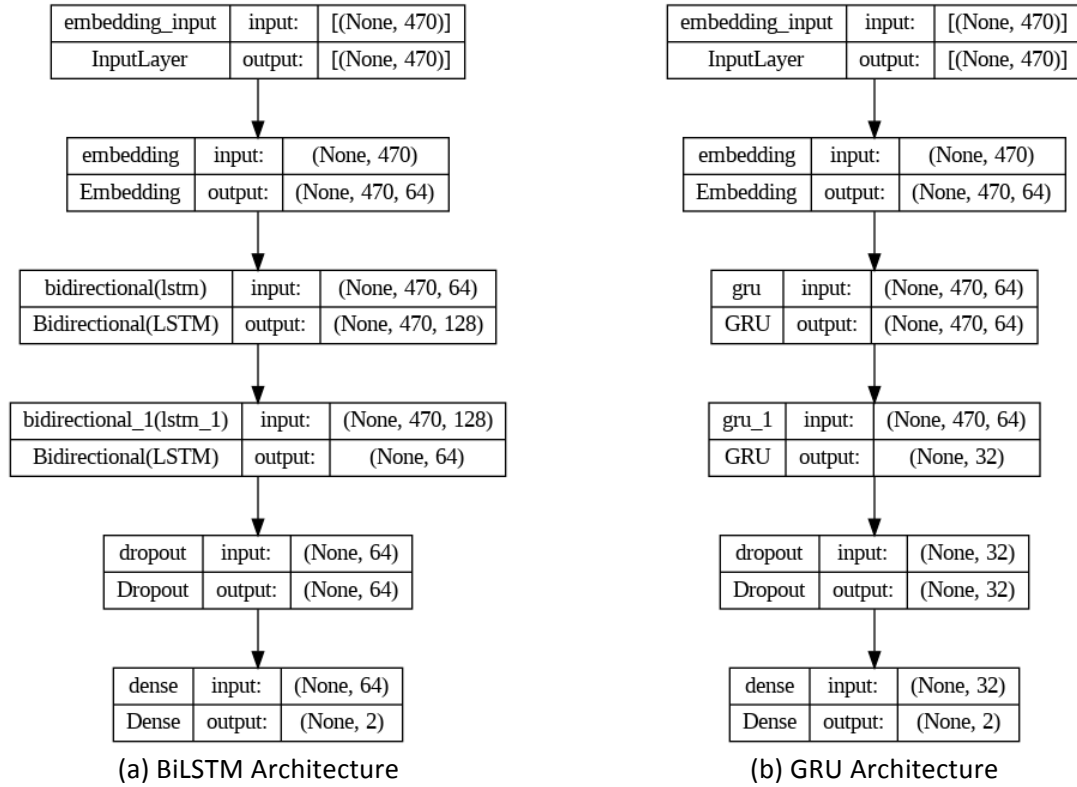
embedding_input	input:	[(None, 470)]
InputLayer	output:	[(None, 470)]

↓

embedding	input:	(None, 470)
Embedding	output:	(None, 470, 64)

↓

bidirectional(lstm)	input:	(None, 470, 64)
Bidirectional(LSTM)	output:	(None, 470, 128)

↓

bidirectional_1(lstm_1)	input:	(None, 470, 128)
Bidirectional(LSTM)	output:	(None, 64)

↓

dropout	input:	(None, 64)
Dropout	output:	(None, 64)

↓

dense	input:	(None, 64)
Dense	output:	(None, 2)

(a) BiLSTM Architecture

embedding_input	input:	[(None, 470)]
InputLayer	output:	[(None, 470)]

↓

embedding	input:	(None, 470)
Embedding	output:	(None, 470, 64)

↓

gru	input:	(None, 470, 64)
GRU	output:	(None, 470, 64)

↓

gru_1	input:	(None, 470, 64)
GRU	output:	(None, 32)

↓

dropout	input:	(None, 32)
Dropout	output:	(None, 32)

↓

dense	input:	(None, 32)
Dense	output:	(None, 2)

(b) GRU Architecture

Figure 3: BiLSTM and GRU Neural Network Architectures for Detecting Propaganda.

BiLSTM by using the hidden state allows the analysis of text sequences in forward and reverse directions, which helps to eliminate the barriers of traditional RNNs in detecting propaganda [29]. GRU has gate mechanisms that allow for more efficient management of gradients in time, which makes it more resistant to the problem of vanishing gradients compared to classical RNNs, which are also able to effectively detect propagation events [30].

In the case of using the BiLSTM and GRU architectures to conduct the experiment, formula (1) will take the form:

$$Eval = k_1 \cdot BiLSTM_r + k_2 \cdot GRU_r, \qquad (2)$$

where k_1 – influence coefficient of the discrete estimate obtained by the neural network BiLSTM, k_2 – influence coefficient of the discrete estimate obtained by the neural network GRU, $BiLSTM_r$ and GRU_r – discrete evaluations of propaganda detection by BiLSTM and GRU neural networks, respectively.

During experiments, neural networks were trained with different parameters (batch, epoch), the results of a comparison of the best models are shown in Table 1.

Table 1
Dependence of metrics on neural network parameters

Parameters:	GRU		BiLSTM	
Batch	32	64	32	64
Epoch	20	20	20	20
Metrics:				

Accuracy	0.97	0.96	0.96	0.95
Loss	0.04	0.06	0.04	0.07

As can be seen from Table 1, GRU has higher accuracy than BiLSTM under the same parameters. Figure 4 shows the distribution of correctly classified texts by the GRU neural network (a) and the distribution of incorrectly classified texts (b).

3573 records were used as validation data, of which 1951 belonged to the "propaganda post" class, and 1622 belonged to the "post without propaganda" class. Of them, 1912 texts of the "propaganda post" class and 1565 texts of the "post without propaganda" class were correctly classified ". 57 texts of the "post without propaganda" class were falsely classified as propaganda by the neural network, and 39 texts of the "propaganda post" class were falsely classified as non-propaganda. The overall accuracy on the validation data is 0.97. As can be seen from the numerical data, the class "post without propaganda" is classified somewhat worse than the class "propaganda post".
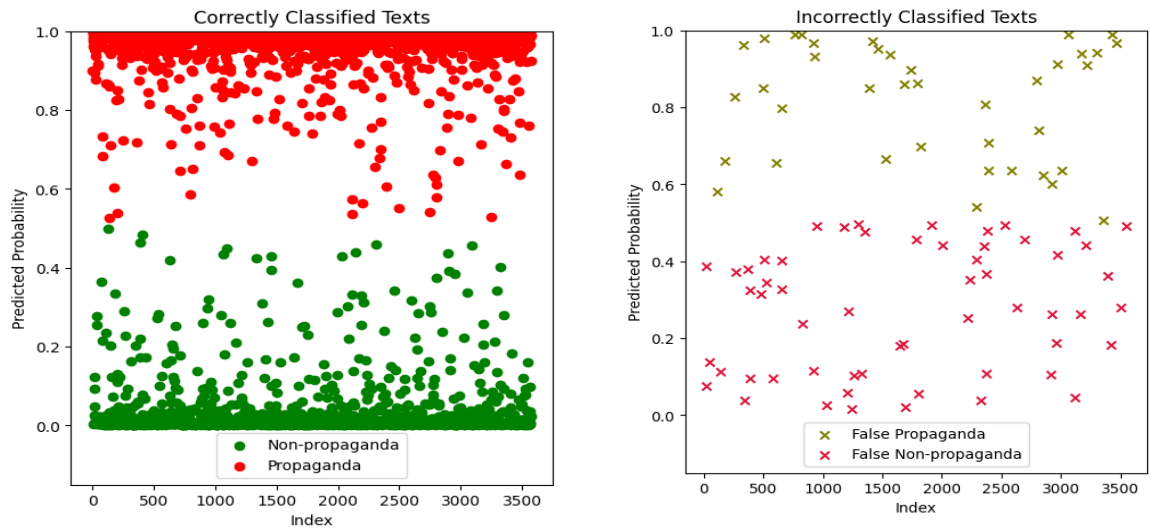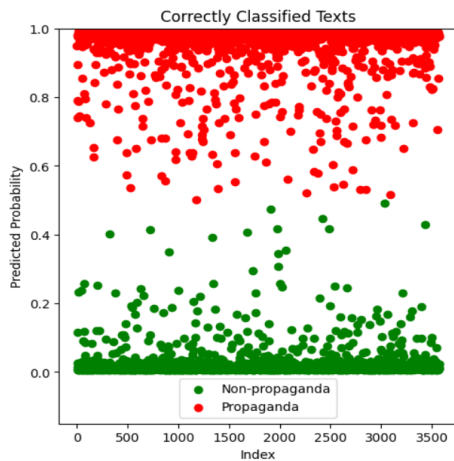

(a) Distribution of correctly classified texts   (b) Distribution of incorrectly classified texts
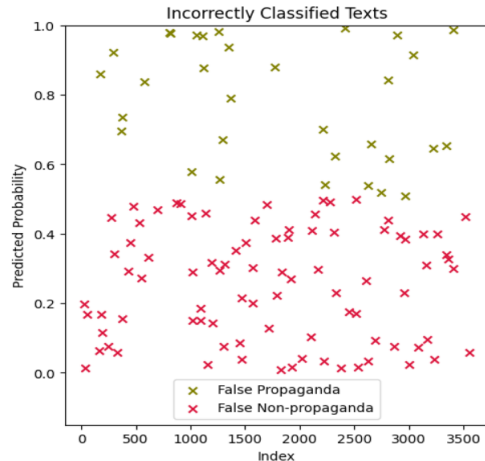**Figure 4:** Distribution of correctly and incorrectly classified texts by the GRU neural network.

Figure 5 shows the distribution of correctly classified texts by the BiLSTM neural network (a) and the distribution of incorrectly classified texts (b).

Out of 3573 validation posts, 1883 posts of the "propaganda post" class and 1572 texts of the "post without propaganda" class were correctly classified. 86 texts of the "post without propaganda" class were falsely classified as propaganda by the neural network, and 32 texts of the "propaganda post" class were falsely classified as non-propaganda. The overall accuracy on the validation data is 0.967.

| (a) Distribution of correctly classified texts | (b) Distribution of incorrectly classified texts |

Figure 5: Distribution of correctly and incorrectly classified texts by the BiLSTM neural network.

As can be seen from Figure 4a and Figure 5a, the texts have a sufficiently high level of interclass resolution, while as can be seen in Figure 4b and Figure 5b, the incorrectly classified data are concentrated closer to the central part of the graphs, which indicates the expediency of the partitioning approach into 3 classes: "propaganda post", "post without propaganda", "suspicious post".

5. Practical Implementation

The software implementation of method of detecting political propaganda shows in Figure 6.

| (a) Neural network learning module | (b) Political propaganda detection module |

Figure 6: The main forms interface of applied software implementation.

To study the effectiveness of the developed method of detecting political propaganda in Internet content, which includes the ensemble use of RNN models of the BiLSTM and GRU architectures, a software implementation of the method was created using the Python language. The interface of the software component responsible for the learning module of neural network models is shown in

Figure 6a. The interface of the software part responsible for the process of detecting propaganda by the developed method is shown in Figure 6b.

With the introduction of the "suspicious post" category, the percentage of errors of the first and second kind decreased. When using the binary approach, 178 samples out of 3573 turned out to be incorrectly classified. However, out of 178 samples, only 71 samples are false, the remaining 107 were classified as "suspicious post". Out of 71 false samples, only 26 texts containing signs of political propaganda were falsely assigned to the "non-propaganda post" class, and 45 texts were falsely assigned to the "propaganda post" class. As for the results of the discrete approach, 130 samples turned out to be incorrectly classified, which in general did not worsen the statistics of the GRU neural network indicators, but out of 130 samples, 37 texts containing signs of political propaganda were wrongly assigned to the "post without propaganda" class, and 52 texts were wrongly assigned to the "propaganda post".

6. Results and Discussion

Accuracy, Precision, Recall and F1 metrics were used to study the effectiveness of detecting political propaganda in textual Internet content using the developed method [11]. The values of the metrics for the discrete and binary variations of the method are shown in Table 2. Although the binary approach gave worse results for the Accuracy metric, it gave better results for the Precision, Recall and F1 metrics, while the discrete Accuracy approach practically did not deteriorate, but at the same time the metrics Precision, Recall and F1 are somewhat inferior to it.

Table 2
Value of metrics for begging and staking

Approach	Accuracy	Precision	Recall	F1
Bagging	0.97	0.973	0.981	0.976
Stacking	0.95	0.977	0.987	0.981

For the experiment, the parameters of discrete approach were as follows: k_1=0.5, k_2=0.5, l_2 = 0.45, l_4 = 0.55. Although binary approach gave worse results on Accuracy metric, it gave better Precision, Recall, and F1 metrics, while the discrete Accuracy approach did not deteriorate, but the Precision, Recall, and F1 metrics were somewhat inferior to it. The chart of metric values is shown in Figure 7.
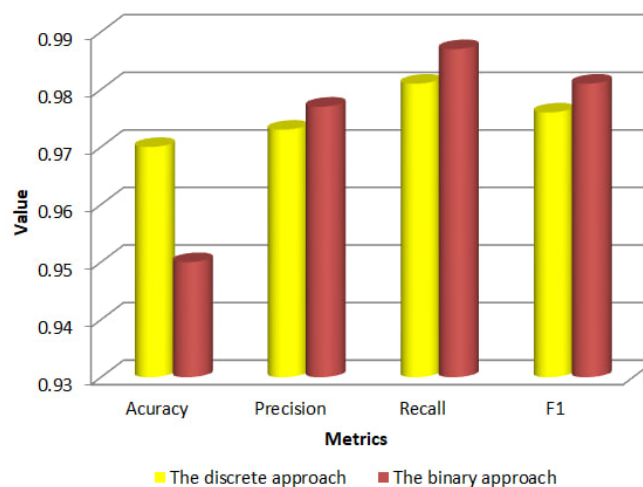

Figure 7: The value of metrics for binary and discrete approaches.

However, the advantages of the discrete method are its flexibility and the ability to be customized depending on the task. Experiments in this direction in the future are promising.

As for the texts that were identified as suspicious, specific signs were found in them. For example: "*let the l-t community calm down. They seem to be the same people (including nationalists), which means they can be joked about like everyone else. 2) Managers are satire, therefore fire. 3) Fortunately, Ukraine is not the USA, so you can joke about real graduates of the Ternopil Medka. P.S. The best joke where the musorina stops Best, takes a bribe. And then he comes back and takes it off in the heat of the moment. P.S.S. Where is the review of the sketch with Khmelnitsky and the Moscow ambassador? There's only one American subject worth anything!*" (in original Ukrainian: «*хай л-т спільнота успокояться . Вони ж начеб то такі самі люди (туди ж і націоналістів) , а значить над ними можна жартувати як і над усіма іншими ..2) Менеджери то сатира тому агонь .3) Україн на щастя не США , тому можна жартувати і над реальними випускниками тернопольської медки.П. С. Найкращий жарт де мусоріна зупиняє Беста , бере хабар . А той потім повертається і знімає його на гарячому.П. С. С. Де огляд скетчу з Хмельницьким та послом московським ? Там один фак американський чтого тільки вартий !*»). The text contains a number of trigger words, such as "nationalists" (in original Ukrainian: «націоналістів»), "Moscow" (in original Ukrainian: «московським»), and the context is similar to propaganda.

There were also erroneously assigned texts in the data set. For example, the following text in the dataset was marked as a "post without propaganda", but its content: "*Russia does not claim the territory of Ukraine, if Ukraine did not attack Russia, there would be no military action. Rather, Russia was attacked by NATO countries on the territory of Ukraine, because the authorities of Ukraine sold the country and betrayed their people and agreed to fight for the interests of Biden, Macron, Sunak and Scholtz until the last living Ukrainian.*" (in original Ukrainian: «*Росія не претендує на територію України, якби Україна не напала на Росію, жодних військових дій не було б. Вірніше на Росію напали країни НАТО на території України, тому що влада України продала країну і зрадила свій народ і погодилася воювати за інтереси Байдена, Макрона, Сунака та Шольца до останнього живого українця.*») is outright propaganda, and both neural network models rated it highly for political propaganda, giving scores of 0.92 (GRU) and 0.97 (BiLSTM), for a total score of 0.944.

7. Conclusions

A method for political propaganda detection in internet content using recurrent neural network models ensemble, which works with the Ukrainian language, has been proposed, and its approbation has been carried out. The method for detecting political propaganda in Internet content using ensemble of recurrent neural network models is intended to identify and analyze potentially propagandistic or manipulative content spread on the Internet. The input data of the method is an ensemble of trained models of recurrent neural networks with tokenizers and a text message for analysis. The output data are the level and percentage of propaganda presence for each neural network model of the ensemble and in general. As part of the research: preparation of an educational Ukrainian-language data set was completed; test training of an ensemble of classifiers based on BiLSTM and GRU neural network architectures was performed; software was developed that implements the created method for political propaganda detection in internet content using recurrent neural network models ensemble, and a study of its effectiveness was conducted.

The applied efficiency research of propaganda detection by an ensemble of classifiers based on the BiLSTM and GRU recurrent neural network architectures was conducted. The proposed approach is capable of detecting political propaganda by an ensemble of RNN models with Accuracy 0.97, Precision 0.973, Recall 0.981, and F1 0.976 in the bagging mode, and Accuracy 0.95, Precision 0.977, Recall 0.987, and F1 0.981 in the stacking mode . The developed method has a limitation: it works with text posts from 200 to 6300 symbols long. For shorter and longer texts, performance degradation is observed.

Further research will be aimed at analyzing the dependence of the considered performance indicators of the proposed method on the features and parameters of the analyzed post, such as size, genre, and subject matter. A promising direction for further research is also an increase in the

number of RNN models in the ensemble to improve performance indicators, and the specialization of models for certain types of propaganda.

References

[1] M. Last, Online Propaganda Detection, Data Mining and Knowledge Discovery Handbook, Cham: Springer International Publishing (2023) 703–719. doi:10.1007/978-3-031-24628-9_31

[2] D. G. Jones, Detecting Propaganda in News Articles Using Large Language Models, Eng. Open Access, 2 (2024) 1–12. doi:10.13140/RG.2.2.34115.17446

[3] P. N. Ahmad, K. Khan, Propaganda Detection And Challenges Managing Smart Cities Information On Social Media, EAI Endorsed Transactions on Smart Cities 7.2 (2023) e2–e2. doi:10.4108/eetsc.v7i2.2925

[4] D. Cavaliere, M. Gallo, C. Stanzione, Propaganda Detection Robustness Through Adversarial Attacks Driven by eXplainable AI, World Conference on Explainable Artificial Intelligence, Cham: Springer Nature Switzerland (2023) 405–419. doi:10.1007/978-3-031-44067-0_21

[5] J. A. Goldstein, J. Chao, Sh. Grossman, A. Stamos, M. Tomz, How persuasive is AI-generated propaganda, PNAS Nexus, 3. 2 (2024). doi:10.1093/pnasnexus/pgae034

[6] G. Faye, B. Icard, M. Casanova, J. Chanson, F. Maine, F. Bancilhon, G. Gadek, G. Gravier, P. Egre, Exposing Propaganda: an Analysis of Stylistic Cues Comparing Human Annotations and Machine Classification, Proceedings of the Third Workshop on Understanding Implicit and Underspecified Language, Association for Computational Linguistics, Malta (2024). 62–72.

[7] M. Abdullah, O. Altiti, R. Obiedat, Detecting Propaganda Techniques in English News Articles using Pre-trained Transformers, 13th International Conference on Information and Communication Systems (ICICS), Irbid, Jordan (2022) 301–308. doi:10.1109/ICICS55353.2022.9811117

[8] A. Horak, R. Sabol, O. Herman, V. Baisa, Recognition of Propaganda Techniques in Newspaper Texts: Fusion of Content and Style Analysis, Expert Systems with Applications, 251 (2024). doi:10.1016/j.eswa.2024.124085

[9] G. Martino, S. Yu, A. Barron-Cedeno, R. Petrov, P. Nakov, Fine-Grained Analysis of Propaganda in News Article, Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (2019) 5640–5650. doi:10.18653/v1/D19-1565

[10] D. B. Rodríguez, V. Dankers, P. Nakov, E. Shutova, Paper bullets: Modeling propaganda with the help of metaphor, Findings of the Association for Computational Linguistics: EACL (2023) 472–489. doi:10.18653/v1/2023.findings-eacl.35

[11] G. D. S. Martino, A. Barron-Cedeno, H. Wachsmuth, R. Petrov, P. Nakov, SemEval-2020 Task 11: Detection of Propaganda Techniques in News Articles, Proceedings of the Fourteenth Workshop on Semantic Evaluation, International Committee for Computational Linguistics, Barcelona (2020) 1377–1414.

[12] K. Hayawi, S. Shahriar, S. S. Mathew, The Imitation Game: Detecting Human and AI-Generated Texts in the Era of ChatGPT and BARD, Journal of Information Science (2024). doi:10.1177/01655515241227531

[13] O. Barmak, O. Mazurets, I. Krak, A. Kulias, A. Smolarz, L. Azarova, K. Gromaszek, S. Smailova, Information technology for creation of semantic structure of educational materials. Proceedings of SPIE – The International Society for Optical Engineering, 11176 (2019) 1117623. doi:10.1117/12.2537064

[14] I. Krak, O. Barmak, O. Mazurets, The Practice Investigation of the Information Technology Efficiency for Automated Definition of Terms in the Semantic Content of Educational Materials. CEUR Workshop Proceedings, 1631 (2016) 237–245. doi:10.15407/pp2016.02-03.237

[15] P. Chen, L. Zhao, Y. Piao, H. Ding, X. Cui, Multimodal Visual-Textual Object Graph Attention Network for Propaganda Detection in Memes, Multimedia Tools and Applications, 83.12 (2024) 36629–36644. doi:10.1007/s11042-023-15272-6

[16] D. Chaudhari, A. V. Pawar, Empowering Propaganda Detection in Resource-Restraint Languages: A Transformer-Based Framework for Classifying Hindi News Articles, Big Data and Cognitive Computing, 7.4, 175 (2023). doi:10.3390/bdcc7040175.

[17] A. Malak, D. Abujaber, A. Al-Qarqaz, R. Abbott, M. Hadzikadic, Combating Propaganda Texts Using Transfer Learning, IAES International Journal of Artificial Intelligence, 12 (2023) 956–965. doi:10.11591/ijai.v12.i2.pp956-965.

[18] O. Zalutska, M. Molchanova, O. Sobko, O. Mazurets, O. Pasichnyk, O. Barmak, I. Krak Method for Sentiment Analysis of Ukrainian-Language Reviews in E-Commerce Using RoBERTa Neural Network, CEUR Workshop Proceedings, 3387 (2023) 344–356. doi: 10.15407/jai2024.02.085

[19] A. Bhattacharjee, H. Liu, Fighting Fire with Fire: Can ChatGPT Detect AI-generated Text, SIGKDD Explor, Newsl 25 (2023) 14–21, doi:10.1145/3655103.3655106

[20] N. K. Kitson, A. C. Constantinou, Z. Guo, A Survey of Bayesian Network Structure Learning, Artif Intell Rev 56 (2023) 8721–8814. doi:10.1007/s10462-022-10351-w

[21] I. Krak, O. Zalutska, M. Molchanova, O. Mazurets, R. Bahrii, O. Sobko, O. Barmak, Abusive Speech Detection Method for Ukrainian Language Used Recurrent Neural Network, CEUR Workshop Proceedings, 3387 (2023) 16–28. doi:10.31891/2307-5732-2024-331-17

[22] S. Mann, D. Yadav, D. Rathee, Identification of Racial Propaganda in Tweets Using Sentimental Analysis Models: A Comparative Study, Proceedings of Fourth Doctoral Symposium on Computational Intelligence, Lecture Notes in Networks and Systems, 726 (2023) 327–341. doi: 0.1007/978-981-99-3716-5_28.

[23] L. Syed, A. Alsaeedi, L. A. Alhuri, H. R. Aljohani, Hybrid Weakly Supervised Learning with Deep Learning Technique for Detection of Fake News from Cyber Propaganda, Array, 19, 100309 (2023). doi:10.1016/j.array.2023.100309.

[24] X. Liu, K. Ma, K. Ji, Zh. Chen, B. Yang, Graph-Based Multi-Information Integration Network with External News Environment Perception for Propaganda Detection, International Journal of Web Information Systems, 20.2 (2024) 195–212. doi:10.1108/IJWIS-12-2023-0242

[25] I. Rizgelienė, G. Korvel, Comparative Analysis of Various Data Balancing Techniques for Propaganda Detection in Lithuanian News Articles, International Baltic Conference on Digital Business and Intelligent Systems, Cham: Springer Nature Switzerland (2024) 227–236. doi:10.1007/978-3-031-63543-4_15

[26] V. Solopova, O. Popescu, C. Benzmüller, Automated Multilingual Detection of Pro-Kremlin Propaganda in Newspapers and Telegram Posts, Datenbank Spektrum (2023) 5–14. doi:10.1007/s13222-023-00437-2.

[27] Y. Li, L. Guo, J. Wang, Y. Wang, D. Xu, J. Wen, An improved Sap Flow Prediction Model Based on CNN-GRU-BiLSTM and Factor Analysis of Historical Environmental Variables, Forests, 14.7 (2023) 1310. doi:10.3390/f14071310

[28] W. Yue, L. Li, Sentiment Analysis using a CNN-BiLSTM Deep Model Based on Attention Classification, International Information Institute, Tokyo, Information, 26.3 (2023) 117–162. doi:10.47880/inf2603-02

[29] A. R. Merryton, M. Gethsiyal Augasta, An Attribute-Wise Attention Model with BiLSTM for an Efficient Fake News Detection, Multimedia Tools and Applications, 83.13 (2024) 38109–38126. doi:10.1007/s11042-023-16824-6

[30] A. H. Alsaedi, A. A. K. Aladhamı, A. M. Alwhelat, A. L. Alshamı, Analysis and Detection of Political Fake News Using Deep Learning with High-Performance Hybrid Model, International Conference on Intelligence Science, Singapore, Springer Nature Singapore (2023) 261–271. doi:10.1007/978-981-99-8976-8_23