# Method for Detecting Propaganda Objects Using Deep Learning Neural Network Models with Visual Analytic

Iurii Krak[1,2], Olha Zalutska[3], Maryna Molchanova[3,*], Olexander Mazurets[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**

Approach to solving the problem of identifying propaganda objects was developed, which allows to find in propaganda texts who and what propaganda techniques are aimed at. Significant problems of propaganda detection were solved, namely: absence of a thorough examination of the connections between propaganda techniques and their targets within texts; lack of synthesized insights regarding propaganda targets and their alternative references in textual materials. Developed method for detecting propaganda objects allows to convert input data in form test text for detect propaganda objects, set of used techniques in the test text and set of neural network models trained to analyze every propaganda techniques, into output data in form semantic model of propaganda for test text. Semantic model of propaganda include set of propaganda objects, set of words representations objects with their semantic evaluation and set of important relations between propaganda techniques and propaganda objects in test text with semantic importance evaluation. Proposed method for detecting propaganda objects has demonstrated full alignment with the results obtained by expert evaluations.

**Keywords**

propaganda objects, propaganda techniques, detecting propaganda, visual analytics

## 1. Introduction

One of the most serious challenges facing humanity in the digital age is propaganda [1]. The main purpose of propaganda is the manipulation of objects in order to achieve certain political, social, economic or cultural goals [2, 3].

Objects of propaganda are understood as individuals, groups, organizations, social strata, as well as phenomena or institutions to which propaganda efforts are directed in order to influence their consciousness, emotions, behavior and public opinion [4].

The aim of the research is to develop a method for detecting propaganda objects using deep learning neural network models, which allows detecting specific objects targeted by specific propaganda techniques. The main contributions of the paper can be summarized as follows:

- Method for detecting propaganda objects is developed, which enables the identification of the specific individuals or entities targeted by propaganda techniques within the analyzed texts.

*Information Technology and Implementation (IT&I-2024), November 20-21, 2024, Kyiv, Ukraine*
* Corresponding author.
yuri.krak@gmail.com (I. Krak); zalutsk.olha@gmail.com (O. Zalutska); m.o.molchanova@gmail.com (M. Molchanova); exe.chong@gmail.com (O. Mazurets); alexander.barmak@gmail.com (O. Barmak)
000-0002-8043-0785 (I. Krak); 0000-0003-1242-3548 (O. Zalutska); 0000-0001-9810-936X (M. Molchanova); 0000-0002-8900-0650 (O. Mazurets); 0000-0003-0739-9678 (O. Barmak)

- The effectiveness of developed method application has been experimentally proven, which allows, in contrast to existing analogues, to expand the list of available propaganda objects at the expense of words representations objects, in addition to searching for NER.
- Visual interpretation of the identified propaganda targets was conducted, enabling clear observation of the subjects influenced within the framework of applied propaganda techniques.

The next section presents an overview of related works in the field of detecting propaganda and its objects. Section 4 offers an overview of the experiment to investigate the effectiveness of the proposed approach. Section 5 offers an overview of the obtained results with comparisons and discussions.

## 2. Related Works

The emergence of diverse media channels in the information age has marked a new chapter in communication and information accessibility, simultaneously amplifying opportunities for propaganda and the manipulation of public opinion [5]. Identifying propagandistic content requires not only recognizing the content but also determining its intended targets or objectives. Identification of objects of propaganda is critically important for understanding the mechanisms of manipulation, therefore it attracts the special attention of scientists.

Existing researches highlight the problem relevance of automated detection of propaganda in text web content. A significant number of studies are currently focused on the intellectualization of propaganda detection processes to address various technological challenges in monitoring media sources [6] and distinguishing propaganda techniques from other forms of manipulative influence [7, 8]. Propaganda encompasses subjects, content, forms, methods [9], and channels of information dissemination [10]. Its primary target is social groups or audiences subjected to influence.

Detecting propaganda in text using natural language processing (NLP) presents considerable challenges due to its reliance on subtle manipulation techniques and contextual dependencies. Researchers in [11] explored the effectiveness of modern large language models for identifying propaganda, conducting experiments on datasets containing news articles. Their findings indicate that modern large language models achieve performance levels comparable to the current state-of-the-art methods [12].

Effective approaches for semantic text content analysis, such as multimodal visual-text object graph attention networks [13] and statistical test analysis [14], have been noted as valuable tools for detecting propaganda [15]. Additionally, the use of transformer-based neural networks [16, 17] and complex neural architectures, including RoBERTa [18], GPT [19], and RNN [20], has emerged as a prominent direction in propaganda detection research.

Certain studies focus on specific components of propaganda, such as racial propaganda [21] or fake news [22]. The authors of [23] emphasize that most existing methods for detecting propaganda primarily analyze linguistic features of its content but often overlook the contextual information from the external media environment where propaganda originates and spreads.

Another problem with propaganda detection is the lack of reliably labeled data sources. To help the scientific community identify propaganda in text news, research [24] proposed a library of propaganda texts (ProText). Verity labels are assigned to ProText repositories after manual and automated verification using fact-checking techniques. The authors used a natural language processing approach to create a system that uses deep learning to automatically identify propaganda in the news. A fine-tuned robustly optimized BERT (RoBERTa) pre-training approach and word embedding using multi-label, multi-class text classification is proposed. Performance estimation accuracies of 90%, 75%, 68%, and 65% were achieved on ProText, PTC, TSHP-17, and

Qprop, respectively. The big data method, especially with deep learning models, can help to fill the unsatisfactory big data in the new text classification strategy.

There are two main approaches to detecting propaganda: a named object detection approach and a text classification approach.

When detecting propaganda as a named entity recognition (NER) task, the problem arises that fragments of text containing propaganda are longer than NER (for example, names of persons or places) and can include tens of words. In [25], the value of the range length that affects the recognition of propaganda is investigated, showing that the difficulty of the task does increase with the increase of the range length. Various commonly used approaches to the task are systematically assessed by evaluating their ability to accurately replicate the length distribution of actual intervals. Additionally, a novel solution is introduced, incorporating an adaptive convolution layer designed to enable information exchange between distant words. This approach enhances length preservation while maintaining overall performance.

Document-level approach to propaganda detection is presented in [26], which explores the potential of using large language models (LLMs), such as those underlying ChatGPT, to identify propaganda features in news articles. The methods discussed build upon the work of Martino et al., who defined a set of 18 distinct propaganda techniques. Using these techniques, an enhanced framework is developed that integrates news articles from Russia Today (RT), a well-known state-controlled media outlet, with the SemEval-2020 Task 11 dataset. The tip and article content are then fed to the gpt-3.5-turbo model OpenAI to determine what propaganda techniques are present and make a final decision on whether an article is propaganda or not. A subset of the results obtained is then analyzed to determine whether the LLM can be effectively used in this way. The authors report an accuracy of 25.12% using the SemEval-2022 dataset.

Thus, based on the results of the analysis of related works in the field of identifying propaganda techniques and propaganda objects, two problems were revealed: the lack of a comprehensive analysis of the relationships between propaganda techniques and objects in the texts; lack of synthesized insights regarding propaganda targets and their alternative references in textual materials. Propaganda detected by searching for named entities only has no visibility into the direction of propaganda through the use of techniques; however, the propaganda techniques detected at document level do not reflect the objects of propaganda direction [27]. At the same time, when identifying propaganda as named entity recognition (NER) task, the problem arises that propaganda objects are not presented by proper names or even synonyms, and their semantic proximity is not reflected in an obvious way. The paper proposes solution to these problems.

## 3. Method Design

It is believed that the propaganda techniques used in the text can be detected by one of the existing neural network approaches [11], while the secondary result will be appropriate neural network models trained to detect individual propaganda techniques. Thus, the proposed approach for identifying propaganda objects and comparing them with the applied techniques is illustrated in Figure 1. This approach provides a comprehensive analysis of the relationships between propaganda techniques and their objects within texts, as well as generalizations for the objects of propaganda and their alternative mentions in the texts. Its distinction from existing methods lies in combining two complementary tasks: the identification of propaganda objects and correlating the identified objects with numerical values assigned to the detected techniques.

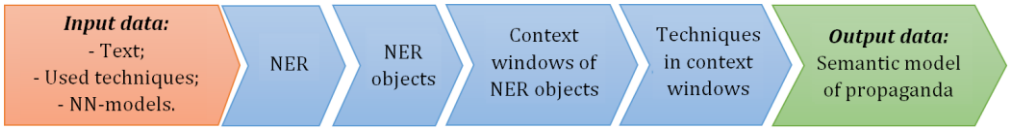


**Figure 1**: Proposed approach for identifying propaganda objects

Within the framework of the proposed approach, 17 separate pre-trained neural network models of the transformer architecture will be used, which allows determining 17 main propaganda techniques, such as: «Appeal to fear-prejudice», «Causal Oversimplification», «Doubt», «Exaggeration», «Flag-Waving», «Labeling», «Loaded Language», «Minimisation», «Name Calling», «Repetition», «Appeal to Authority», «Black and White Fallacy», «Reductio ad hitlerum», «Red Herring», «Slogans», «Thought terminating Cliches», «Whataboutism» [28].

The proposed neural networks are trained on labeled data collected by the Analysis Project team (https://propaganda.qcri.org/index.html), which analyzed the texts, detecting all segments that incorporate propaganda techniques along with identifying their specific types. Specifically, they developed a dataset of news articles that were manually annotated at the segment level based on eighteen distinct propaganda techniques. Dataset includes 788 articles.

The method of detecting propaganda objects created in accordance with the proposed approach uses neural network tools, and allows detecting the belonging of objects to the used propaganda techniques, and uses the following presentation of propaganda semantic model *SMP* in the test text *T*:

$$SMP = PO' \cup TT' \cup RTO' \cup Metadata, \quad (1)$$

where $PO'$ – set of words representations objects in the test text $T$ with their semantic evaluation ($PO' \subset PO$); $PO$ – set of propaganda objects in the test text $T$ detected by NER; $TT'$ – set of used techniques in the test text ($TT' \subset TT$); $TT$ – set of all possible techniques; $RTO'$ – set of important relations between techniques $TT'$ and objects $PO'$ in text $T$ with semantic importance evaluation; *Metadata* – auxiliary components complex.

Auxiliary components complex *Metadata* in (1) include components that are not an informative display of the propaganda model in the text, but are used to obtain initial data:

$$Metadata = T \cup NM' \cup CW \cup CW', \quad (2)$$

where $T$ – pre-processed test text; $NM'$ – set of neural network models trained to analyze every propaganda techniques; $CW$ – set of context windows which contain every emergence of words representations objects in the test text $T$; $CW'$ – set of united context windows associated with every words representations objects in the test text $T$ for neural network analyze.

Method for detecting propaganda objects using deep learning neural network models and includes stages: building set of propaganda objects $PO$ by named entity recognition; text pre-processing and expansion of $PO$ at expense of words representations objects; building context windows $CW$ and association of $CW$ by $PO'$; detecting level of propaganda techniques use in $CW'$ by neural network models $NM'$; building set of important relations $RTO'$ between techniques $TT'$ and objects $PO'$ for test text $T$ (Figure 2).

Method for detecting propaganda objects as input data has test text for detect propaganda objects $T$, set of used techniques in the test text $TT'$ and set of neural network models $NM'$, trained to analyze every propaganda techniques.

The first stage is named entity recognition (NER) using by «STANZA» neural network library. Since named entities can contain repetitions, all repetitions are removed at the level of lems at this stage as well. The initial data of first stage is the set of propaganda objects $PO$ by NER without repetitions.

The second stage will be to search for words-objects that are close in meaning to each named entity. This need arises because objects of propaganda are a slightly broader concept than NER contained in the set of $PO$. They also include aspects of culture, groups of objects united by certain characteristics, etc. A pre-trained «FastText» model developed by Facebook AI Research will be used to find NER-like objects. «FastText» supports the «CBOW» and «Skip-gram» models, which makes it possible to effectively analyze the context of words and identify semantic relationships between them [29]. The use of «FastText» in this context is appropriate, as the model allows the detection of similar words and objects based on contextual vectors, which is useful for expanding the range of detected propaganda objects beyond named entities. The «FastText» model is retrained before being used on propaganda texts. The initial data of the stage is expansion of $PO$ at

the expense of words representations objects *PO'*. The minimum semantic proximity is set depending on the task, determined empirically. No threshold was applied in this study.

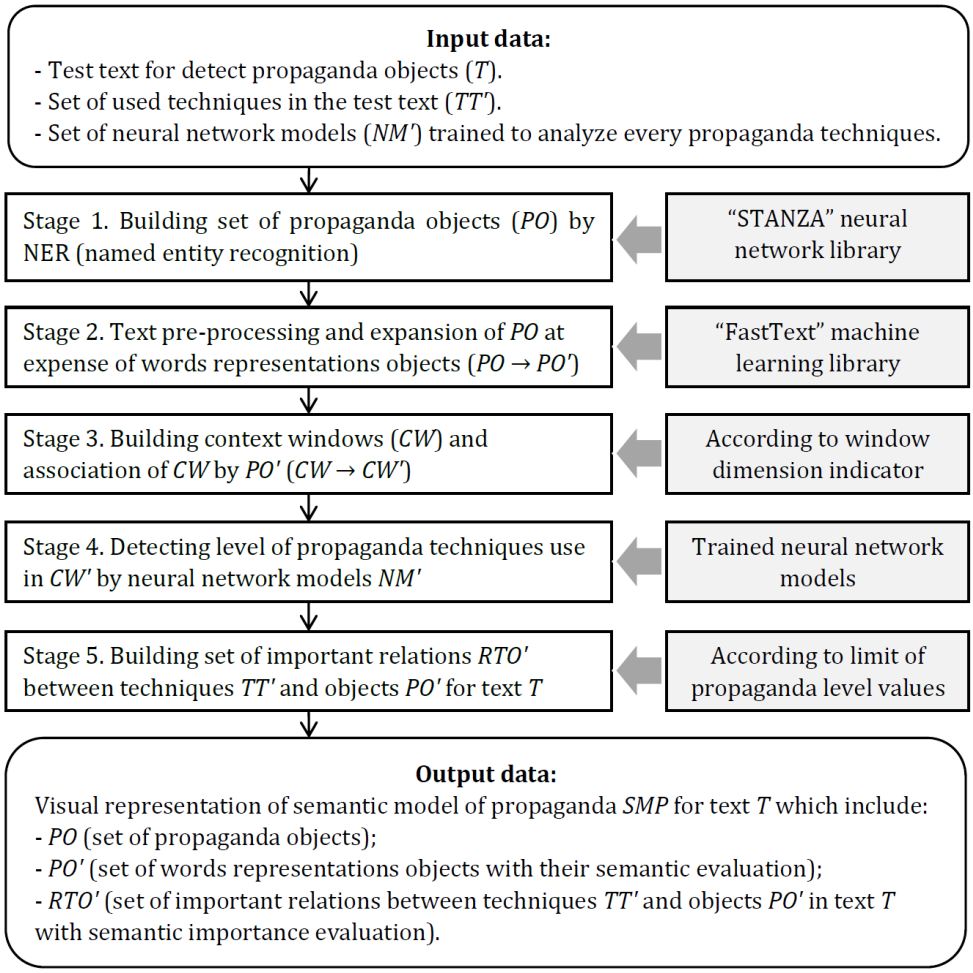


**Figure 2**: Scheme of method for detecting propaganda objects

The next step is the step of constructing *CW* context windows for each propaganda object with *PO´*. In the framework of the work, the context window is understood as the sentence where the specified object of propaganda is found [30]. If one context window contains several objects of propaganda, the windows are not duplicated (for one propaganda object).

According to *CW'*, detecting level of propaganda techniques use in *CW'* by neural network models *NM'* is carried out. The assessment of the context windows' belonging to the used techniques is carried out by vectorizing the context windows *CW'* with corresponding vectors of traffic jams *NM'* and analyzing its belonging to each of techniques presented in text *T*.

The last stage is the building set of important relations *RTO´* between techniques *TT´* and objects *PO´* for text *T*. If the strength of the manifestation of the technique is below a certain threshold value, within the framework of the evaluation of the element of the set *CW´*, the technique is not considered applied to the group of objects.

The initial data of the proposed approach are visual representation of semantic model of propaganda *SMP* for text *T* which include: set of propaganda objects; set of words representations objects with their semantic evaluation; set of important relations between techniques *TT'* and objects *PO'* in text *T* with semantic importance evaluation.

Therefore, an approach has been developed that allows detection of objects of propaganda that go beyond the detection of NER, and also allows to assess the affiliation of detected objects of propaganda to the propaganda techniques used in the text.

Accordingly, developed method for detecting propaganda objects allows to convert input data in form test text for detect propaganda objects ($T$), set of used techniques in the test text ($TT$) and set of neural network models ($NM$) trained to analyze every propaganda techniques into output data in form semantic model of propaganda $SMP$ for test text $T$. Semantic model of propaganda $SMP$ include set of propaganda objects $PO$, set of words representations objects with their semantic evaluation $PO'$ and set of important relations between techniques $TT'$ and objects $PO'$ in test text $T$ with semantic importance evaluation $RTO'$.

## 4. Experiment

To evaluate the effectiveness of the developed method for detecting propaganda objects, specialized software was created. This software facilitated the identification of propaganda objects and their comparison with the applied propaganda techniques. The obtained results in the form of visual analytics were compared with the results of the analysis of these sources by authoritative resources and experts in detecting and countering propaganda.

To conduct the experiment, software was created in the form of a web application in the Python programming language. Interface of the experimental software is shown in Figure 3.

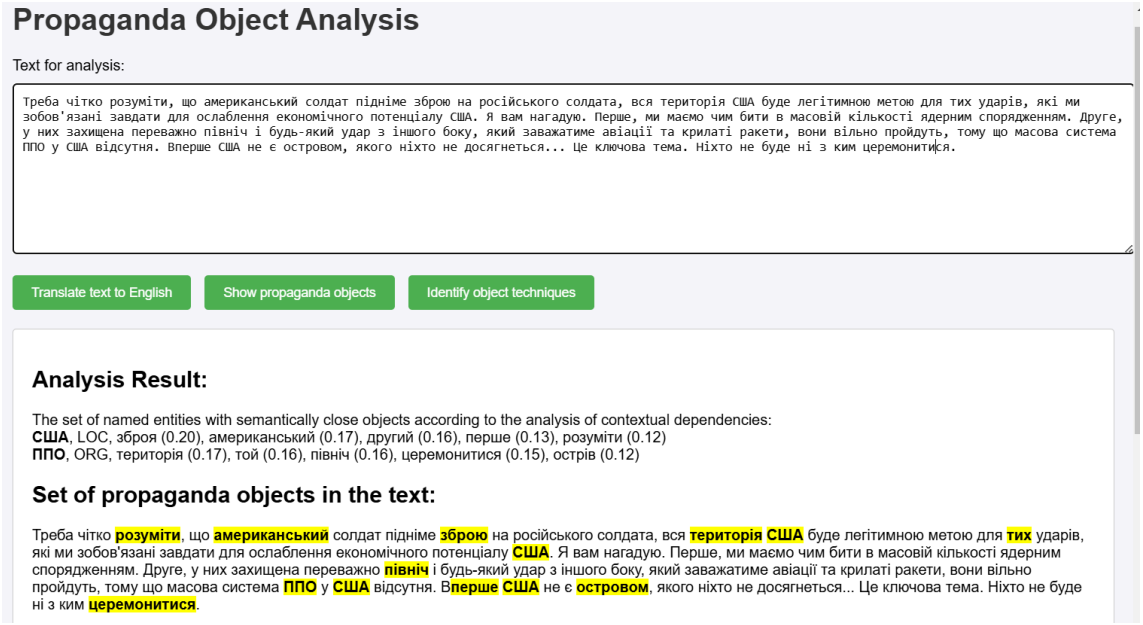


**Figure 3**: An example of obtaining the result with visual analytics by developed software

For the analysis, posts from social media were selected, processed by the «Strategic Communications Center» (https://spravdi.gov.ua/) and the «Disinformation Countering Center» (https://cpd.gov.ua/). These posts have been previously labeled and include expert conclusions, which can be compared with the original data produced by the proposed approach. Specialized software was developed for this analysis, incorporating a method for detecting propaganda objects.

The web application utilizes 17 pre-trained neural network transformation models (derived from previous research), the «Stanza» neural network library for named entity recognition (NER), the «Flask» framework, and the «FastText» pre-trained model, which has been retrained on propaganda texts. The software provides functionality for detecting the techniques used in text, identifying propaganda objects, and analyzing the association between detected objects and the applied techniques. Figure 4 shows the step-by-step illustration of proposed approach.

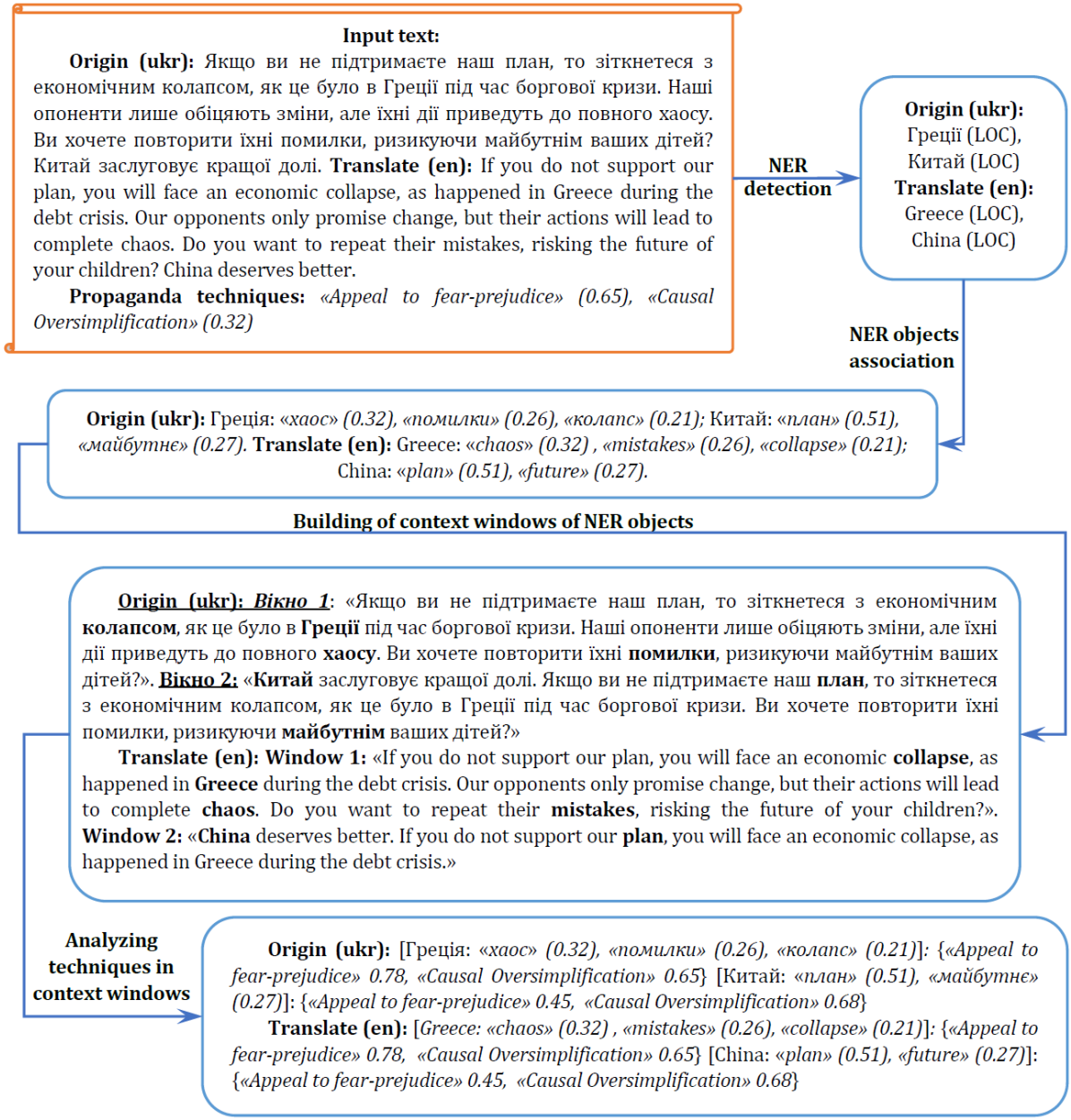


**Figure 4**: An example of propaganda objects data conversion by method

For example in Figure 4, within the text «*If you do not support our plan, you will face an economic collapse, as happened in Greece during the debt crisis. Our opponents only promise change, but their actions will lead to complete chaos. Do you want to repeat their mistakes, risking the future of your children? China deserves better.*» (in original Ukrainian: «*Якщо ви не підтримаєте наш план, то зіткнетеся з економічним колапсом, як це було в Греції під час боргової кризи. Наші опоненти лише обіцяють зміни, але їхні дії приведуть до повного хаосу. Ви хочете повторити їхні помилки, ризикуючи майбутнім ваших дітей? Китай заслуговує кращої долі.*») the first step is to search NER. Two named entities are found in this text: «*Greece* (*LOC*)» (in original Ukrainian: «*Греція*») and «*China* (*LOC*)» (in original Ukrainian: «*Китай*»). The set of NERs found must be further expanded by selecting semantically close objects. After passing the NER stage, the set of objects will be completed, and for this example it will become: {*Greece*: «*chaos*» (0.32), «*mistakes*» (0.26), «*collapse*» (0.21)}; {*China*: «*plan*» (0.51), «*future*» (0.27)} (in original Ukrainian: {*Греція*: «*хаос*» (0.32), «*помилки*» (0.26), «*колапс*» (0.21)}; {*Китай*: «*план*» (0.51), «*майбутнє*» (0.27)}).

The next stage is the construction of context windows for the found objects.

Window 1: {*Greece* (NER): «*chaos*» (0.32) , «*mistakes*» (0.26), «*collapse*» (0.21)}: «*If you do not support our plan, you will face an economic collapse, as happened in Greece during the debt crisis. Our opponents only promise change, but their actions will lead to complete chaos. Do you want to repeat their mistakes, risking the future of your children?*» (in original Ukrainian: {*Греція* (NER)*: «хаос» (0.32), «помилки» (0.26), «колапс» (0.21)*}: «*Якщо ви не підтримаєте наш план, то зіткнетеся з економічним колапсом, як це було в Греції під час боргової кризи. Наші опоненти лише обіцяють зміни, але їхні дії приведуть до повного хаосу. Ви хочете повторити їхні помилки, ризикуючи майбутнім ваших дітей?*»).

Window 2: {*China* (NER): «*plan*» (0.51), «*future*» (0.27)}: *«China deserves better. If you do not support our plan, you will face an economic collapse, as happened in Greece during the debt crisis»* (in original Ukrainian: {*Китай* (NER)*: «план» (0.51), «майбутнє» (0.27)*}: «*Китай заслуговує кращої долі. Якщо ви не підтримаєте наш план, то зіткнетеся з економічним колапсом, як це було в Греції під час боргової кризи. Ви хочете повторити їхні помилки, ризикуючи майбутнім ваших дітей?*»).

The last step is to search for techniques in context windows. The assessment of the presence of the technique is carried out by evaluating the context window with a neural network trained to identify the techniques expressed in the text. For this example: {*Greece: «chaos» (0.32), «mistakes» (0.26), «collapse» (0.21)*} (in original Ukrainian: {*Греція: «хаос» (0.32), «помилки» (0.26), «колапс» (0.21)*}) «Appeal to fear-prejudice» technique was rated at 0.78, « Causal Oversimplification» technique was rated at 0.65. At the same time, for {*China: «plan» (0.51), «future» (0.27)*} (in original Ukrainian: {*Китай: «план» (0.51), «майбутнє» (0.27)*), «Appeal to fear-prejudice» was rated at 0.51, «Causal Oversimplification» was rated at 0.68.

Considering the content of the context windows, Window 1 has more pronounced manifestations of both techniques (especially «Appeal to fear-prejudice»). Window 2 has a less pronounced «Appeal to fear-prejudice» technique and a more pronounced «Causal Oversimplification» technique due to the simplified interpretation of causes and effects.

## 5. Results and discussion

When training neural network models of the transformer architecture based on pre-trained BERT models, an accuracy of more than 80% was achieved for detecting propaganda techniques [18], the accuracy results of the BERT models are shown in Table 1.

Regarding the study of the effectiveness of the proposed method for identifying propaganda techniques and the objects they are aimed at, here the results correlate with the conclusions of researchers from the «Strategic Communications Center» (https://spravdi.gov.ua/) and «Disinformation Countering Center» (https://cpd.gov.ua/). For example, a post from the propaganda channel and the expert's conclusion are shown on Figure 5.

As can be seen from Figure 5, the post is aimed at:

- justifying the use of nuclear weapons against the United States;
- further Western intervention threatens to escalate the war;
- portraying the West as a weak opponent (discrediting).

As part of the analysis obtained by using the developed software installation, the following data were obtained:

- used propaganda techniques with manifestation powers: «Appeal to fear-prejudice» with rating of 0.598, «Loaded Language» with rating of 0.46 and «Exaggeration» with rating of 0.356;

- propaganda objects: NER together with words semantically close to them with closeness ratings, software output data: «*США* [*зброя, американський, перше, розуміти*]; *ППО* [*територія, той, північ, церемонитися, острів*]»;
- assessment of important relations of propaganda objects to the techniques used, software output data: «*США* [*Appeal to fear-prejudice 0.578; Loaded Language 0.473; Exaggeration 0.311*], *ППО* [*Appeal to fear-prejudice 0.611; Loaded Language 0.4; Exaggeration 0.379*]»;
- visual presentation of found objects in custom text.

**Table 1**
Accuracy of propaganda techniques detection by neural network models

| Techniques of propaganda | Accuracy |
|---|---|
| «Appeal to fear-prejudice» | 0.87 |
| «Causal Oversimplification» | 0.82 |
| «Doubt» | 0.93 |
| «Exaggeration» | 0.80 |
| «Flag-Waving» | 0.92 |
| «Labeling» | 0.96 |
| «Loaded Language» | 0.97 |
| «Minimisation» | 0.90 |
| «Name Calling» | 0.92 |
| «Repetition» | 0.94 |
| «Appeal to Authority» | 0.89 |
| «Black and White Fallacy» | 0.91 |
| «Reductio ad hitlerum» | 0.87 |
| «Red Herring» | 0.80 |
| «Slogans» | 0.86 |
| «Thought terminating Cliches» | 0.80 |
| «Whataboutism» | 0.79 |


**Figure 5**: Analysis of post which contains propaganda, from «Disinformation Countering Center» (https://cpd.gov.ua/)

The use of the «Appeal to Fear-Prejudice» technique in the text is followed by the intention to create fear about the potential use of nuclear weapons, intimidation with massive strikes and the description of the USA as a vulnerable target creates a feeling of anxiety and uncertainty in the reader.

The use of the «Loaded Language» technique is followed by expressions such as «вся територія США буде легітимною метою» and «ми маємо чим бити» deliberately selected to evoke strong emotions such as anger or fear.

The «Exaggeration» technique is expressed by exaggerating the weakness of the US defense systems («масова система ППО у США відсутня») and the ability to deliver an irresistible blow, which is an attempt to give the reader the impression of aggressor's absolute superiority. An example of the analysis performed by the program is shown in Figure 6.

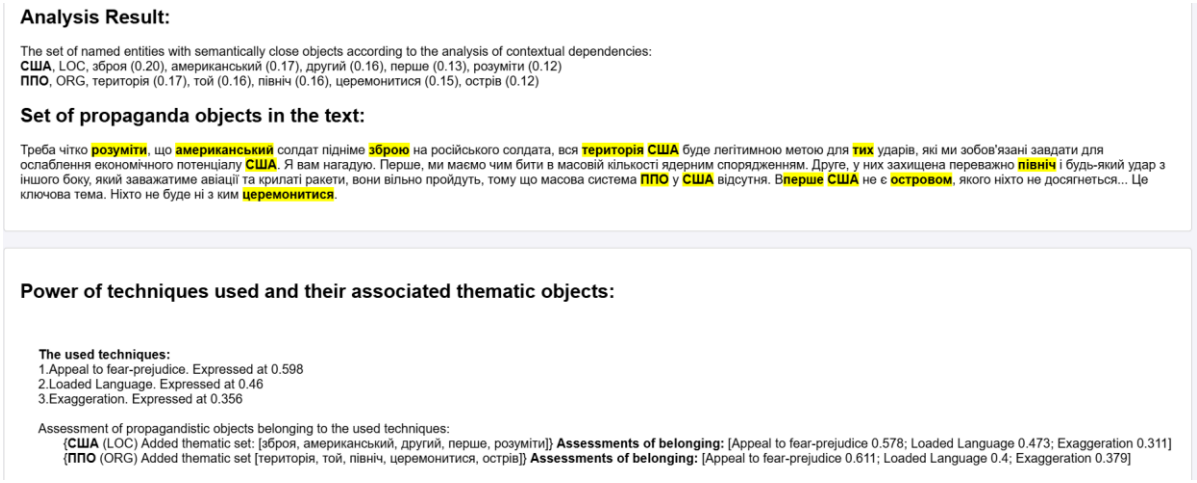


**Figure 6**: Analysis of test text with propaganda by developed software

As shown in Figure 6, the analysis not only identified propaganda objects and the techniques used but also determined how the identified objects were linked to the techniques applied.

Thus, the study evaluating the effectiveness of the developed method revealed that the technique for detecting propaganda objects produces results that align perfectly with the expert findings. Additionally, by treating propaganda as a unified model and employing visual analytics for the results, it was possible to conduct a thorough analysis of the relationships between propaganda techniques and objects, as well as to generalize the objects of propaganda and their alternative references in the texts.

## 6. Conclusions

As part of the main goal, approach to solving the problem of identifying propaganda objects was developed, which allows to find in propaganda texts who and what propaganda techniques are aimed at. Significant problems of propaganda detection were solved, namely: absence of a thorough examination of the connections between propaganda techniques and their targets within texts; lack of synthesized insights regarding propaganda targets and their alternative references in textual materials. The effectiveness of the proposed approach has been experimentally demonstrated, offering a significant advantage over existing methods. In addition to performing named entity recognition (NER) with the «STANZA» neural network library, it expands the list of identified propaganda objects in texts by utilizing the «FastText» machine learning library. Furthermore, the approach provides an assessment that correlates the detected objects with the applied techniques. For user convenience, a visual representation of the identified propaganda objects is also provided,

enabling a clear visualization of the influence targets within the framework of the used propaganda techniques.

Developed method for detecting propaganda objects use deep learning neural network models and includes stages: building set of propaganda objects by named entity recognition; text pre-processing and expansion of propaganda objects at expense of words representations objects; building context windows and they association by every words representations objects as context multiwindows; detecting level of propaganda techniques use in context multiwindows by neural network models; building set of important relations between propaganda techniques and propaganda objects for test text. Accordingly, developed method for detecting propaganda objects allows to convert input data in form test text for detect propaganda objects, set of used techniques in the test text and set of neural network models trained to analyze every propaganda techniques, into output data in form semantic model of propaganda for test text. Semantic model of propaganda include set of propaganda objects, set of words representations objects with their semantic evaluation and set of important relations between propaganda techniques and propaganda objects in test text with semantic importance evaluation.

To evaluate the effectiveness of the developed method for detecting propaganda objects, specialized software was created. This software facilitated the identification of propaganda objects and their comparison with the techniques employed in the content. The obtained results in the form of visual analytics were compared with the results of the analysis of these sources by authoritative resources and experts in detecting and countering propaganda. As a result, proposed method for detecting propaganda objects has demonstrated full alignment with the results obtained by expert evaluations. Further research is planned to be directed to the automation of the process of comparing the obtained results with the results obtained by experts. Also, further research will be aimed at developing a method for automated dynamic determination of minimum threshold values for detecting propaganda techniques occurrences, which will allow identifying the key features of combination of objects and propaganda techniques used in test text, while not overloading the model by noise in the form of insignificant semantic manifestations of the author's sentiment.

## Declaration on Generative AI

Authors have not employed any Generative AI tools.

## References


[1] P. 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

[2] 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

[3] 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

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

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

[6] J. 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

[7] 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 (2024). 62-72.
[8] 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) (2022) 301-308. doi:10.1109/ICICS55353.2022.9811117
[9] 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 (2024). doi:10.1016/j.eswa.2024.124085
[10] 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 (2019) 5640-5650. doi:10.18653/v1/D19-1565
[11] G. 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 (2020) 1377-1414.
[12] K. Hayawi, S. Shahriar, 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] 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
[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] 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
[16] D. Chaudhari, A. Pawar, Empowering Propaganda Detection in Resource-Restraint Languages: A Transformer-Based Framework for Classifying Hindi News Articles, Big Data and Cognitive Computing, 7.4 (2023) 175. 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] N. Kitson, A. Constantinou, Z. Guo, A Survey of Bayesian Network Structure Learning, Artificial Intelligence Revue 56 (2023) 8721-8814. doi:10.1007/s10462-022-10351-w
[20] 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, 3688 (2024) 16-28. doi:10.31891/2307-5732-2024-331-17
[21] S. Mann, D. Yadav, D. Rathee, Identification of Racial Propaganda in Tweets Using Sentimental Analysis Models: A Comparative Study, Lecture Notes in Networks and Systems, 726 (2023) 327-341. doi: 0.1007/978-981-99-3716-5_28.
[22] L. Syed, A. Alsaeedi, L. Alhuri, H. Aljohani, Hybrid Weakly Supervised Learning with Deep Learning Technique for Detection of Fake News from Cyber Propaganda, Array 19 (2023) 100309. doi:10.1016/j.array.2023.100309.

[23] 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
[24] P. Ahmad, Y. Liu, G. Ali, M. Wani, M. ElAffendi, Robust Benchmark for Propagandist Text Detection and Mining High-Quality Data, Mathematics, 11 (2023) 2668. doi:10.3390/math11122668.
[25] P. Przybyla, K. Kaczynski, Long Named Entity Recognition for Propaganda Detection and Beyond, Proceedings of the International Conference of the Spanish Society for Natural Language Processing (2023). doi:10.1007/s44196-024-00456-1
[26] D. Jones, Detecting propaganda in news articles using large language models, Engineering, 2, (2024) 1-12. doi:10.13140/RG.2.2.34115.17446.
[27] I. Krak, V. Didur, M. Molchanova, O. Mazurets, O. Sobko, O. Zalutska, O. Barmak, Method for Political Propaganda Detection in Internet Content Using Recurrent Neural Network Models Ensemble, CEUR Workshop Proceedings, 3806 (2024) 312-324.
[28] S. Yoosuf, Y. Yang, Fine-Grained Propaganda Detection with Fine-Tuned BERT, Proceedings of the Second Workshop on Natural Language Processing for Internet Freedom: Censorship, Disinformation, and Propaganda (2019) 87-91. doi:10.18653/v1/D19-5011
[29] I. Ghozali, K. Sungkono, R. Sarno, R. Abdullah, Synonym based feature expansion for Indonesian hate speech detection, International Journal of Electrical and Computer Engineering, 13. 1 (2023) 1105-1112. doi:10.11591/ijece.v13i1
[30] L. Rongsheng, J. Xu, Z. Cao, H. Zheng, H. Kim, Extending Context Window in Large Language Models with Segmented Base Adjustment for Rotary Position Embeddings, Applied Sciences, 14. 7 (2024) 3076. doi: 10.3390/app14073076