ISSN 1814-4225 (print)
ISSN 2663-2012 (online)

UDC 004.8:327.56 doi: 10.32620/reks.2024.4.04

Halyna PADALKO[1,2,3], Vasyl CHOMKO[2], Sergiy YAKOVLEV[4,5], Plinio P. MORITA[2,6,7]

[1] National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine
[2] University of Waterloo, Waterloo, ON, Canada
[3] Balsillie School of International Affairs, Waterloo, ON, Canada
[4] Lodz University of Technology, Lodz, Poland
[5] V. N. Karazin Kharkiv National University, Kharkiv, Ukraine
[6] University of Toronto, Toronto, ON, Canada;
[7] University Health Network, Toronto, ON, Canada;

CLASSIFICATION OF DISINFORMATION IN HYBRID WARFARE: AN APPLICATION OF XLNET DURING THE RUSSIA'S WAR AGAINST UKRAINE

The spread of disinformation has become a critical component of hybrid warfare, particularly in Russia's war against Ukraine, where social media serves as a battlefield for influence and propaganda. This study ***develops*** *a comprehensive methodology for classifying disinformation in the context of hybrid warfare, focusing on Russia's war against Ukraine. The* ***objective*** *of this study is to address the challenges of disinformation detection, particularly the increased spread of propaganda due to hybrid warfare. The* ***study focuses on*** *the use of transformer-based language models, specifically, XLNet, to classify multilingual, context-sensitive disinformation. The* ***tasks*** *of this study are to analyze current research and develop a methodology to effectively classify disinformation using the XLNet model. The proposed* ***methodology*** *includes several key components: data preprocessing to ensure quality, application of XLNet for training on diverse datasets, and hyperparameter optimization to handle the complexities of disinformation data. The study used datasets containing pro-Russian and neutral/pro-Ukrainian tweets, and the XLNet model demonstrated strong performance metrics, including high precision, recall, and F1-scores across different dataset sizes.* ***Results*** *showed that accuracy initially improved with increasing data volume but declined slightly with numerous datasets, suggesting the need for balancing data quality and quantity. The proposed methodology addresses the gaps in automated disinformation detection by integrating transformer-based models with advanced preprocessing and training techniques. This research improves the capacity for real-time detection and analysis of disinformation, thus contributing to public information governance and strategic communication efforts during wartime.*

Keywords: *hybrid warfare; disinformation detection; machine learning; XLNet; social media analysis; transformer models.*

1. Introduction

In recent years, the global landscape has experienced a surge in hybrid warfare, marked by Russia's war against and an increasing reliance on disinformation campaigns that leverage advanced technologies, including artificial intelligence (AI) [1]. Russia's war against Ukraine exemplifies the multifaceted nature of modern conflicts, where cyber warfare and information manipulation serve as powerful tools of influence and disruption [2]. The accessibility of AI technologies has enabled state and non-state actors to craft and disseminate disinformation at an unprecedented scale, affecting Ukraine and other democracies worldwide [3]. This phenomenon underscores the evolution of warfare from conventional battlefields to digital platforms, with disinformation campaigns eroding the foundational concepts of truth and democratic governance.

The year 2024 has proven to be a pivotal period for testing the resilience of democratic societies [4]. Nearly half of the global population participated in significant elections marred by foreign information manipulation and interference efforts aimed at shaping public opinion and influencing election outcomes [5]. Disinformation campaigns targeting electoral processes have intensified, posing substantial threats to democratic integrity by distorting the informational landscape available to citizens [6]. These campaigns, such as Doppelganger, are often coordinated across multiple platforms and exploit existing political and social divides, fueling polarization and mistrust among the electorate [7]. Consequently, the

Creative Commons Attribution
NonCommercial 4.0 International

spread of manipulated content can undermine the foundation of informed voting, thus diminishing citizens' ability to make informed decisions [8].

The sensitive intersection of information governance and freedom of speech remains a central challenge for democracies in effectively combating disinformation. While these issues are crucial for maintaining democratic integrity, governments worldwide have struggled to implement coherent policy frameworks to address them. The need for robust regulatory responses reflects the complexity involved in balancing free speech with the need for truthful information [9]. In response to this policy gap, technological interventions have emerged as promising solutions, with AI and machine learning (ML) tools leading the charge [10]. AI-driven approaches hold significant potential in identifying, analyzing, and mitigating the risks associated with foreign information manipulation [11]. By leveraging advanced algorithms, these tools can detect patterns of coordinated inauthentic behavior, predict content that may rapidly gain traction, and conduct nuanced content analysis to flag harmful or misleading narratives.

ML, in particular, offers robust methodologies for data classification and cluster analysis, making it an invaluable tool for fighting disinformation [12]. Classification algorithms can accurately categorize content based on its likelihood of being manipulative, whereas clustering techniques can identify thematic narratives that could signal coordinated campaigns [13]. The ability of ML to automate and enhance these processes provides policymakers with evidence-based insights necessary for informed decision-making [14] . For instance, governments can employ ML frameworks to identify foreign interference campaigns in real-time, enabling timely interventions to mitigate their impact.

The continued advancement of AI and ML technologies presents an opportunity for democratic societies to enhance their resilience against disinformation and foreign interference [15]. As these technologies evolve, so does their capacity to support public policy and governance efforts [16]. By equipping governments with tools to detect, analyze, and respond to information manipulation, AI is a critical asset in preserving democratic values in an increasingly digitized world. Through rigorous academic and policy-oriented research, the potential of AI-driven solutions to combat disinformation can be fully realized, paving the way for more resilient and secure democracies.

This study aimed to develop a deep learning model based on the XLNet architecture to efficiently analyze war-related content on social media and classify information into pro-Russian, pro-Ukrainian, and neutral narratives. In this study, we contribute to the growing field of automated disinformation detection and narrative analysis in the context of hybrid warfare. By leveraging XLNet, a transformer-based model known for its superior performance in language-understanding tasks, we address the complexities of multilingual, context-sensitive, and often subtle messaging that characterize propaganda and influence campaigns during Russian war against Ukraine.

In this paper, section 2, namely, Current Research Analysis, discusses the current state of research on disinformation, propaganda, and informational disorders in the context of Russia's war against Ukraine. Section 3, Methodology, presents the XLNet model for disinformation classification. Section 4, namely Results discuss data collection, model tuning, and model performance. Section 5, namely, the Discussion, discusses the proposed methodology and highlights its novelty, applicability, and limitations. The conclusions describe the outcomes of the research.

2. Current Research Analysis

The reviewed studies provide a comprehensive analysis of various aspects of disinformation and information manipulation related to Russia's war against Ukraine. These investigations cover diverse strategies and tools used by both state-affiliated entities and individuals to influence narratives, including agenda-setting, framing, propaganda through bots, and the involvement of diaspora communities. Several ML models, such as Graph Neural Networks (GNN), ensemble text classifiers, ARIMA models, and large language models, have been employed to examine manipulation tactics, sentiment, community interactions, and disinformation detection on platforms like Twitter, Reddit, and Telegram. Despite using advanced models, the studies face limitations like dataset biases, challenges in generalizability, high computational requirements, and a focus on specific platforms, which restricts the broad applicability of their findings.

The study [17] explored nuanced strategies of information manipulation during Russian war against Ukraine, focusing on agenda-setting, framing, and priming tactics. The authors released the VoynaSlov dataset comprising over 38 million social media posts from Russian media outlets, and they used various NLP models, including the Structured Topic Model (STM), Contextualized Neural Topic Model (CTM), and large pre-trained language models like XLM-RL, to examine these strategies across media types and time periods. Their findings revealed significant differences in manipulation tactics based on media control, platforms and wartime contexts. However, this study is limited by biases in the data collection process and the challenges associated with effectively deploying NLP models during an emerging crisis.

The article [18] explores the use of GNNs to detect fake news, specifically in the context of disinformation campaigns, such as those observed in Russia's information warfare against Ukraine. This study focuses on automating the analysis of negative psychological influences in online media through knowledge graphs (KG) and GNN-based models, including GraphSAGE, GAT, and GCN. By encoding relationships within knowledge graphs, these methods help identify harmful content spread across social media. Despite promising results, the study highlights several limitations: the dependency on large, labeled datasets, challenges with model stability and accuracy across different platforms, and the requirement for substantial computational resources, especially for real-time monitoring.

The article [19] investigated how users on social media platforms, particularly on Reddit's /r/Russia subreddit, act as "visual audience gatekeepers" by selectively sharing images to influence public perception, especially during Russia's war against Ukraine. Through critical visual content analysis, the study explores how gatekeepers create a visual echo chamber that reinforces their social reality and ideological perspectives. The main findings reveal that during polarizing events, users amplify specific narratives, often showcasing Russia favorably while condemning perceived adversaries, leading to radicalized and biased visual content. A limitation of this study is its focus on a single subreddit, which may not represent broader audience dynamics on other platforms or contexts.

The article [20] examines the topics and sentiments expressed by Ukrainian-speaking Telegram users during the first six months of the Russian war against Ukraine using machine learning techniques to analyze social media data. This study implements topic modeling through Non-negative Matrix Factorization with Kullback-Leibler Divergence and sentiment analysis using pretrained models to categorize themes and emotional tones of messages. A notable limitation is the dataset's focus on a single platform (Telegram), which may not capture the full range of social discourses surrounding war.

The article [21] analyzes Russia's bot-driven propaganda and Ukraine's counter-narratives on social media during the key stages of the 2022 Russia's war against Ukraine. Using TweetBERT for topic modeling and integrating the BEND framework with Moral Foundations Theory, this study examines how bots manipulated narratives to justify Russian actions and counter NATO, while Ukraine used similar tactics to promote solidarity and resilience. The primary limitation of this study is its exclusive focus on bot-generated content, which may not fully capture human interactions or the overall influence on public opinion.

The article [22] investigated the impact of Twitter's labeling policy on Russian state-affiliated media accounts, examining whether the label reduced these accounts' reach and influence following the onset of Russia's war against Ukraine. Using an ARIMA model to track engagement metrics before and after Twitter implemented labeling on February 28, this study measures changes in tweet reach, focusing on retweet counts. A major limitation is the lack of a causal link between Twitter's labeling policy and engagement decline due to concurrent events, such as the restriction of Russian media in Europe and Russia's blocking of Twitter.

The article [23] introduces the OLTW-TEC method, which is andvanced machine learning approach for detecting disinformation in Ukrainian-language content by utilizing an ensemble of text classifiers and a sliding window for dynamic online learning. The proposed method combines multiple classifiers to adapt to changes in the data, ensuring high accuracy and relevance in real-time scenarios. A notable limitation of the proposed method is its high computational resource demand, which may hinder scalability, especially for large-scale or resource-constrained applications.

The article [24] investigates the handling of Russian disinformation about war using three popular LLM-powered chatbots: Perplexity, Google Bard, and Bing Chat. Using an AI audit approach, this study examines the consistency, accuracy, and use of disclaimers in chatbot responses to prompts tied to Russian disinformation narratives. A major limitation of this study is the inherent stochasticity in LLM outputs, which leads to significant variability in responses, often resulting in the unintended amplification of false narratives.

The article [25] presents an Entity-Aware Approach (EAA) to detect logical fallacies in Kremlin-related social media content, specifically targeting disinformation about the Ukraine war. Using named entity recognition (NER), the EAA replaces named entities with general labels to improve model performance by reducing confusion in fallacy detection, particularly when applied to Kremlin tweets. The results demonstrate that, combined with the DeBERTa language model, EAA outperforms baseline models on both the domain-nonspecific dataset (LOGIC) and domain-specific datasets (RuFal). However, the current study is limited by its reliance on a single NER approach and dataset. Thus, future work should consider additional datasets and ensemble methods.

The article [26] investigated the role of Ukrainian and Russian diaspora communities in spreading disinformation on social media, specifically focusing on content related to the Donbas conflict and the MH17 crash. This study uses a combination of social network analysis and ML classification techniques to identify user communities and classify them by diaspora affiliation (Ukrainian, Russian, or other). A significant limitation of the research is the lack of multilingual data, as the study is limited to English-language tweets, which may not capture the full scope of diaspora engagement in disinformation campaigns.

Table 1 presents an overview of existing disinformation analysis studies conducted during Russia's war in Ukraine.

Table 1

An overview of studies on disinformation analysis during Russian war in Ukraine

Paper	Task	Method	Findings
Challenges and Opportunities in Information Manipulation Detection: An Examination of Wartime Russian Media [17]	Assess strategies of information manipulation used during the war, specifically focusing on tactics such as agenda-setting, framing, and priming.	NLP models: Structured Topic Model (STM), Contextualized Neural Topic Model (CTM), and pre-trained models like XLM-RL	Significant differences in the manipulation tactics used, depending on the type of media control (state-affiliated or independent), the platform (Twitter or VKontakte), and the context (pre-war vs. wartime).
An analysis of approach to the fake news assessment based on the graph neural networks [18]	Detect fake news in online media by identifying negative psychological influences within disinformation campaigns.	GNNs like GraphSAGE, GAT, and GCN, combined with knowledge graphs to model relationships and analyze textual data.	GNN models, especially GraphSAGE, effectively classify content with high accuracy, with GraphSAGE achieving the best performance among tested models.
Visual audience gatekeeping on social mediaplatforms: A critical investigation on visualinformation diffusion before and during the Russo–Ukrainian War [19]	To explore visual audience gatekeeping on social media during polarizing events.	Critical visual content analysis on images shared on Reddit's /r/Russia subreddit during the war.	Users create a visual echo chamber, amplifying pro-Russian perspectives and demonizing adversaries, and that heightened social tensions intensify these biases and lead to more radical visual narratives.
First Six Months of War from Ukrainian Topic and Sentiment Analysis [20]	To capture and analyze the topics and sentiments of Ukrainian Telegram users during the initial phase of the war.	Non-negative Matrix Factorization with Kullback-Leibler Divergence and sentiment analysis	Key topics include armed conflict, political figures, and humanitarian issues, with sentiment largely negative but showing some positivity around significant events; this analysis highlights real-time social perceptions but may be limited by platform-specific discourse.
Analyzing digital propaganda and conflict rhetoric: a study on Russia's bot-driven campaigns and counter-narratives during the Ukraine crisis [21]	To examine bot-driven propaganda and counter-narratives on social media during the Russian war against Ukraine	TweetBERT for topic modeling, coupled with the BEND framework and Moral Foundations Theory.	Pro-Russian bots spread narratives focusing on loyalty and protection themes, while pro-Ukraine bots emphasize justice and resistance. However, the study suggests that bot-generated propaganda might be less effective than anticipated due to limited influence on actual public opinion
The fight against disinformation and its consequences: measuring the impact of "Russia state-affiliated media" on Twitter [22]	To analyze the impact of Twitter's labeling policy on the reach of Russian state-affiliated media accounts.	ARIMA model, Structured Topic Modelling (STM),	The findings indicate a decline in reach for tagged accounts, especially for journalists, but the simultaneous implementation of external restrictions complicates attribution to Twitter's policy alone.

Continuation of Table 1

Paper	Task	Method	Findings
OLTW-TEC: online learning with sliding windows for text classifier ensembles [23]	To develop a robust method for identifying disinformation in Ukrainian-language online content.	Ensemble of classifiers (logistic regression, svm, lstm, transformers) with online learning and a sliding window technique.	OLTW-TEC achieved a classification accuracy of 93%, showing high precision (0.95) for fake news and excellent recall (0.99) for true news, but scalability issues due to computational demands remain a concern.
Stochastic lies: How LLM-powered chatbots deal with Russian disinformation about the war in Ukraine [24]	To analyze how LLM-powered chatbots respond to prompts involving Russian disinformation about the Ukraine war.	An AI audit assessing the consistency, accuracy, and use of disclaimers in responses.	Findings reveal that more than a quarter of chatbot outputs contain inaccuracies, and less than half of responses mention or debunk Russian perspectives, with high variability in consistency across chatbot instances
An Entity-Aware Approach to Logical Fallacy Detection in Kremlin Social Media Content [25]	To detect logical fallacies in Kremlin social media content using an Entity-Aware Approach	DeBERTa model, using named entity recognition.	EAA outperforms baseline models by at least 0.83% on LOGIC and 3.09% on RuFal, but further NER models and data are needed to enhance robustness and generalizability.
Analyzing the Role of Ukrainian and Russian Diaspora in Disinformation Campaigns [26]	To analyze the role of Ukrainian and Russian diaspora in disinformation campaigns on Twitter.	Social network analysis and ML classification.	The analysis finds that while diaspora communities participate in discussions about the war, their role in spreading disinformation appears limited; this may suggest other actors are more active in these campaigns, although the lack of multilingual data limits these conclusions

The collection of studies highlights the complexity of understanding and combating disinformation, especially in the context of ongoing war. Through the use of sophisticated machine learning techniques, each study provides valuable insights into different facets of information manipulation, revealing both the potential of AI-based solutions and their inherent limitations. Issues such as model scalability, dataset biases, and the stochastic nature of LLM outputs emphasize the need for continued development in both computational techniques and cross-domain research. A multidisciplinary approach with expanded datasets and improved models is crucial for effectively addressing the evolving challenges of disinformation campaigns.

3. Methodology

3.1. XLNet Model

XLNet is an advanced NLP model that was developed to overcome the limitations of earlier models such as BERT. It is based on the Transformer architecture, but what sets XLNet apart is its novel permutation-based training approach [27]. Unlike BERT, which uses a masked language modeling technique, XLNet employs a permutation-based autoregressive training objective that allows the model to effectively capture bidirectional context without sacrificing the natural language structure. This permutation approach ensures that the model can learn dependencies between words in a more generalized manner, thus improving its understanding of the linguistic relationships within a sequence.

The core component of XLNet is its autoregressive formulation based on permutation language modelling. Unlike BERT, which relies on masked language modelling, XLNet generates predictions by maximizing a sequence's likelihood under all possible factorization order permutations. Specifically, given a sequence $x = (x_1, x_2, \dots, x_T)$, XLNet maximizes the following objective:

$$L_{XLNet} = E_{\pi \epsilon P(T)} \sum_{t=1}^{T} \log P\left(x_{\pi_t} | x_{\pi_1}, \dots, x_{\pi_{t-1}}\right), \quad (1)$$

where π denotes a possible permutation of the sequence indices, x_i represents the i-th token in the sequence, P(T) is the joint probability of a sequence of tokens T, E is embedding and the training process involves computing the joint probability of the tokens within the sequence, considering multiple permutations. This autoregressive nature allows XLNet to maintain the bidirectional context without masking tokens, which ensures that the model fully captures the interdependencies between all words in the sentence during training. In contrast, BERT's masked language modelling may result in discrepancies between training and inference due to artificially masked inputs.

The attention mechanism in XLNet is based on Transformer-XL's segment-level recurrence, which improves the model's ability to capture long-term dependencies across sequences. The recurrence mechanism enables the flow of contextual information from one segment to the next, overcoming the fixed-length limitation inherent in many previous Transformer architectures. Formally, let $h_t^{(l)}$ represent the hidden state at layer l and position t for a given sequence segment. Then, XLNet extends the hidden states across segments, thereby modelling the recurrence as follows:

$$h_t^{(l)} = \mathrm{TransformerXL}\left(h_t^{(l-1)}, h_{t-1}^{(l)}\right), \tag{2}$$

where the attention for each layer can refer back to prior segments to retain the context across multiple sentences. This technique is crucial for detecting disinformation because content shared across social media often involves long-range dependencies between messages or posts.

Another key innovation in XLNet is the two-stream attention mechanism, which differentiates between the predicted content (content stream) and the positional query used for prediction (query stream). This mechanism allows XLNet to separately model the target token and context during training. The two-stream attention is defined as follows:

$$q_t = f_q(h_t, x_t), \tag{3}$$

$$h_t = f_c(h_t, q_t), \tag{4}$$

where q_t represents the query stream and h_t is the hidden state of the content stream. By decoupling the two, XLNet can more robustly handle context and target token dependencies, resulting in better predictive performance across tasks involving nuanced linguistic structures such as multilingual disinformation.

The Adam optimizer optimizes XLNet using adaptive learning rates. The permutation-based training objective introduces significant computational complexity, so techniques such as memory-efficient attention and gradient clipping stabilize training. The model's hyperparameters, including the learning rate and warm-up steps, are tuned specifically to ensure convergence, particularly given the noisy nature of the social media data used for disinformation classification.

Incorporating XLNet into the classification of war-related content provided a significant advantage in capturing the underlying disinformation context. By leveraging bidirectional context without the drawbacks of masked inputs and by employing segment-level recurrence and two-stream attention mechanisms, XLNet demonstrated high accuracy and robustness in classifying multilingual tweets as either pro-Russian or neutral/pro-Ukrainian. The performance gains demonstrate the effectiveness of XLNet's unique permutation-based approach in handling complex and highly contextual disinformation.

3.2. Performance Metrics

In this study, the performance of the XLNet model was evaluated using standard classification metrics: precision, recall, F1 score, support, and accuracy. These metrics provide insight into the model's effectiveness in distinguishing between classes.

Precision measures the proportion of correctly predicted positive observations to all predicted positive observations. It is calculated as follows:

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}, \tag{5}$$

where TP (True Positives) represents the number of correctly identified positive instances, and FP (False Positives) represents the number of incorrectly predicted positive instances.

Recall, also known as the sensitivity or true positive rate, measures the proportion of correctly predicted positive observations to all observations in the actual class. It is expressed as follows:

$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}, \tag{6}$$

where FN (False Negatives) represents the number of instances incorrectly predicted as negative.

The F1 score is the harmonic mean of precision and recall, which balances the two. In particular, it is useful when the class distribution is imbalanced. The formula is as follows:

$$\mathrm{F1} = 2 \times \frac{\mathrm{Precision} \times \mathrm{Recall}}{\mathrm{Precision} \times \mathrm{Recall}}. \tag{7}$$

Support refers to the number of actual occurrences of each class in the dataset. This helps provide context for the performance metrics, indicating the number of samples available for evaluation in each class.

Accuracy measures the proportion of correctly predicted instances (both positive and negative) to the total number of instances. It is calculated as follows:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \qquad (8)$$

where TN (True Negatives) represents the number of correctly identified negative instances.

These metrics collectively provide a comprehensive evaluation of the XLNet model's ability to effectively classify disinformation. High precision indicates a low false positive rate, whereas high recall indicates a low false negative rate. The F1 score balances these two metrics to provide a single performance measure. Accuracy provides a general overview of the model's correctness, and support provides additional context to understand the reliability of these metrics for each class.

4. Results

4.1 Data

To collect and process tweets, we employed the UbiLab Misinformation Analysis System (U-MAS) [28], which is a tool built to streamline data extraction and analysis. U-MAS employs a Python script, which depends on two configuration files: twitter-keys.txt, containing the bearer token for API authentication, and twitter_meta.txt, which describes the monthly intervals for data queries. U-MAS provides researchers with specialized access to Twitter's V2 API, granting them a high level of access to both historical and live public data, along with additional features that enhance the completeness and reliability of the datasets collected. The extracted data is stored in JSON format in an Azure Blob Storage container, where they underwent preprocessing as per the research requirements.

In U-MAS, a performance metric is computed by assigning different weights to various engagement metrics, including likes, replies, quotes, and retweets. This allows the system to rank tweets and eliminate duplicates, keeping only the ones with the highest interaction levels. The data were validated by two researchers specializing in Russian disinformation campaigns. The validated data, which included 'N' of the most significant tweets, were used to train the classification models. The metric used to evaluate a tweet's performance was based on its interaction count. A further qualitative analysis performed by domain experts identified pro-Russian and pro-Ukrainian tweets in the dataset. The results of this analysis were used as labels to train the sentiment classification model.

In total, U-MAS was used to extract 42,000 English-language tweets from September 2022. This month was chosen because of significant military events and humanitarian crises that triggered a surge in Russian disinformation and narrative control efforts.

To facilitate a robust examination of pro-Russian sentiment, our methodology also involved developing a search strategy that could systematically identify content pushing a pro-Russian agenda while filtering out pro-Ukrainian tweets. Out of the extracted dataset, the top 5,000 most engaging tweets were manually categorized into two distinct classes: pro-Russian tweets that aligned with Russia's strategic messaging and neutral or pro-Ukrainian tweets that did not support Russian interests. This labeling process was conducted by two researchers experienced in analyzing Russian propaganda efforts. Each researcher independently assigned the tweets to the two classes, and the results were compared to minimize individual bias in classification decisions.

The dataset was also balanced to provide equal representation of both classes, with approximately half of the tweets representing pro-Russian propaganda and the other half representing neutral or pro-Ukrainian content. This balanced distribution was critical for effectively training the machine learning models and preventing bias toward one class. By carefully curating and labelling the dataset in this manner, we aimed to create a robust and reliable dataset that would facilitate a meaningful evaluation of the machine learning models used in this study.

4.2. Model tuning

The process of setting up and training the XLNet model began with comprehensive data preprocessing. This process cleaned the text data by removing unwanted elements, such as hashtags, special characters, URLs, and stop words, to improve the overall quality of the dataset. Stemming was applied to reduce words to their root forms, ensuring uniformity across the data. The dataset was then split into training (80%), validation (10%), and test sets (10%) to provide an effective basis for model evaluation and tuning. The training set was used to teach the model, the validation set was used to facilitate hyperparameter tuning, and the test set was used to evaluate the final performance.

Following preprocessing, an XLNet model was initialized to perform binary classification targeting pro-Russian and neutral/pro-Ukrainian tweets. Tokenization was conducted to convert the text data into a numerical format that the model could interpret. In particular, the text was tokenized into sequences with a maximum length of 256 tokens to maintain consistency across the samples. The sequence length was selected to strike a balance between capturing sufficient context from each tweet and maintaining computational efficiency. Tokenization also involved padding shorter tweets to this fixed length and truncating longer ones to fit the set limit.

The model architecture was optimized by carefully selecting key hyperparameters to improve training stability and accuracy. Training was performed for three epochs because this was found to be optimal for balancing training time with model performance gains. A batch size of 8 was used for both training and evaluation phases, which allowed for efficient memory usage while still providing reliable model updates. The learning rate schedule incorporated a warm-up phase of 500 steps to ensure that the model did not diverge early in training. The weight decay was set to 0.01 to mitigate the risk of overfitting and encourage the model to generalize better by penalizing overly complex solutions.

Training was performed using the Adam optimizer, which dynamically adjusted the learning rates during training to ensure convergence. During the evaluation phase, early stopping was employed to halt training once the performance on the validation set ceased improving, which prevented overfitting. Performance metrics, including accuracy, precision, recall, and F1-score, were recorded to assess the effectiveness of the model across different classes. The final trained model was evaluated on the test set, and it demonstrated its ability to accurately classify tweets into appropriate categories, effectively balancing computational efficiency and predictive power.

4.3. Experimental Results

A structured table 2 representing the performance of the XLNet model. This table summarizes the performance metrics (Precision, Recall, F1-Score, Support, and Accuracy) of the XLNet model for different quantities of tweets and classes (0 and 1).

We used different sizes of the data samples (1000 to 5000 tweets which were divided according to the model tuning methodology (80% / 10% / 10%). The XLNet model's performance, as shown in Table 2, shows a steady ability to categorize tweets into pro-Russian and neutral/pro-Ukrainian classes with consistently high precision across different data sizes. Precision, in this context, measures the proportion of correct positive predictions among all positive predictions. For the smallest set of 1,000 tweets, precision for both categories remained consistently high at 0.95, reflecting the model's ability to minimize false positives. As the dataset grew, precision remained largely stable, experiencing only minor fluctuations. For instance, at the 2,000-tweet level, Class 0 achieved perfect precision, suggesting that all pro-Russian tweets were identified without any false positives. Even with larger datasets of 3,000, 4,000, and 5,000 tweets, precision values remained above 0.91 for both classes, indicating that the model effectively differentiated between relevant and irrelevant content.

Recall is another critical metric that measures the model’s ability to identify all relevant items in each class. For the 1,000-tweet dataset, recall for Class 1 was 0.97, slightly outperforming Class 0, which had a recall of 0.92. This trend persisted as the dataset size increased, with Class 1 consistently showing higher recall rates than Class 0. For example, at the 5,000-tweet level, Class 1 had a recall of 0.95, whereas Class 0 had 0.89. This difference suggests that the model was slightly more proficient at recognizing all neutral/pro-Ukrainian tweets than pro-Russian tweets, possibly due to more consistent language patterns in neutral/pro-Ukrainian content. Nevertheless, recall remained above 0.89 for all data sizes, reinforcing the model’s reliability in identifying relevant tweets across both categories.

The F1-score, which represents the harmonic mean of precision and recall, serves as a balanced measure of the model’s overall effectiveness. For the 1,000-tweet dataset, the F1-score for Class 0 was 0.94, whereas Class 1 scored 0.96, indicating strong performance in tweet classification. As the dataset expanded, the F1-scores remained consistently high, with values between 0.91 and 0.98. At 2,000 tweets, the F1-score for Class 1 reached 0.98, demonstrating the model's ability to maintain a high standard as more data were introduced. Even at higher volumes of 4,000 and 5,000 tweets,

Table 2

Performance of XLNet model

Tweets number	Class	Precision	Recall	F1 score	Support	Accuracy
1000	0	0.95	0.92	0.94	46	0.95
	1	0.95	0.97	0.96	54	
2000	0	1	0.94	0.97	77	0.98
	1	0.96	1	0.98	123	
3000	0	0.93	0.89	0.91	122	0.93
	1	0.93	0.96	0.94	178	
4000	0	0.93	0.90	0.91	172	0.93
	1	0.93	0.95	0.94	228	
5000	0	0.93	0.89	0.91	214	0.92
	1	0.91	0.95	0.93	286	

the F1-scores remained above 0.91 for both classes, demonstrating the model's ability to strike a balance between precision and recall, thereby reducing both false positives and false negatives.

Support, which represents the number of data instances in each class, also played a significant role in evaluating the model's performance. As the dataset size increased, the number of support for each class increased, providing a broader basis for assessing the model's effectiveness. At the 5,000-tweet level, Class 0 had a support value of 214, while Class 1 had 286. The consistent model performance across different levels of support indicates that the model handled varying class distributions well. High support values also helped ensure that performance metrics were not skewed by insufficient data in either class, which contributed to a more reliable evaluation of the results.

The accuracy, which measures the overall correctness of the model predictions, remained high across all data volumes. For the 1,000 tweet dataset, the accuracy was 95%, which indicates reliability even with limited data. The accuracy peaked at 98% for 2,000 tweets, indicating that the model excelled with a moderate dataset size. However, as the tweet volume increased to 3,000, 4,000, and 5,000, accuracy dropped slightly to 93%. This minor decline may be due to the increased diversity and complexity of the data, which introduced more challenging cases. Despite this decrease, an accuracy consistently above 90% across all dataset sizes confirms that the XLNet model remains a powerful tool for classifying social media content, particularly in the context of information warfare and propaganda analysis.

Figure 1 shows the accuracy of the XLNet model versus the number of tweets.

It provides important insights into how the model's performance evolved with increasing data size. Initially, the model's accuracy improved significantly from 95% to 98% as the number of tweets increased from 1,000 to 2,000. This improvement suggests that the model benefited from the additional training data, which provided more context and examples to learn from, resulting in better classification of tweets. However, beyond the 2,000-tweet mark, the accuracy began to decline gradually as more data were added. This decline in performance, which decreased to approximately 93% at 5,000 tweets, indicates that the added complexity in larger datasets introduced more challenging variations and edge cases, which may have made it harder for the model to maintain peak accuracy.

The trend observed in the graph also implies that there may be an optimal data volume for training the XLNet model for this specific task. While increasing the amount of data often helps improve model performance, in this case, the increased number of tweets greater than 2,000 appeared to introduce diminishing returns, potentially due to increased noise or diverse language patterns that made it harder for the model to generalize effectively. This highlights the importance of carefully balancing the dataset size and quality during training. It may also suggest that further hyperparameter tuning or data augmentation is necessary to address the reduction in performance observed with larger datasets, particularly if the goal is to maintain a high level of accuracy in complex real-world scenarios involving information warfare and propaganda analysis.

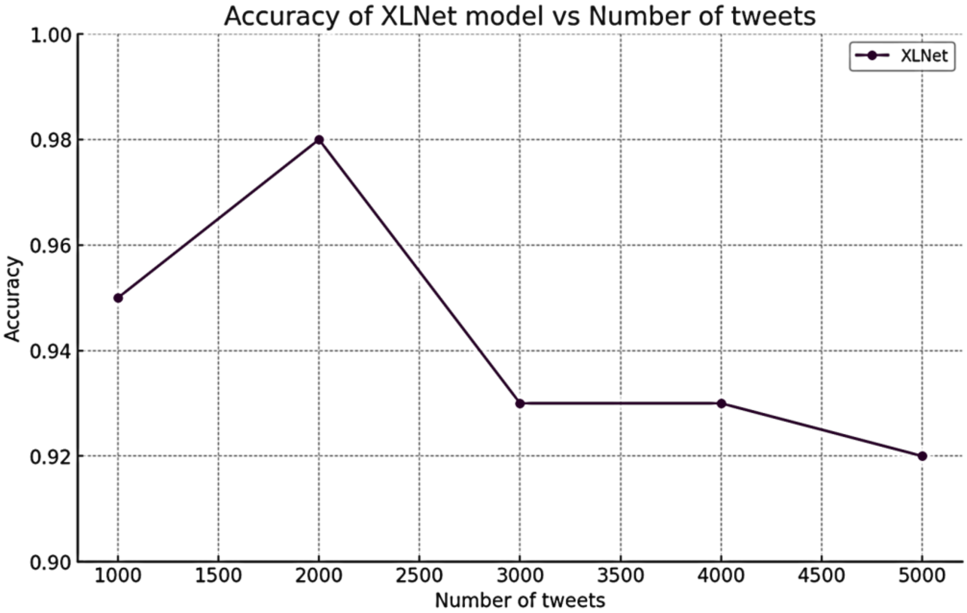

Figure 1. Accuracy of the XLNet model

5. Discussion

The XLNet model demonstrated a strong ability to classify tweets into pro-Russian and neutral/pro-Ukrainian categories, maintaining high precision, recall, and F1-scores across different dataset sizes. Initially, accuracy improved significantly as the number of tweets increased from 1,000 to 2,000, which indicates that the model benefited from a larger and more diverse training dataset. The high accuracy (98%) suggests that adding data improved the model's capacity to generalize and accurately predict unseen instances. However, as the dataset size increased to 3,000, 4,000, and 5,000 tweets, accuracy declined slightly to 93%. This drop in performance may be due to the increased complexity and noise in larger datasets, which introduced challenging variations that the model found difficult to handle effectively.

This accuracy trend suggests that there is an optimal data volume for training the model to achieve the best results. Although an increased dataset size can provide more information for training, it can also introduce variability and noise, which increases complexity of the learning process more complex. To address the reduction in accuracy for larger dataset sizes, hyperparameter tuning or data augmentation may be necessary. In addition, exploring advanced regularization techniques can help the model maintain higher accuracy levels even as the complexity of the data increases. These findings highlight the importance of finding a balance between dataset size and quality during training to ensure optimal model performance, especially in the context of disinformation detection where data variability is high.

Despite its promising performance, the XLNet model has several limitations that must be addressed. A key limitation of the proposed method is its sensitivity to the size and quality of the training dataset. An increase in the dataset size beyond a certain point led to a decline in model accuracy, which was likely due to the introduction of noise and increased complexity that the model struggled to handle. This indicates that XLNet may require extensive preprocessing and noise reduction steps to ensure the quality of training data. Another limitation is the model's computational cost, as XLNet is resource-intensive and requires significant computational power for both training and inference. This can be a barrier for researchers and practitioners with limited access to high-performance computing resources. In addition, the model’s reliance on extensive hyperparameter tuning to achieve optimal performance can make it challenging to deploy in real-world scenarios where time and resources are limited. Lastly, the generalizability of the model to different domains remains an open question because its performance may vary significantly depending on the nature of the data and specific characteristics of the task under consideration.

6. Conclusions

This study presented an application of the XLNet model to classify tweets related to Russia’s war against Ukraine and distinguish between pro-Russian and neutral/pro-Ukrainian content. The model demonstrated strong performance with consistently high precision, recall, F1-scores, and accuracy across varying dataset sizes. While accuracy initially improved with increasing data volume, it began to decline slightly as the dataset size increased. This finding suggests that an optimal data size exists for training models on complex disinformation datasets, where balancing data quality and quantity is crucial for achieving high classification accuracy.

The findings of this research have implications for improving automated disinformation detection methods, particularly in the context of hybrid warfare and online propaganda. Future work could focus on addressing the identified challenges, such as the model’s decreasing accuracy with larger datasets, by refining hyperparameters, employing regularization techniques, or incorporating ensemble methods. In addition, exploring multilingual capabilities could further enhance the model's ability to detect disinformation across different languages and platforms. Overall, the XLNet-based approach holds significant promise for supporting the analysis of social media content and mitigating the impact of disinformation campaigns in modern conflicts.

The scientific novelty of this research lies in the application of the XLNet model for disinformation detection in the context of hybrid warfare, specifically, in Russia’s war against Ukraine. By employing XLNet, a transformer-based architecture known for its superior contextual understanding, this study addresses the complexities of multilingual, context-sensitive propaganda in a highly dynamic social media environment. Unlike earlier models, XLNet's permutation-based training allows it to better capture nuanced relationships within the data, making it particularly effective for distinguishing between subtle pro-Russian and neutral/pro-Ukrainian content.

From a practical perspective, the novelty of this research is evident in its focus on applying an advanced language model to real-world disinformation detection, providing a framework for the automated classification of propaganda and influence campaigns on social media. The insights gained from applying XLNet can support policymakers and analysts in identifying and mitigating the effects of disinformation, thereby contributing to more effective information governance strategies during wartime. In addition, this study presents a practical understanding of the challenges related to data quality and volume, which are crucial for improving automated disinformation detection tools.

Future research may focus on enhancing the XLNet model's ability to handle larger datasets by exploring advanced regularization techniques, data augmentation methods, or the use of ensemble models to maintain high accuracy even with complex data. In addition, incorporating multilingual capabilities broadened the scope of the proposed model, which allowed it to detect disinformation in various languages and across multiple platforms. Furthermore, future studies could evaluate the generalizability of the model to other types of conflicts and disinformation scenarios, providing a more comprehensive understanding of the model's applicability in diverse settings.

Contributions of authors: conceptualization – **Halyna Padalko**; methodology – **Halyna Padalko, Vasyl Chomko**; formulation of tasks – **Halyna Padalko, Sergiy Yakovlev**; analysis – **Halyna Padalko, Vasyl Chomko, Sergiy Yakovlev, Plinio Pelegrini Morita**; development of model – **Halyna Padalko, Vasyl Chomko**; verification – **Halyna Padalko, Vasyl Chomko**; visualization – **Halyna Padalko, Vasyl Chomko**; writing – original draft preparation – **Halyna Padalko**; writing – review and editing – **Vasyl Chomko, Sergiy Yakovlev, Plinio Pelegrini Morita**.

Conflict of interest

The authors declare that they have no conflict of interest concerning this research, whether financial, personal, authorship, or otherwise, that could affect the research and its results presented in this paper.

Data availability

The data used in this study is available upon request to the corresponding author.

Use of Artificial Intelligence

The authors confirm that they used generative artificial intelligence methods for editing in their work.

All the authors have read and agreed to the publication of the final version of this manuscript.

References

1. Bontridder, N., & Poullet, Y. The Role of Artificial Intelligence in Disinformation. *Data & Policy*, 2021, vol. 3, article no. E32, DOI: 10.1017/dap.2021.20.

2. Bahruz, E.T. Manipulation as a Form of information-psychological War. *Revista Universidad y Sociedad*, 2023, vol. 15, no. 5, pp. 143–150. Available at: http://scielo.sld.cu/scielo.php?pid=S2218-36202023000500143&script=sci_abstract (Accessed 1 Sep. 2024).

3. Manheim, K., & Kaplan, L. Artificial Intelligence: Risks to Privacy and Democracy. *Yale Journal of Law & Technology*, 2019, vol. 106, 83 p. Available at: https://yjolt.org/artificial-intelligence-risks-privacy-and-democracy (Accessed 1 Sep. 2024).

4. Ewe, K. *Elections Around the World in 2024*. TIME, 2023. Available at: https://time.com/6550920/world-elections-2024/ (Accessed 1 Sep. 2024).

5. Interference 2024. *Interference Tracker 2024*, 2024. Available at: https://interference2024.org/ (Accessed 1 Sep. 2024).

6. Müller, M. M. *Looking Doppelganger: an Analysis of Evolving State-Sponsored Disinformation Tactics*. Utwente.nl, 2024. Available at: https://purl.utwente.nl/essays/102708 (Accessed 1 Sep. 2024).

7. Serrano-Puche, J. Digital Disinformation and emotions: Exploring the Social Risks of Affective Polarization. *International Review of Sociology*, 2021, vol. 31, no. 2, pp. 231–245. DOI: 10.1080/03906701.2021.1947953.

8. Tenove, C. Protecting Democracy from Disinformation: Normative Threats and Policy Responses. *The International Journal of Press/Politics*, 2020, vol. 25, no. 3, pp. 517–537. DOI: 10.1177/1940161220918740.

9. Raad, A. Protecting Freedom of Thought: Mitigating Technological Enablers of Disinformation. *Centre for International Governance Innovation*, 2024. Available at: https://www.cigionline.org/publications/protecting-freedom-of-thought-mitigating-technological-enablers-of-disinformation/ (Accessed 1 Sep, 2024).

10. Mohammadi, A., Meniailov, I., Bazilevych, K., Yakovlev, S., & Chumachenko, D. Comparative study of linear regression and SIR models of COVID-19 propagation in Ukraine before vaccination. *Radioelectronic and Computer Science*, 2021, vol. 3, pp. 5-18. DOI: 10.32620/reks.2021.3.01.

11. Choraś, M., Demestichas, K., Gielczyk, A., Herrero, A., Ksieniewicz, P., Remoundou, K., Urda, D., & Wozniak, M. Advanced Machine Learning techniques for fake news (online disinformation) detection: A systematic mapping study. *Applied Soft Computing*, 2021, vol. 101, article no. 107050. DOI: 10.1016/j.asoc.2020.107050.

12. Padalko, H., Chomko, V., Yakovlev, Y., & Chumachenko, D. Ensemble Machine Learning Approaches for Fake News Classification. *Radioelectronic and Computer Systems*, 2023, no. 4, pp. 5–19. DOI: 10.32620/reks.2023.4.01.

13. Tianda, I. M., Ubadah, M. N., Mardianto, M. F. F., Munawwarah, A., & Ana, E. Clustering Fake News with K-Means and Agglomerative Clustering Based on Word2Vec. *International Journal of Mathematics and Computer Research*, 2024, vol. 12, no. 02, pp. 3999–4007. DOI: 10.47191/ijmcr/v12i2.01.

14. Chumachenko, D., Piletskiy, P., Sukhorukova, M., & Chumachenko, T. Predictive Model of Lyme Disease Epidemic Process Using Machine Learning Approach. *Applied Sciences*, 2022, vol. 12, no. 9, article no. 4282. DOI: 10.3390/app12094282.

15. Akhtar, P., Ghouri, A.M., Khan, H.U.R., Haq, M.A., Awan, U., Zahoor, N., Khan, Z., & Ashraf, A. Detecting fake news and disinformation using artificial intelligence and machine learning to avoid supply chain disruptions. *Annals of Operations Research*, 2022, vol. 327, pp. 633-657. DOI: 10.1007/s10479-022-05015-5.

16. Chumachenko, D., Butkevych, M., Lode, D., Frohme, M., Schmailzl, K. J. G., & Nechyporenko, A. Machine Learning Methods in Predicting Patients with Suspected Myocardial Infarction Based on Short-Time HRV Data. *Sensors*, 2022, vol. 22, no. 18, article no. 7033. DOI: 10.3390/s22187033.

17. Park, C. Y., Mendelsohn, J., Field, A., & Tsvetkov, Y. Challenges and Opportunities in Information Manipulation Detection: an Examination of Wartime Russian Media, *Findings of the Association for Computational Linguistics: EMNLP 2022*, 2022, pp. 5209-5235. DOI: 10.18653/v1/2022.findings-emnlp.382.

18. Pilkevych, I. A., Fedorchuk, D. L., Romanchuk, M. P., & Naumchak, O. M. Approach to the Fake News Detection Using the Graph Neural Networks. *Journal of Edge Computing*, 2023, vol. 2, no. 1, pp. 24–36. DOI: 10.55056/jec.592.

19. Durani, K., Eckhardt, A., Durani, W., Kollmer, T., & Augustin, N. Visual audience gatekeeping on social media platforms: A critical investigation on visual information diffusion before and during the Russo–Ukrainian War. *Information Systems Journal*, 2023, vol. 34, iss. 2, pp. 415-468. DOI: 10.1111/isj.12483.

20. Maathuis, C., & Kerkhof, I. First Six Months of War from Ukrainian topic and sentiment analysis. *European Conference on Social Media*, 2023, vol. 10, no. 1, pp. 163–173. DOI: 10.34190/ecsm.10.1.1147.

21. Marigliano, R., Hui, L., & Carley, K. M. Analyzing Digital Propaganda and Conflict rhetoric: a Study on Russia's bot-driven Campaigns and counter-narratives during the Ukraine Crisis. *Social Network Analysis and Mining*, 2024, vol. 14, no. 1, article no. 170. DOI: 10.1007/s13278-024-01322-w.

22. Aguerri, J. C., Santisteban, M., & Miró-Llinares, F. The Fight against Disinformation and Its consequences: Measuring the Impact of 'Russia state-affiliated Media' on Twitter. *Crime Science*, 2024, vol. 13, no. 1, 17. DOI: 10.1186/s40163-024-00215-9.

23. Lipianina-Honcharenko, K., Bodyanskiy, Y., Kustra, N., Ivasechko, A. OLTW-TEC: Online Learning with Sliding Windows for Text Classifier Ensembles. *Frontiers in Artificial Intelligence*, 2024, vol. 7, article no. 1401126. DOI: 10.3389/frai.2024.1401126.

24. Makhortykh, M., Sydorova, M., Baghumyan, A., Vziatysheva, V., & Kuznetsova, E. Stochastic lies: How LLM-powered Chatbots Deal with Russian Disinformation about the War in Ukraine. *Harvard Kennedy School Misinformation Review*, 2024, vol. 5, no. 4. DOI: 10.37016/mr-2020-154.

25. Shultz, B. An Entity-Aware Approach to Logical Fallacy Detection in Kremlin Social Media Content. *ASONAM '23: Proceedings of the International Conference on Advances in Social Networks Analysis and Mining*, 2023, pp. 780-783. DOI: 10.1145/3625007.3627988.

26. Maathuis, C., De Ridder, C., & Stuurman, S. Analyzing the Role of Ukrainian and Russian Diaspora in Disinformation Campaigns. *European Conference on Social Media*, 2023, vol. 10, no. 1, pp. 153–162. DOI: 10.34190/ecsm.10.1.1118.

27. Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R., & Le, Q. XLNet: Generalized Autoregressive Pretraining for Language Understanding. *33rd Conference on Neural Information Processing Systems (NeurIPS 2019)*, 2019, article no. 161263.

28. Hussain, I. Z., Kaur, J., Lotto, M., Butt, Z. A., & Morita, P. P. Tweeting for Health Using Real-Time Mining and AI-Based Analytics: Design & Development of as Misinformation Data Ecosystem for Twitter (Preprint). *Journal of Medical Internet Research*, 2022, vol. 25, article no. e44356. DOI: 10.2196/44356.

Received 10.07.2024, Accepted 18.11.2024

КЛАСИФІКАЦІЯ ДЕЗІНФОРМАЦІЇ В УМОВАХ ГІБРИДНОЇ ВІЙНИ: ЗАСТОСУВАННЯ XLNET ВПРОДОВЖ ВІЙНИ РОСІЇ ПРОТИ УКРАЇНИ

Г. А. Падалко, В. Д. Чомко, С. В. Яковлев, П. П. Моріта

Поширення дезінформації стало критичним компонентом гібридної війни, що особливо помітно у російській війні проти України, де соціальні медіа стали полем битви за вплив та пропаганду. **Метою** цього дослідження є розробка комплексної методології для класифікації дезінформації в контексті гібридної війни, зосереджуючи увагу на війні Росії проти України. **Об'єктом** цього дослідження є вирішення проблем виявлення дезінформації, особливо в умовах збільшення пропаганди через гібридну війну. **Предметом** дослідження є використання мовних моделей на основі трансформерів, зокрема XLNet, для класифікації багатомовної, контекстно-залежної дезінформації. **Завданнями** цього дослідження є проведення аналізу сучасних досліджень та розробка методології ефективної класифікації дезінформації за допомогою моделі XLNet. Запропонована **методологія** включає кілька ключових компонентів: попередню обробку даних для забезпечення їх якості, застосування XLNet для навчання на різноманітних наборах даних та оптимізацію гіперпараметрів для врахування складнощів дезінформаційних даних. У дослідженні використовувалися набори даних, що містили проросійські та нейтральні/проукраїнські твіти, причому модель XLNet демонструвала високі показники, включаючи високу точність, повноту та міру F1 для різних обсягів даних. **Результати** показали, що точність спочатку покращувалася зі збільшенням обсягу даних, але трохи знижувалася при дуже великих наборах даних, що вказує на необхідність балансу між якістю та кількістю даних. Запропонована методологія заповнює прогалини в автоматизованому виявленні дезінформації шляхом інтеграції моделей на основі трансформерів із сучасними методами попередньої обробки та навчання. Це дослідження покращує можливості для виявлення та аналізу дезінформації в режимі реального часу, сприяючи управлінню суспільною інформацією та стратегічній комунікації під час конфліктів.

Ключові слова: гібридна війна; виявлення дезінформації; машинне навчання; XLNet; аналіз соціальних медіа; трансформерні моделі.

Падалко Галина Анатоліївна – здобувачка каф. математичного моделювання та штучного інтелекту, Національний аерокосмічний університет ім. М. Є. Жуковського «Харківський авіаційний інститут», Харків, Україна, здобувачка департаменту політичних наук в Університеті Ватерлу, Канада, дослідниця Школи міжнародних відносин Балсіллі, Ватерлу, Канада.

Чомко Василь Дмитрович – студент магістерської програми з прикладних наук у галузі інженерії системного проєктування в Університеті Ватерлу, Канада.

Яковлев Сергій Всеволодович – д-р фіз.-мат. наук, проф., заступник директора Інституту комп'ютерних наук та штучного інтелекту Харківського національного університету імені В. Н. Каразіна, Харків, Україна; Професор-дослідник Інституту математики Лодзького технологічного університету, Лодзь, Польща.

Моріта Пелегріні Плініо – доцент Школи Наук Публічного Здоров'я, Університет Ватерлу, Ватерлу, Канада

Halyna Padalko – PhD Student, Department of Mathematical Modeling and Artificial Intelligence, National Aerospace University "Kharkiv Aviation Institute," Kharkiv, Ukraine, Visiting PhD student, Department of Political Science at the University of Waterloo, Waterloo, Canada, Research Fellow, Balsillie School of International Affairs, Waterloo, Canada.
e-mail: hpadalko@uwaterloo.ca, ORCID: 0000-0001-6014-1065.

Vasyl Chomko – Master of Applied Science Student, Systems Design Engineering at the University of Waterloo. Canada,
e-mail: vchomko@uwaterloo.ca, ORCID: 0009-0009-4419-6651.

Sergiy Yakovlev – Corresponding Member of the National Academy of Sciences of Ukraine, Doctor of Physical and Mathematical Sciences, Professor, Deputy Director of Institute of Computer Science and Artificial Intelligence at V. N. Karazin Kharkiv National University, Kharkiv, Ukraine; Research Professor at the Institute of Mathematics, Lodz University of Technology, Lodz, Poland,
e-mail: sergiy.yakovlev@p.lodz.pl, ORCID: 0000-0003-1707-843X.

Plinio Pelegrini Morita – Associate Professor, School of Public Health Sciences, University of Waterloo, Waterloo, Canada,
e-mail: plinio.morita@uwaterloo.ca, ORCID: 0000-0001-9515-6478.