РОЗДІЛ ІІ. ІНФОРМАЦІЙНО-КОМП'ЮТЕРНІ ТЕХНОЛОГІЇ

DOI: 10.25140/2411-5363-2023-4(34)-82-90
UDC 004.02:[323.266:316.472.4]

Liudmyla Mishchenko[1], Iryna Klymenko[2]

[1]PhD student of Department of Computer Engineering,
National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute" (Kyiv, Ukraine)
E-mail: liudamishchenko@gmail.com. **ORCID**: https://orcid.org/0000-0002-5101-2517

[2]Prof., ScD of Department of Computer Engineering
National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute" (Kyiv, Ukraine)
E-mail: ikliryna@gmail.com. **ORCID**: http://orcid.org/0000-0001-5345-8806

METHOD FOR DETECTING FAKE NEWS THROUGH WRITING STYLE

In the era of digital technologies, distinguishing truth from misinformation is a challenging task. Fake news, characterized by deceitful narratives, poses a significant threat. Traditional fact-checking methods often overlook the nuances of linguistic stylistic coloring. This study employs an advanced writing style analysis that extends beyond conventional methodologies. Several linguistic dimensions of texts are considered in this research, emphasizing on pre-processing and function development. The experiments are based on various datasets. Thus, the developed method for detecting fake news utilizes a multidimensional approach. The proposed development includes meticulous verification of the dataset, pre-processing, and function development, focusing on emotionally charged vocabulary, word groups used in reports indicating event likelihood, mild cursing, and non-standard lexicon. Significant differences in linguistic features were identified, contributing to a nuanced understanding of the construction and creation of deceptive texts. The research results demonstrate that this method accurately distinguishes genuine from fake news articles based on writing style. This study represents significant progress in identifying phony news through writing style analysis, aiding in combating misinformation in the era of digital technologies.

Keywords: *fake news; real news; writing style analysis; emotionally colored words; emotional language; probabilistic words; profanity detection.*
Table: 2. Fig.: 1. References: 10.

Urgency of the research. The proliferation of misinformation and fake news in the digital sphere poses a severe threat to society. With the exponential growth of information sources and the ease of disseminating content online, distinguishing between credible and false information has become increasingly challenging. This research addresses the critical need to combat the spread of fake news, providing methods to analyze the text through linguistic patterns and writing styles.

The topic covered in this research aligns closely with the interdisciplinary realm encompassing language, computer science, and artificial intelligence. It intersects with natural language processing, machine learning, and linguistic analysis. The study aims to develop sophisticated algorithms and models capable of differentiating between genuine and fabricated news articles based on nuanced linguistic features. This amalgamation of linguistic expertise and technological innovation is pivotal in enhancing information reliability and trustworthiness in our digital age.

In essence, the urgency of this research lies in its ability to bridge the gap between linguistic analysis and technological advancements, providing a robust framework to combat the proliferation of fake news.

Target setting. The problem statement for the research on "Method for Detecting Fake News Through Writing Style" centers on developing robust techniques to differentiate between authentic and fake news articles based on linguistic patterns. The critical challenge lies in combating the proliferation of misinformation, which undermines public trust and distorts societal discourse.

Solving this problem is essential as fake news disrupts information integrity, affects decision-making processes. Detecting deceptive news through writing styles and advanced computational methods contributes to increasing reliability of methods for detecting fake news, ensuring the dissemination of accurate information to the public.

The significance of addressing this problem extends to various fields such as computer science, journalism, and information technology. It aligns with the urgent need to create reliable tools that leverage linguistic analysis and machine learning to enhance media literacy, safeguard democratic processes, and promote ethical journalism practices in the digital age.

Actual scientific researches and issues analysis. The problem of fake news in the era of digital media and the rapid circulation of information, characterized by the deliberate dissemination of deceptive or false information, has drawn significant attention from numerous researchers. Detecting fake news is a complex task, and many scientific works delve into the nuances of this issue with a focus on writing style analysis.

The research "Fake News Detection on Social Media: A Data Mining Perspective" [1] outlines the problem of identifying fake news on social media platforms, introducing a structure for intelligent data analysis. Emphasis is placed on the importance of linguistic and stylistic features in detecting deceptive content, with a specific analysis of writing style. However, the work highlights the need for a multidimensional approach that includes linguistic analysis and pays more attention to user behavior, source reliability, and domains not covered within the proposed method.

In another scholarly work [2], a theoretical approach to early detection of fake news is employed. Linguistic features and writing style are crucial indicators, focusing on language differences between true and false news. However, while the researchers propose a theoretical analysis model, practical implementation and actual experiments still need to be improved. Moreover, their reliance on templated article writing may not align with contemporary standards.

This comprehensive review [3] discusses fundamental theories and methods for detecting fake news. It emphasizes the role of linguistic characteristics and writing style in differentiating genuine news from fake news. However, the work only covers certain aspects of writing style, such as emotional tone or cursing.

In the study by other authors [4], the language of hostility is frequently encountered in fake news. This work identifies hostile expressions based on writing style analysis, emotional tone, and linguistic markers associated with harmful content, contributing to a comprehensive understanding of detecting fake news. However, certain aspects of this work still need resolution, such as achieving more detailed classification and a broader range of categories in the database.

Researchers in [5] develop automatic methods for detecting fake news based on writing style. They create two classifiers: a neural network and a model based on stylometric characteristics. The stylometric model analysis focuses on an affective lexicon typical of fake news. However, the study does not consider profanity, words describing event likelihood, or emotional coloration in the text. These factors significantly impact the results and broaden the possibilities of news analysis.

Natural Language Processing (NLP) involves creating algorithms and models to identify and autonomously differentiate accurate news articles from deceptive or fabricated content. As discussed in a previous article [6], NLP technology was central to detecting fake news. All subsequent developments aim to improve existing methods' accuracy, classification, or speed.

Uninvestigated parts of general matters defining. Despite making a valuable contribution to the detection of fake news through writing style analysis, it's important to acknowledge certain limitations:

- limited representativeness of the dataset: Research outcomes rely on the representativeness of the input data. If the dataset fails to encompass the full spectrum of writing styles and characteristics present in digital spaces, the model's generalizability may be limited;
- static nature of analysis: The study primarily focuses on static analysis of writing styles, potentially overlooking the dynamic evolution of language models. A more proactive approach considering temporal changes could enhance the model's effectiveness;
- exclusion of multimedia elements: Fake news often includes multimedia elements (images, videos) to deceive audiences. Solely focusing on textual content analysis may restrict applicability to rich multimedia disinformation;

- human interpretation: Understanding the decision-making process of the model's outcomes requires a more detailed discussion. Exploring alternative AI methods could enhance transparency in the model's results;
- intercultural sensitivity: To improve research outcomes, it's crucial to consider intercultural differences in writing styles and linguistic characteristics fully. Adapting the model to diverse cultural contexts is essential for its global applicability.

Acknowledging these shortcomings lays the groundwork for future research or refining and expanding the proposed methodology for detecting fake news.

The research objective. This research aims to advance the field of detecting fake news through a complex writing style analysis. Delving into the intricate dimensions of linguistic stylistics, the study seeks to identify characteristic patterns that distinguish deceptive narratives from authentic content. Another essential objective is to contribute to a comprehensive understanding of writing styles as a powerful tool in addressing issues related to fake news in the era of digital technologies. The research is focused on developing reliable methods to differentiate between genuine and deceptive narratives based on writing style models through careful feature engineering, text vectorization, and model selection.

The statement of basic materials. The implementation of the proposed method is a complex approach, which by such other methods and means. To achieve the goal and test the ideas in practice, it is necessary to perform the following points: collection and preparation of data, development of a technique for distinguishing stylistic color (departure from the standard algorithm and a feature of the method), vectorization of text selection, model and its training. A detailed description of each step and evaluation of the model's performance are described below.

Data Collection and Preparation. The fundamental step in this scientific research involves a meticulous check and preparation of the dataset, a process crucial for building a solid foundation upon which further analysis is grounded. Choosing the appropriate dataset is paramount as it is a significant basis encompassing various writing styles and subtleties. The richness of diverse input data combined systematically aids in training the model.

The dataset chosen for the proposed method comprises a collection of textual artifacts gathered from various online sources. This ensures a representative sample encompassing diverse writing styles, tones, and contexts. Such a selection encapsulates many writing styles and texts authored by different individuals, allowing an analysis of stylistic content within the text, such as profanity, emotionally charged language, reporting lexicon, casual language, and words indicating event likelihood. This diversity is an integral part of the multifaceted language of writing. It also enables the analysis of news texts used across various online platforms, ranging from news articles to social media discourse.

The subsequent stage involves a meticulous data preparation process. It consists of systematic steps to refine the raw text corpus. Text preprocessing techniques are applied to standardize the dataset, ensuring uniformity while preserving authentic linguistic characteristics inherent in various sources. This phase involves tokenization, stemming, and lemmatization, guaranteeing a consistent and standardized representation of linguistic artifacts.

Moreover, the dataset undergoes a thorough cleansing process to mitigate the influence of stop words (known as noise) and irrelevant artifacts. Punctuation marks are unrelated, and extraneous symbols are removed, creating an appropriate environment for subsequent analyses. Overall, the preparation process results in a dataset cleansed of irrelevant information for training purposes while retaining its authentic emotional tone [7].

Feature Engineering. Since the research aims to recognize writing style entities, the traditional development of software functions must be altered. Based on this, specific lexical directions have been highlighted:

- **emotionally charged lexicon**: this includes words and phrases that carry a strong emotional tone, often evoking certain feelings, reactions, and moods in the reader or listener. These words can elicit positive and negative emotional responses and are frequently used to persuade, influence, or engage the audience emotionally. Examples include "heartfelt", "horrifying", "charming", "tragic", "inspiring".
- **reportage description words**: these are crucial in storytelling, dialogues, interviews, and journalistic articles; these words help attribute statements and actions to specific individuals, enriching the text and making it more appealing and informative. Examples include say, tell, ask, whisper, assert, suggest, report, explain, agree, deny.
- **probability words**: these convey uncertainty, likelihood, or probability. They often express the possibility of an event or state, creating a sense of probability or randomness. In different contexts, probability words help convey the degree of reliability or uncertainty. Examples include probable, possible, likely, maybe, perhaps, seems, might.
- "**soft swear words**" typically refer to expressions or phrases considered relatively mild or less offensive than more substantial forms of swearing. These expressions can still convey disappointment, irritation, or emphasis but are generally less vulgar or socially unacceptable. People often use soft swear words to express their feelings without explicit or harsh language. Examples of soft swear words include "darn," "nonsense," "crap," "rascal," "scoundrel," "balderdash".

The examples of words are singled out as text units, while soft swear words and profanity are usually phrases or expressions. This poses an additional challenge, requiring a focus on the nuances of syntactic complexity, discourse coherence, and semantic subtleties, aiming to comprehend the word's meaning and the essence of the expression.

A similar classification model based on linguistic features described in the article [8] already exists and can serve as a basis for development. This approach has proven effective, although the authors note the necessity to expand the dataset to include new classifications. Such expansions are proposed in this research.

Text Vectorization. The phase of text vectorization in this scientific research is a critical moment for accurate analysis considering the context. This transition from linguistic nuances to numerical representations requires the application of modern methods to encapsulate contextual intricacies.

The process commences with applying contextual embeddings to capture word relationships in the text. The TF-IDF algorithm generates numerical values for words while retaining an understanding of their semantics within the context in which they're used. Compared to static embeddings, this approach provides a more dynamic and contextually sensitive representation, acknowledging the importance of linguistic expression.

Often pre-trained on extensive linguistic corpora, models prove effective in encoding context within textual data, enriching the vectorization process with detailed representations of language structures.

This approach goes beyond the traditional "bag of words" paradigm, acknowledging the limitations of such methodologies in capturing language subtleties. By considering the context of words within a document, vectorization becomes a dynamic exploration of specific words or phrases and their contextual nuances, enhancing the precision of representing core language structures.

The obtained numerical word representations form the basis for subsequent analysis steps but often result in high-dimensional vectors. This occurs because they encapsulate linguistic information within the texts and complex contextual connections between words, enabling the examination of the semantic meaning of expressions and phrases.

Model Selection and Training. The formation of the architecture of the neural model is a critical stage. It involves designing the model's architecture and training process, considering its interpretability and adaptability. Ensemble models form the analytical core due to their interpretability. The training process meticulously calibrates parameters, balancing accuracy and generalization without overfitting. This approach demonstrates the practical utility of models in analyzing linguistic features. The training process refrains from arbitrary optimization, focusing on preserving stylistic features in the input data and adapting to the dynamic evolution of writing styles.

The selected model is based on a hierarchical linguistic tree and a neural network that retains information about the tree. It also involves researching the differential linguistic style of fake news and truth and adopting the peculiarities of some fake news models [9]. These models create a hierarchical linguistic tree of news documents, comparing the linguistic style of each news document based on the words used by its author and how these words are recursively structured into phrases, sentences, paragraphs, and finally, the document. By integrating the hierarchical linguistic tree with the neural network, the proposed method learns and classifies representations of news documents, capturing their locally sequential and globally recursive structures, which are linguistically significant.

An essential step for the proposed algorithm is the word classification stage. A specific approach was chosen for this developmental stage to accomplish classification tasks with high accuracy without extensive computation. The combination of TF-IDF (term frequency-inverse document frequency) and the simple Naive Bayes classifier is a well-established and widely accepted solution in text classification. TF-IDF provides a means to highlight features that capture the importance of words in textual documents. Simultaneously, the Naive Bayes classifier, as a probabilistic classifier, makes predictions based on conditional probabilities of features given class labels.

The Naive Bayes classifier is a family of simple probabilistic classifiers based on the general assumption that all features are independent of each other, given a variable category, and is often used as a baseline in text classification [10].

As words acquire vector values after the computation of the TF-IDF score, they become a discrete quantity. Therefore, the Naive Bayes algorithm was applied to this word representation to execute the classification task.

In the subsequent step, the Naive Bayes classifier is trained using a training set of fake and verified news. During training, the model calculates the probability of the appearance of words in each class. These probabilities are derived from TF-IDF vectors associated with each class.

Once trained, the model can classify new documents without additional training. Initially, new documents are transformed into vectors using the TF-IDF vectorizer used during training. Then, the Naive Bayes model is employed to predict the class label for each document.

Model Evaluation. To evaluate the effectiveness of the model in the classification of real and fake news articles, taking into account the goal of the work, the methods of assessment Accuracy, Precision, Recall and F1 were chosen.

Accuracy – the accuracy of the results, compared to other studies or data sets. Calculated according to formula (1).

$$Accuracy = \frac{(True\ Positives + True\ Negatives)}{True\ Positives + False\ Positives + True\ Negative + False\ Negative} \tag{1}$$

Precision (2) measures the accuracy of optimistic predictions made by the model. It is the ratio of correctly predicted positive observations to the total number of positive observations. A high Precision score indicates a low level of false positive results.

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives}. \tag{2}$$

Recall, also known as sensitivity, measures the model's ability to find all positive instances. It is the ratio of correctly predicted positive observations to the actual positive observations.

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives}. \quad (3)$$

The F1 score is a combination of precision and recall, and calculates (4) a balanced measure between the two, indicating the optimal mean between the precision and recall measures.

$$F_1 = 2\frac{Precision \cdot Recall}{Precision + Recall} \quad (4)$$

In the context of fake news detection, Precision will help estimate the proportion of correctly identified genuine news articles among all articles that are believed to be genuine. At the same time, Recall will estimate how many genuine news articles were correctly identified out of all genuine news articles. Accuracy acts as a fundamental indicator that calculates the overall correctness of system classifications.

The classification accuracy values were calculated based on the evaluation methods described above. Building on the training results, a classification accuracy of 92% was achieved, indicating the model's ability to distinguish real news articles from fake ones. However, in this case, an essential indicator that should be considered in similar studies is the precisely defined and limited criteria for classification.

A rather high result was obtained according to the Recall assessment. Choosen input data contains more than half of the true news, but to train the model, it is necessary to select a set with more false information. Then further analysis will increase the accuracy rate of fake news recognition.

To calculate the results, the model was trained and based on it, 4 different sets of news from different sources were analyzed: social networks Facebook and X, verified news databases EUvsDisinfo, ZaboronaMedia. The total sample size is 5,000 news items. All news are verified by journalists and marked as "true" or "fake", where 3,750 of them are fake, the rest are true. The results of the classification of the proposed approach compared to news brands from fact-checkers are shown in Table 1.

Table 1. Obtained results of binary classification

Result	True positive	True negative	False positive	False negative
Number of entities	3150	1050	200	600

In accordance with the above-described evaluation method, the following indicators were calculated:

- $Accuracy = \frac{3150 + 1050}{5000} = 0.84$ or 84 %;
- $Precision = \frac{3150}{3150 + 200} = 0.94$ or 94 %;
- $Recall = \frac{3150}{3150 + 600} = 0.84$ or 84 %;
- $F_1 = 2\frac{0.94 \cdot 0.84}{0.94 + 0.84} = 0.88$ or 88 %.

Thus, the value of the accuracy of the classification Accuracy and the F1 indicator were calculated. Based on the results of the model's training, a classification accuracy of 84 % was achieved, which indicates the ability of the model to distinguish real news articles from fake ones. Although in this case it is also important for accuracy, an important indicator that should not be neglected in similar studies is precisely defined and limited classification criteria.

To justify the proposed method of recognizing fake news by written style, a number of experiments were conducted. The obtained results showed that the amount of light swearing and obscene expressions, emotionally colored words, words of probability make up a larger percentage of the text in fake news, compared to true ones. However, the words to describe the reports are a smaller share. The percentage of such groups of words in news texts is shown in Table 2.

Table 2. Presence of a certain group of words in true and fake news in percentage

Type of emotionally colored words	**Real news**, %	**Fake news**, %
Mild curses	7	12
Emotional words	36	55
Profanity	3	8
Reporting Words	23	17
Probabilistic Words	10	26

For visual representation of the influence of emotionally colored words on the contextual content of the text of fake news, a graph of the percentage ratio of a certain group of words in false or manipulated and true texts is plotted. It is shown in Figure 1.

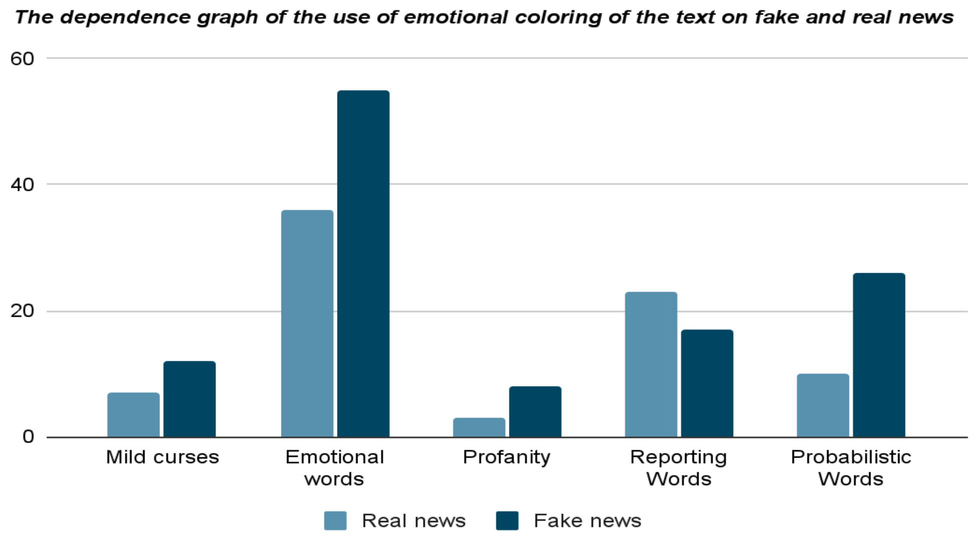

Fig. 1. Comparison of the percentage content of emotionally colored words in the news text.

The graph shows that fake news contains a significantly higher percentage of emotional, probabilistic words, light taunts, and profanity. However, directive words are used less frequently. Typically, authors of fake news try to refrain from concluding less often. Thus, they encourage readers to make subjective or inaccurate conclusions.

Conclusions. The research aims to distinguish fake news from genuine news through writing style analysis. Traditional fact-checking methods often overlook linguistic nuances in texts that conceal deceptive narratives. Thus, this study delves into the detailed stylistic features of specific language expressions, such as emotionally charged words, profanity, probabilistic terms, and reports. A carefully curated dataset represented diverse linguistic shades prevalent in the digital space. Preprocessing methods were meticulously chosen to ensure dataset standardization while preserving its internal diversity.

This work represents a significant step towards developing reliable methods to detect fake news using writing style analysis. It explores intricate dimensions of linguistic expression and highlights the critical role of function development, text vectorization, model selection, and teaching the differentiation between deceptive narratives and truthful content.

The results indicate a high classification accuracy of 92 %, demonstrating the model's effectiveness in recognizing real and fake news articles based on writing style. The obtained results were analyzed for the presence of a certain group of words in the news texts and presented as a percentage. The findings reveal that fake news contains a higher occurrence of emotionally charged words, probabilistic terms, and profanity. Moreover, words indicating reports or journalism are used less frequently, often avoiding explicit conclusions, prompting readers to form conclusions and subjectively interpret events.

The research underscores the importance of a nuanced understanding of writing styles as a potent tool in combating fake news issues in the digital era. The findings provide a foundation for future advancements in this field, pointing towards avenues for studying multimodal analysis, deep learning architectures, contextual analysis, user behavior, cross-lingual detection, source verification, real-time analysis, ethical considerations, and educational and informational initiatives.

Therefore, this work contributes to ongoing efforts to mitigate the impact of fake news in society by understanding linguistic nuances prevalent in deceptive narratives. It successfully enhances the precision of fake news detection systems and broadens the identification methods of misleading information. Thus, the research goal can be achieved. This work forms a crucial foundation for future developments aimed at combating the spread of misinformation and upholding the integrity of digital information systems.

References

1. Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD explorations newsletter, 19(1),* 22-36.

2. Zhou, X., Jain, A., Phoha, V. V., & Zafarani, R. (2020). Fake news early detection: A theory-driven model. *Digital Threats: Research and Practice, 1(2),* 1-25.

3. Zhou, X., & Zafarani, R. (2020). A survey of fake news: Fundamental theories, detection methods, and opportunities. *ACM Computing Surveys (CSUR), 53(5),* 1-40.

4. Schmidt, A., & Wiegand, M. (2017, April). A survey on hate speech detection using natural language processing. In *Proceedings of the fifth international workshop on natural language processing for social media* (pp. 1-10).

5. Przybyla, P. (2020, April). Capturing the style of fake news. In *Proceedings of the AAAI conference on artificial intelligence.* (Vol. 34, No. 01), (pp. 490-497).

6. L. Mishchenko, I. Klymenko. (February, 2023). Method For Detecting Fake News Based On Natural Language Processing. *The VI International Scientific and Practical Conference «Modern ways of solving the problems of science in the world».* (pp. 375-378). Warsaw, Poland.

7. Caihong, Z. (2022). Quantitative Evaluation of Vocabulary Emotional Color in Language Teaching. *Occupational Therapy International.* https://www.hindawi.com/journals/oti/2022/5203122/.

8. Pathak, A., & Srihari, R. K. (2019, July). BREAKING! presenting fake news corpus for automated fact checking. In *Proceedings of the 57th annual meeting of the association for computational linguistics: student research workshop*. (pp. 357-362).

9. Zhou, X., Li, J., Li, Q., & Zafarani, R. (2023). *Linguistic-style-aware Neural Networks for Fake News Detection.* arXiv preprint arXiv:2301.02792.

10. Xu, S. (2018). Bayesian Naïve Bayes classifiers to text classification. *Journal of Information Science, 44(1),* 48-59.

Отримано 01.12.2023

УДК 004.02

Людмила Міщенко[1], Ірина Клименко[2],

[1]аспірантка кафедри обчислювальної техніки
Національний технічний університет України "Київський політехнічний інститут ім. Ігоря Сікорського", (Київ, Україна)
E-mail: liudamishchenko@gmail.com. **ORCID:** https://orcid.org/0000-0002-5101-2517

[2]професор, доктор технічних наук кафедри обчислювальної техніки
Національний технічний університет України "Київський політехнічний інститут ім. Ігоря Сікорського", (Київ, Україна)
E-mail: ikliryna@gmail.com. **ORCID:** http://orcid.org/0000-0001-5345-8806

МЕТОД ВИЯВЛЕННЯ ФЕЙКОВИХ НОВИН ЗА СТИЛЕМ НАПИСАННЯ

У сучасну цифрову епоху відрізнити правду від дезінформації є складною проблемою, а фейкові новини, становлять значну загрозу можливості маніпуляції думкою суспільства. Традиційні методи перевірки фактів рідко звертають увагу на емоційне мовне забарвлення, властивими оманливому вмісту. Це дослідження націлене на розробку підходу до виявлення фейкових новин шляхом аналізу стилю написання у поєднанні з обробкою природної мови.

Базуючись на різноманітному наборі даних, у дослідженні розглянуто різні лінгвістичні аспекти, з особливим акцентом на попередній обробці та розробці функцій. На відміну від традиційних методів, цей багатовимірний підхід спрямований на розкриття нюансів у стилі написання як справжніх, так і фейкових новинних статей.

Дослідження передбачає ретельну підготовку та обробку вхідного набору даних, розробку функцій, зосереджуючись на ключових лінгвістичних елементах, таких як емоційно забарвлена лексика, звітні слова, ймовірнісні слова, легкі лайки та нецензурна лексика. Завдяки ретельному аналізу таких груп слів дослідження спрямоване на виявлення характерних моделей стилю написання фейкових новин, сприяючи розробці надійних механізмів їх виявлення.

Процес оцінювання ретельно перевіряє точність розпізнавання справжніх та фейкові статті новин запропонованим способом на основі визначених стилів написання. Результати підтверджують доволі високу точність підходу та здатність виявляти відмінності в лінгвістичних особливостях написання новин.

Це дослідження впливає на підвищення точності виявлення фейкових новин за допомогою аналізу стилю написання. Поєднуючи ретельну перевірку набору даних, передові методи попередньої обробки та комплексну розробку функцій, дослідження сприяє боротьбі з дезінформацією в епоху цифрових технологій. Такий підхід доповнює уже існуючі методи та робить їх більш комплексним для розпізнавання фейкових новин. Це демонструє його потенціал для підвищення медіаграмотності, захисту демократичних процесів, коли дезінформація становить серйозну проблему для цілісності інформації.

Ключові слова: *фейкові новини; справжні новини; аналіз стилю письма; емоційно забарвлені слова; емоційна мова; ймовірнісні слова; виявлення ненормативної лексики.*

Табл.: 2. Рис.: 1. Бібл.: 10.

Mishchenko, L., Klymenko, I. (2023). Method for detecting fake news through writing style. *Technical sciences and technologies*, 4(34), 82-90.