architecture addresses critical limitations such as restricted memory capacity, lack of hardware cryptographic support, and the necessity to minimize energy consumption. Classical secure boot models are largely unsuitable for these conditions; therefore, a lightweight Secure Boot variant has been developed using lightweight hash functions, in particular SPONGENT-128/128/8, which ensures a basic level of software authenticity and integrity at minimal computational cost. Implementation on STM32 and ESP8266 microcontrollers demonstrated high accuracy in detecting firmware modifications, energy efficiency (less than 11 μJ per verification), and a startup delay not exceeding 30 ms. The study discusses secure storage of the reference hash, the use of ECDSA-based digital signatures, and scaling options for different hardware configurations. The proposed approach is promising for integration into medical, industrial, military, and other applied IoT solutions where balancing security, energy efficiency, and performance is crucial. The conclusion outlines directions for future research, including obfuscation techniques, secure OTA updates, and support for emerging cryptographic primitives.

Keywords. *IoT; lightweight Secure Boot; embedded software; cryptographic hash; SPONGENT; energy efficiency*

UDC 004.8
DOI https://doi.org/10.36059/978-966-397-538-2-7

METHOD FOR ANALYZING THE UKRAINIAN LANGUAGE TEXTS SENTIMENT USING NATURAL LANGUAGE PROCESSING

O. Zalutska[0000-0003-1242-3548]
Khmelnytskyi National University, Ukraine
EMAIL: zalutsk.olha@gmail.com

Abstract. The paper focuses on intelligent sentiment analysis of text related to named entities. The proposed method combines a neural network-based natural language processing model, a lexical NLP library, and a Ukrainian sentiment dictionary. It provides results in the form of sentiment scores for named entities at the sentence and text levels, as well as an overall sentiment evaluation of the analyzed content. The relevance of the research is determined by the growing need for accurate sentiment analysis in the context of large-scale digital information flows. Identifying emotional attitudes toward specific persons, organisations, or events has essential applications in monitoring public opinion, brand perception, political discourse, and financial market analysis. The scientific novelty lies in developing and implementing a method that supports Ukrainian-language texts and evaluates sentiment across negativity, neutrality, positivity, and emotionality dimensions. The practical significance is creating a software system capable of semantic sentiment analysis of textual content, achieving higher effectiveness than translation-based approaches. The developed method can analyze public opinion, social media reactions, market trends, and individual texts.

Keywords: *named entity recognition, emotional tone, emotional tone detection, Stanza, VADER*

Problem Statement

Intelligent sentiment analysis of textual information, particularly in the context of named entities, has become increasingly relevant today, where vast amounts of data are constantly exchanged and processed. In the era of digital technologies and information overload, quickly and accurately identifying the emotional context and subjective attitudes toward specific named entities – such as individuals, organizations, or events – has gained crucial importance [1]. Applying these methods to named entities has numerous practical implications, including monitoring brand perception in reviews and mentions, identifying public sentiment toward political figures, events, or products in social media, and analyzing the impact of news on market indicators and financial risks. Moreover, sentiment analysis at the entity level not only reveals the general emotional tone of a text but also distinguishes attitudes toward specific entities, which may vary within a single document [2]. From a natural language processing perspective, sentiment analysis related to named entities requires the design of advanced algorithms and neural models capable of correctly interpreting linguistic but also semantic, contextual, and culturally specific features of language [3]. The study introduces innovations in sentiment analysis, particularly by enhancing the method of intelligent sentiment evaluation relating to named entities. The proposed approach generates comprehensive output data based on a neural network NLP model, a lexical NLP library, and a Ukrainian sentiment dictionary. This includes identified named entities, sentiment scores for each entity at the sentence level, aggregated sentiment scores for the entire text, and overall sentiment assessments at both sentence and text levels [4].

Unlike existing methods, the developed approach is adapted explicitly to Ukrainian texts and enables sentiment evaluation for named entities within individual sentences and across the whole text. Furthermore, it provides sentiment analysis along four key dimensions: negativity, neutrality, positivity, and emotionality.

Analysis of Previous Studies

Existing methods and tools for intelligent analysis of the tone of textual information in relation to named entities. The review of methods and tools for intelligent sentiment analysis concerning named entities demonstrates that this task requires a combination of advanced natural language processing techniques and machine learning models. The process is generally divided into three main stages:

– named entity recognition, where algorithms identify and classify persons, organizations, locations, and other entities in text;

– sentiment determination, which analyzes the local context of each entity to detect its emotional tone;

– contextual analysis considers the document's broader semantic and thematic background to ensure a more accurate interpretation of sentiment [5].

To implement these stages, various approaches are employed, such as Bidirectional Encoder Representations from Transformers (BERT), Long Short-Term Memory (LSTM) networks, Conditional Random Fields (CRFs), and NLP libraries like spaCy and NLTK. These methods are among the most powerful and widely used entity detection and sentiment analysis tools. In addition, lexicon-based and hybrid approaches such as VADER, TextBlob, and Google Cloud Natural Language API are actively applied for sentiment evaluation. However, most of these solutions are primarily optimized for English, which limits their effectiveness for Ukrainian texts due to linguistic and cultural differences [6].

The analysis shows that while lexicon-based tools such as VADER or TextBlob can effectively detect sentiment in short and informal texts, they often require adaptation or developing specialized resources for Ukrainian. On the other hand, deep learning models such as BERT and LSTM provide higher accuracy and adaptability by capturing semantic and contextual dependencies. However, they demand large annotated datasets and significant computational resources [7].

The study highlights traditional and state-of-the-art approaches to entity-level sentiment analysis, outlining their strengths and limitations. It also emphasizes the growing relevance of neural network-based methods, which offer more profound and precise sentiment interpretation. In the modern information environment, the ability to automatically determine attitudes toward specific named entities is becoming increasingly important for applications such as monitoring public opinion, analyzing social media reactions, studying consumer behavior, and evaluating market trends [8].

Research on the analysis of scientific publications in the intellectual study of text tone in relation to named entities. The review focused on several key aspects of sentiment analysis methods: the type of algorithm, the use of supervised or unsupervised learning, the specific models applied, characteristics of datasets including size and collection methods, vectorization techniques, and evaluation metrics. Table 1 summarizes the findings from selected studies, highlighting the predominance of supervised learning methods, while unsupervised approaches were used much less frequently.

The comparative analysis of the reviewed studies highlights the predominance of supervised learning methods in sentiment analysis, as most authors relied on these approaches [8, 9, 10, 11, 12]. In contrast, unsupervised methods were applied only occasionally in a single study [10]. A wide range of algorithms has been explored, including classical machine learning models such as logistic regression (LR), support vector machines (SVM), random forest classifiers (RFC), and naïve Bayes, as well as deep learning models like CNNs, LSTMs, and advanced transformer-based architectures such as BERT, RoBERTa, XLNet, ALBERT, BART, and DistilBERT [8, 9, 10, 11, 12]. Simpler lexicons- and rule-based tools like TextBlob, VADER, and Stanza were also tested, demonstrating their relevance for rapid sentiment detection in short texts [9].

Table 1

Results of research scientific papers

№ in the References	Methods	Training type	Metrics	Vectorisation
8	LR, SVM, RFC, NBC	Supervised	Accuracy, Precision, Recall, F1-score	TF-IDF
9	RoBERTa	Supervised	Accuracy, Precision, Recall, F1-score	-
10	TextBlob, VADER, Stanza	Both	-	-
11	SVM, XGBoost, RNN, RNN + Attention, CNN, Dense Network, BERT, XLNet, RoBERTa, ALBERT, BART, DistilBERT	Supervised	Accuracy, Precision, Recall, F1-score, MCC	BoW, TF-IDF, Harvard IV-4, LM, Word2Vec, GloVe, 23FastText, ELMo, Doc2Vec, Skip-Thought Vectors, InferSent
12	Ensemble of LR, RC, and Weightless Neural Network (WNN)	Supervised	F1-score	TF-IDF

Evaluation across studies was mainly based on accuracy, precision, recall, and F1-score, with some works also employing the Matthews correlation coefficient (MCC) for a more robust performance measure [11]. Different vectorization techniques were adopted, ranging from traditional bag-of-words (BoW) and TF-IDF to semantic embeddings like Word2Vec, GloVe, FastText, ELMo, InferSent, and Skip-Thought Vectors [10, 11]. Notably, [10] focused on unigram and bigram TF-IDF features, while [11] emphasized emotion recognition in dialogue systems, a domain increasingly relevant for conversational agents. Dialogue-related approaches were divided into party-dependent models (e.g., DialogueRNN) and party-independent ones (e.g., AGHMN), with additional frameworks like BiERU, GNTB, and TFE proposed to capture high-quality emotional features [9]. Ensemble methods, such as combining LR, RC, and weightless neural networks, were also investigated to enhance robustness [12].

These studies reveal a trend toward leveraging deep neural networks and transformer-based architectures for large-scale sentiment analysis, highlighting the complementary role of simpler models and lexicon-based tools. Integrating diverse vectorization techniques and advanced architectures opens new possibilities for capturing subtle emotional nuances in text [13].

The analysis revealed various algorithms, including Logistic Regression, SVM, Random Forest, CNN, LSTM, RoBERTa, TextBlob, and VADER, with particular attention given to SVM due to its frequent application. Evaluation metrics most commonly employed were accuracy, precision, recall, and F1-score, while advanced studies also incorporated MCC. Vectorization methods varied from classical BoW and TF-IDF to modern embeddings such as Word2Vec, GloVe, FastText, ELMo, and BERT-based encodings [14].

Unresolved Issues

The studies illustrate that sentiment analysis of textual data is an intensively researched domain with diverse approaches. Research efforts are primarily directed toward improving the processing of massive text datasets through advanced machine learning and deep neural networks, capable of automatically learning complex dependencies between text and its emotional expression. These innovations pave the way for more accurate sentiment analysis and a deeper understanding of emotional dynamics in textual content.

Objective of the Article

The object of research is the process of determining the tone of textual information related to named entities. The subject of research is methods, algorithms, information technologies, models, and tools for determining the tone of textual information related to named entities.

Method of intelligent analysis of the tone of textual information in relation to named entities

The methodology of intellectual analysis of texts' emotional tome, especially in determining attitudes toward named entities, plays a significant role in decision-making in various fields, including commerce, media, finance, and politics. This field of research opens up broad prospects for developing and improving text analysis algorithms that take named entities into account [15].

Figure 1 shows the stages of the method of intelligent analysis of the tone of textual information in relation to named entities, which allows the selected text under study, using a neural network model of natural language processing, a lexical library for natural language processing, and a Ukrainian-language tonal dictionary, to obtain output data in the form of a conclusion about the tonality of the text under study, which includes a set of named entities, a numerical assessment of tonality in relation to each of the named entities for individual sentences; the value of the overall tone assessment for the entire text in regard to each of the named entities, the value of the tone assessment of the text for individual sentences of the text, and the value of the overall tone assessment for the entire text.

The main goal of the proposed approach is to identify positive, negative, or neutral sentiment in a text and to establish links between these sentiments and the specific named entities mentioned. This allows for determining how textual information is perceived concerning particular entities. The input data consists of textual posts describing events, activities, or individuals, primarily sourced from

"Ukrainska Pravda" [12]. While the method can handle texts of any length, processing extensive texts may reduce efficiency and speed, requiring more computational resources. Significant texts can be divided into smaller segments or sentences before being fed into the model to optimise performance.

A central component of the method is a neural language model responsible for assessing the emotional tone of the text. For this purpose, the Stanza model was selected [16]. In the first step, named entities are extracted using the neural model. The model identifies all mentions of an entity, including variations due to inflexion or case. The second step applies lemmatisation and stemming to consolidate these mentions. Lemmatization reduces words to their base form, often using Part of Speech (POS) tags to account for context and distinguish between words with different meanings [17]. For example, accurately extracting a person's name requires differentiating it from other words in the text, which can be achieved using NLP libraries such as NLTK, spaCy, or TextBlob. During the second step, variations of the same entity (e.g., «Петренка», «Петренко», «Петренкові») are grouped under the nominative form for sentiment evaluation.

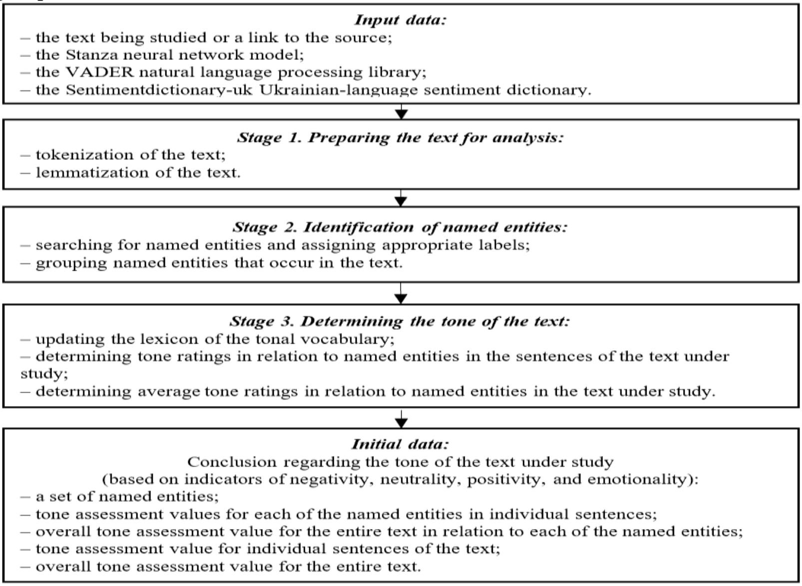

Figure 1. Stages of the method of intellectual analysis of the tone of textual information in relation to named entities

The third step involves assessing text sentiment regarding each named entity using the VADER approach [18]. VADER (Valence Aware Dictionary for sEntiment Reasoning) is an NLP algorithm that combines a lexicon-based approach with grammatical and syntactic rules to determine polarity and sentiment intensity. Its lexicon contains approximately 7,500 words, phrases, emoticons, and abbreviations, each rated on a scale from -4 (extremely negative) to +4 (extremely positive) by 10 independent experts. Words not present in the lexicon are scored as 0 (neutral).

VADER also calculates a compound metric, a normalized indicator of overall sentiment ranging from -1 (strongly negative) to +1 (strongly positive), with values near 0 reflecting neutrality or mixed emotions. This metric provides a single numeric representation of the overall sentiment of the text, incorporating the intensity of individual words [19].

Table 2 illustrates an example of how sentiment scores are computed based on the writing style of the text, demonstrating the method's ability to capture nuanced emotional tones related to named entities.

Table 2

An example of calculating the emotional tone of a text, depending on the style of writing the message

Input text	Negative level	Neutral level	Positive level	Compound
Цей комп'ютер був гарною покупкою.	0	0.58	0.42	0.44
Цей комп'ютер був ***дуже*** *гарною покупкою.*	0	0.61	0.39	0.49
*Цей комп'ютер був дуже гарною покупкою****!!***	0	0.57	0.43	0.58
Цей комп'ютер був дуже гарною покупкою!! ***:)***	0	0.44	0.56	0.74
Цей комп'ютер був ***ДУЖЕ*** *гарною покупкою****!! :)***	0	0.393	0.61	0.82

To calculate the compound, VADER scans the text for known sentiment features, modifies the intensity and polarity according to the rules, sums up the ratings of the features found in the text, and normalizes the final rating to (-1, 1) using the following function (1):

$$compound = \frac{x}{\sqrt{x^2 + \alpha}}, \tag{1}$$

where x is the sum of the scores for all words in the text that have an emotional connotation (each word has a predefined score according to its positive or negative connotation); α is a parameter used to normalize the sum of the scores of word x (in VADER, it is set to 15, which is considered to approximate the maximum expected value of x). The function normalizes the sum of scores to not exceed the scale from -1 to 1. This makes sentiment assessment more comparable between texts of different lengths and emotional intensity [20].

Step 4 involves determining the tone based on the values obtained in the previous step using the corresponding scale, where compound >= 0.5 is negative, compound > -0.5 or compound < 0.5 is neutral, and compound <=-0.5 is positive.

The last step of the method is to form a conclusion about the tonality of the text post in relation to named entities. The process can return the emotional tonality value for each sentence and an overall assessment of the named entity. For example, a person or organization may appear several times in the text and have a different tone in different sentences.

The output data of the method is a conclusion about the tonal state of the text in relation to the named entity. It should be noted that if a person, event, or place appears several times in the text, the method will return a value for the tone for each sentence separately, as well as an overall assessment in the context of the entire text.

Thus, intelligent analysis of the emotional tone of text, focusing on named entities, is an essential tool for automating text data processing, which contributes to the formation of more reasonable and informed decisions based on in-depth analysis of extensive text databases. As a result of this section, a method for the intelligent analysis of the tone of textual information in relation to named entities was developed, and the steps of its operation were described in detail. To implement the intelligent analysis of the tone of textual information in relation to named entities, it is necessary to retrain VADER on a Ukrainian-language dataset, which will allow determining sentiment in Ukrainian-language posts without resorting to machine translation tools. The corresponding dataset should be marked as follows: word; discrete sentiment (from the range: -2, -1, 0, 1, 2).

The Ukrainian tonal dictionary [21] sentimentdictionary-uk [22] was selected. The Ukrainian tonal dictionary contains an extensive database: 3442 words of the Ukrainian language with explicitly defined emotional tonality. Each word has an assigned tonality value ranging from -2 to 2. This dictionary was constructed based on two key sources:

– the tone-dict-uk-manual.tsv file was formed by averaging the ratings provided by several experts. This process ensured an objective assessment of the emotional tone of each word.

– the tone-dict-uk-auto.tsv file was created by automatically expanding the base tone-dict-uk-manual dictionary. This was done using machine learning methods that apply vector representations of words, such as word2vec and lex2vec. This process also included human post-processing to ensure data accuracy and consistency.

The data in the dictionary is organised in a format that includes two columns separated by a tab: the first column contains the word, and the second contains its discrete tone. It is important to note that most words in the dictionary are given in their basic grammatical form. In addition, where possible, adverbs have been replaced with cognate adjectives to ensure greater analysis accuracy. The Ukrainian sentiment dictionary sentimentdictionary-uk contains 14,856 words and phrases presented in the Dictionary. Creating a specific sentiment dictionary for a particular language is a key aspect in developing sentiment analysis systems. For the Ukrainian language, there is only one freely available tonal dictionary (the lang-uk project) with 6,859 words, which was last updated in September 2016. Given the dynamics of language and changes in word usage, there is a need to develop a new tonal dictionary that reflects contemporary word usage. This need created a new dictionary, “tonSUM.2.0,” available in various formats.

Explanation of the dictionary layout: column 1 contains the word or phrase (Word//word combination); columns 2 to 9 contain the sentiment scores for the word or phrase listed in column 1; the 10th column shows the average sentiment score, which is calculated automatically (using built-in functions) by dividing the sum of the specified scores by their number. Figure 2 shows the distribution of words and phrases in the Ukrainian Sentiment Dictionary dataset. Score -2: 874 reviews, score -1: 1370 reviews, score 1: 1081 reviews, score 2: 117 reviews.

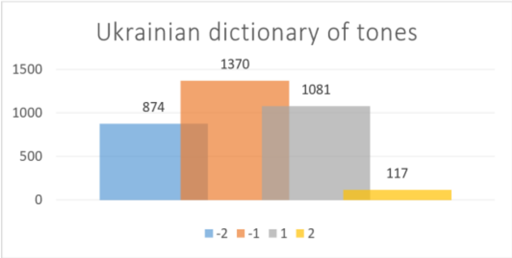

Figure 2. Distribution of words and phrases in the Ukrainian tonal dictionary dataset

The sentimentdictionary-uk dictionary contains a significantly larger amount of data. Figure 3 shows the distribution of words and phrases in the sentimentdictionary-uk dataset. The dataset contains 14,855 records: 1,006 reviews with a rating of -2, 2,093 reviews with a rating of -1, 5,093 reviews with a rating of 0, 5,844 reviews with a rating of 1, and 819 reviews with a rating of 2.

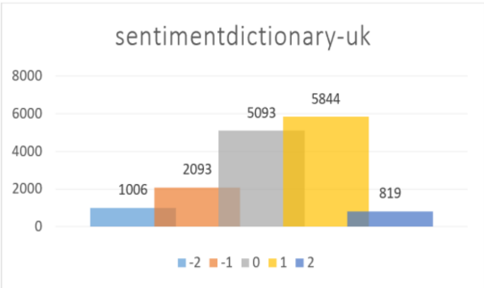

Figure 3. Distribution of words and phrases in the dataset “sentimentdictionary-uk”

Thus, combining data from datasets, a sample was obtained to supplement the VADER dictionary, containing 18,297 entries.

Approach to verifying library training for natural language processing for text sentiment analysis. A dataset containing 7,656 records was formed, of which 6,655 were used for training and 1,331 (about 20% of the training set) for validation. This set is characterized by the presence of Russian language and slang and Russian-language elements in 37% of Ukrainian-language documents. This reflects specific linguistic dynamics in social networks after the start of the war. There is also a high frequency of spelling errors in the reviews.

The dataset obtained earlier contains binary ratings: recommend/do not recommend, and to validate VADER, it is necessary to check whether the method returns the correct "score" rating. Therefore, after downloading the validation sample, it will be checked whether the method returns a score >= 0.5 and the validation sample label "1" (recommended) – the test is considered passed. Accordingly, if the method returns a score <= -0.5 and the validation sample label "0" (do not recommend), the check is considered successful.

Experiment

To study the effectiveness of the method of intelligent analysis of the tone of textual information in relation to named entities, it is necessary to compare the results of the information system with automatic machine translation and subsequent determination of the tone of textual information in the text.

First, the software product was tested based on the intelligent determination of the tone of textual information in relation to named entities using the input parameter – an article from the resource "Ukrainska Pravda". The following sentence was entered for the study: "*Петро Василенко – безжальний вбивця! Хто б міг подумати, що виконавець пісні Мамині світлиці такий жорстокий!*". The program for determining the emotional tone of named entities, which is configured to work with English-language texts, returns the following results. Translated sentence: *"Petro Vasilenko is a ruthless killer! Who would have thought that the performer of the song of the mother's room is so cruel!"*. For comparison, the exact text was analyzed using Stanza and VADER, which had been pre-populated with tone dictionaries for the Ukrainian language. The method returned the following results. The result of running the program code to test the method of determining the emotional tone of text in relation to named entities for Ukrainian-language texts is shown in Figure 4. Using empirical approaches, a sample of data was collected to compare the performance of the methods (Table 3).

```
INFO:stanza:Using device: cpu
INFO:stanza:Loading: tokenize
INFO:stanza:Loading: mwt
INFO:stanza:Loading: lemma
INFO:stanza:Loading: ner
INFO:stanza:Done loading processors!
Entity: Петро Василенко, Type: PERS
Entity: Мамині, Type: PERS
Петро василенко – безжальний вбивець ! хто б могти подумати , що виконавець пісня мамині світлиця такий жорстокий !
Sentiment: {'neg': 0.533, 'neu': 0.349, 'pos': 0.118, 'compound': -0.8885}
```

Figure 4. Result of executing the program code for determining the emotional tone of the text in relation to named entities for Ukrainian-language texts

Table 3

An example of calculating the emotional tone of a text, depending on the style of writing the message

Input text		Negative level	Neutral level	Positive level	Compound
Нарешті жителям нашої громади покращили прибудинкові території! Тепер дітки можуть гратись на сучасних майданчиках, а мами можуть не хвилюватись за їх безпеку! Дякуємо!	Proposed method	0.00	0.744	0.256	0.6757
	Machine translation	0.00	0.886	0.114	0.4754
«Російський терорист не знає меж! Путін ніколи не зупиниться, українцям потрібно готуватись до тривалої війни», - коментар Мельника	Proposed method	0.411	0.519	0.07	-0.8808
	Machine translation	0.219	0.638	0.143	-0.3964
42-річний житель Хмельницького району, що вже відсидів свій строк за вбивство, сяде за вбивство вчергове. Підсудний вбив товариша по чарці, прийнявши його за російського солдата.	Proposed method	0.333	0.528	0.139	-0.8779
	Machine translation	0.292	0.601	0.107	-0.7068

To justify the superiority of the implemented method, experiments were conducted in which both methods (proprietary and machine translation-based) were compared on the same dataset. It is important to note that the implemented method shows a more distinct and precise distribution of sentiments (negative, neutral, positive), which indicates its greater sensitivity to the emotional nuances of the text.

The compound value indicates the overall sentiment of the text. The proprietary method gives a more accurate overall value that corresponds to the expected sentiment of the text, which indicates its advantage.

A heat map (Figure 5) and a bar chart (Figure 6) were used to visualize the results. Each row represents a sentiment analysis method (the implemented method and machine translation), and the columns reflect different aspects of sentiment: negativity, neutrality, positivity, and the overall compound indicator.

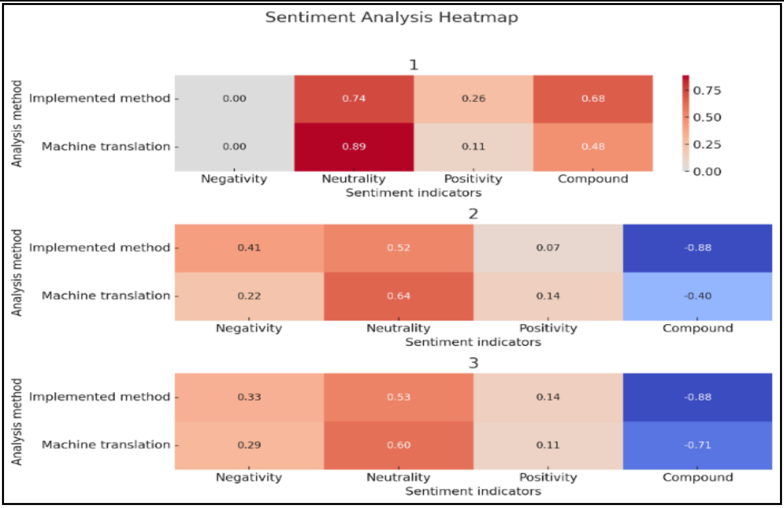

Figure 5. Heat map based on the obtained results

Based on data analysis and visualization using a heat map and bar chart, it can be concluded that the implemented method was more effective than using machine translation tools, particularly in determining the tone of textual information relating to named entities. In particular, our method demonstrated higher positive tone and compound values in the first text, indicating its ability to recognise positive nuances better. At the same time, for texts with negative sentiment, our method also showed a greater ability to identify negative emotions, as evidenced by lower (more negative) compound values compared to machine translation. This ability to more accurately determine sentiment is essential when analyzing context related to named entities, as it requires a deeper understanding of linguistic nuances and culturally specific expressions that machine translation is often unable to interpret correctly.

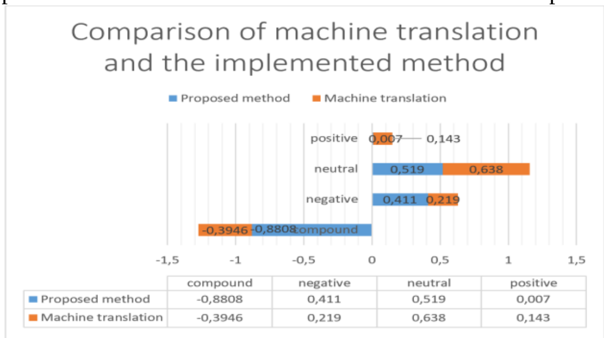

Figure 6. Distribution diagram of results

Overall, the implemented method showed better results in analyzing the tone of texts, making it a more reliable tool for sentiment analysis, especially in cases where contextual dependence and language specificity must be considered.

Conclusions and Prospects

The study addresses the problem of determining the sentiment of textual content ccerning named entities. A method for intelligent sentiment analysis of text information in relation to named entities has been developed, which enables, for a given text, the identification of key named entities and the assessment of sentiment toward each entity. The method combines a neural network-based natural language processing model, a lexical processing library, and a Ukrainian sentiment lexicon to produce outputs that include: a set of named entities, sentiment scores per entity for individual sentences, overall sentiment scores per entity for the entire text, sentiment scores for individual sentences, and overall text sentiment scores.

A corresponding software implementation was created to validate the proposed method. The study conducted the following tasks: a review of the current state of intelligent sentiment analysis regarding named entities, development of the sentiment analysis method, implementation of the process in software, and evaluation of its practical effectiveness. The results demonstrate the scientific novelty and innovation of the developed method. It allows for the analysis of Ukrainian-language texts. It provides sentiment assessment at both the sentence and text levels for each named entity, measuring negativity, neutrality, positivity, and emotionality.

Comparative evaluation indicates that the proposed method is more effective than approaches relying on translation into English followed by sentiment assessment. Overall, the developed approach can be applied to monitor public opinion and direct semantic analysis of individual texts, providing a reliable tool for sentiment assessment concerning named entities in Ukrainian-language content.

References

1. Taherdoost, H., & Madanchian, M. (2023). Artificial intelligence and sentiment analysis: A review in competitive research. *Computers, 12*(2), 37. https://doi.org/10.3390/computers12020037

2. Mazurets, O., Sobko, O., Dydo, R., Zalutska, O., & Molchanova, M. (2025). Transformer-based multilabel classification for identifying hidden psychological conditions in online posts. *CEUR Workshop Proceedings, 4004*, 347–361. https://ceur-ws.org/Vol-4004/paper26.pdf

3. Molchanova, M., Didur, V., Sobko, O., & Mazurets, O. (2025). Detection of web propaganda patterns by transformer neural networks: Improving efficiency via dataset balancing. *CEUR Workshop Proceedings, 3988*, 112–126. https://ceur-ws.org/Vol-3988/paper10.pdf

4. Murava, V., Zalutska, O., Didur, V., & Mazurets, O. (2025, June 25–27). Software architecture of information system for exchanging LLM thematic prompts. In *Proceedings of the IV International Scientific and Practical Conference Global Trends in the Development of Information Technology and Science* (pp. 121–127).

Stockholm, Sweden. https://isu-conference.com/wp-content/uploads/2025/06/Stockholm_Sweden_25.06.25.pdf

5. Andrushchenko, D., Klimenko, V., & Mazurets, O. (2025, May 28–30). Vector databases search for adaptive filtering of scientific articles. In *Proceedings of the XXII International Scientific and Practical Conference Scientific Trends in the Development of Modern Technologies* (pp. 189–195). Krakow, Poland. https://eu-conf.com/wp-content/uploads/2025/05/SCIENTIFIC-TRENDS-IN-THE-DEVELOPMENT-OF-MODERN-TECHNOLOGIES.pdf

6. Mazurets, O., & Ovcharuk, O. (2024, November 21). Efficiency research of method for detecting mental disorders by analysis of user content. In *Proceedings of the 11th International Conference on Information Technology Implementation (Satellite)* (pp. 46–47). Kyiv, Ukraine.

7. Mazurets, O., & Vit, R. (2024, November 21). Practical application of method of thematic classification of text information using LDA. In *Proceedings of the 11th International Conference on Information Technology Implementation (Satellite)* (pp. 151–152). Kyiv, Ukraine.

8. Mazurets, O., Tymofiiev, I., & Dydo, R. (2024, November 15). Approach for using neural network BERT-GPT2 dual transformer architecture for detecting persons depressive state. In *Proceedings of the VI International Scientific and Practical Conference Ricerche scientifiche e metodi della loro realizzazione* (pp. 147–151). Bologna, Italy. https://archive.logos-science.com/index.php/conference-proceedings/issue/view/29

9. Rendón-Cardona, P., Gil-Gonzalez, J., Páez-Valdez, J., & Rivera-Henao, M. (2022). Self-supervised sentiment analysis in Spanish to understand the university narrative of the Colombian conflict. *Applied Sciences, 12*(11), 5472. https://doi.org/10.3390/app12115472

10. Li, W., Shao, W., Ji, S., & Cambria, E. (2022). BiERU: Bidirectional emotional recurrent unit for conversational sentiment analysis. *Neurocomputing, 467*, 73–82. https://doi.org/10.1016/j.neucom.2021.09.057

11. Mohan, S., Solanki, A. K., Taluja, H. K., Anuradha, & Singh, A. (2022). Predicting the impact of the third wave of COVID-19 in India using hybrid statistical machine learning models: A time series forecasting and sentiment analysis approach. *Computers in Biology and Medicine, 146*, 105354. https://doi.org/10.1016/ j.compbiomed.2022.105354

12. Tymofiiev, I., Mazurets, O., Hardysh, D., & Molchanova, M. (2024, November 6–8). Neural network dual architecture for depression detection using cloud services. In *Proceedings of the XLVI International Scientific and Practical Conference Scientific Research in the Era of Digital Technologies: Challenges and Opportunities* (pp. 84–88). Barcelona, Spain. https://isu-conference.com/wp-content/uploads/2024/11/Scientific_research_in_the_era_of_digital_technologies_challenges_and_opportunities_November_6_8_2024_Barcelona_Spain.pdf

13. Dharma, E., & Muntina, E. (2022). The accuracy comparison among word2vec, GloVe, and fastText towards convolution neural network (CNN) text

classification. *Journal of Theoretical and Applied Information Technology, 100*(2), 54–64. http://www.jatit.org/volumes/Vol100No2/5Vol100No2.pdf

14. Yurchenko, D., Mazurets, O., Didur, V., & Molchanova, M. (2024, November 13–15). Approach to using cloud services for visual analytics of neural network analysis of texts emotional tonality. In *Proceedings of the XLVII International Scientific and Practical Conference The Future of Scientific Discoveries: New Trends and Technologies* (pp. 108–113). Marseille, France. https://isu-conference.com/wp-content/uploads/2024/11/The_future_of_ scientific_discoveries_new_trends_and_technologies_November_13_15_2024_Marseille_France.pdf

15. Hladun, O., Mazurets, O., Molchanova, M., & Sobko, O. (2024, November 25–27). Real time detection the person emotion state using neural network. In *Proceedings of the II International Scientific and Practical Conference Scientific Research: Modern Innovations and Future Perspectives* (pp. 119–123). Montreal, Canada. https://www.eoss-conf.com/wp-content/uploads/2024/11/Montreal_Canada_25.11.24.pdf

16. Ukrainska Pravda. (n.d.). Retrieved from https://www.pravda.com.ua

17. Blazhuk, V., Mazurets, O., & Zalutska, O. (2024, October 23–25). An approach to using the mBERT deep learning neural network model for identifying emotional components and communication intentions. In *Proceedings of the XLIV International Scientific and Practical Conference The Impact of Scientific Research on the Development of the Modern World* (pp. 79–84). Dubrovnik, Croatia. https://isu-conference.com/wp-content/uploads/2024/10/The_impact_of_scientific_research_on_the_development_of_the_modern_world_October_23_25_2024_Dubrovnik_Croatia.pdf

18. Mazurets, O., Molchanova, M., Klimenko, V., & Prosvitliuk, M. (2024, June 12–14). Practice implementation of neural network model BART-Large-CNN for text annotation. In *Proceedings of the XXVII International Scientific and Practical Conference Prospects of Scientific Research in the Conditions of the Modern World* (pp. 97–102). Rotterdam, Netherlands. https://isu-conference.com/wp-content/uploads/2024/06/Prospects_of_scientific_research_in_the_conditions_of_the_modern_world_June_12_14_2024_Rotterdam_Netherlands.pdf

19. Sobko, O., Mazurets, O., Didur, V., & Chervonchuk, I. (2024, June 5–7). Recurrent neural network model architecture for detecting a tendency to atypical behavior of individuals by text posts. In *Proceedings of the XXVI International Scientific and Practical Conference Theoretical and Practical Aspects of Modern Research* (pp. 113–117). Ottawa, Canada. https://isu-conference.com/wp-content/uploads/2024/06/Theoretical_and_practical_aspects_of_modern_research_June_5_7_2024_Ottawa_Canada.pdf

20. Mazurets, O., Sobko, O., Molchanova, M., Zalutska, O., & Yurchak, A. (2024, May 31). Practical implementation of neural network method for stress features detection by social internet networks posts. In *Proceedings of the II International Scientific Theoretical Conference Global Science: Prospects and*

Innovations (pp. 160–167). Berlin, Germany. https://previous.scientia.report/index.php/archive/issue/view/31.05.2024/89

21. Lang-uk. (n.d.). *Ukrainian tonal dictionary*. GitHub. https://github.com/lang-uk/tone-dict-uk

22. Sentimentdictionary-uk. (n.d.). GitHub. https://github.com/Oksana504/sentimentdictionary-uk

23. Abiola, O., Abayomi-Alli, A., Tale, O. A., et al. (2023). Sentiment analysis of COVID-19 tweets from selected hashtags in Nigeria using VADER and Text Blob analyser. *Journal of Electrical Systems and Information Technology, 10*(5). https://doi.org/10.1186/s43067-023-00070-9

24. Bing, L. (2022). *Sentiment analysis and opinion mining*. Springer Nature. https://link.springer.com/book/10.1007/978-3-031-02145-9

МЕТОД АНАЛІЗУ ЕМОЦІЙНОЇ ТОНАЛЬНОСТІ УКРАЇНОМОВНИХ ТЕКСТІВ ІЗ ВИКОРИСТАННЯМ ЗАСОБІВ ОБРОБКИ ПРИРОДНОЇ МОВИ

О. Залуцька[0000-0003-1242-3548]
Хмельницький національний університетУкраїна
EMAIL: zalutsk.olha@gmail.com

Анотація. *Дослідження присвячене інтелектуальному аналізу емоційної тональності текстів, пов'язаних з іменованими сутностями. Запропонований метод поєднує модель обробки природної мови на основі нейронної мережі, лексичну бібліотеку NLP та україномовний словник термінів, що містять емоційну тональність Він надає результати у вигляді оцінок емоційного забарвлення іменованих сутностей на рівні речень і текстів, а також загальну оцінку емоційного тону проаналізованого контенту.Актуальність дослідження визначається зростаючою потребою в точному аналізі емоційної тональності в контексті великомасштабних цифрових інформаційних потоків. Визначення емоційного ставлення до конкретних осіб, організацій або подій має важливе значення для моніторингу громадської думки, сприйняття бренду, політичного дискурсу та аналізу фінансових ринків.Наукова новизна полягає у розробці та впровадженні методу, який підтримує тексти українською мовою та оцінює сентимент за такими вимірами, як негативність, нейтральність, позитивність та емоційність. Практичне значення полягає у створенні програмної системи, здатної здійснювати семантичний аналіз сентименту текстового контенту, досягаючи вищої ефективності, ніж підходи, засновані на машинному перекладі. Розроблений метод може аналізувати громадську думку, реакції в соціальних мережах, ринкові тенденції та окремі тексти.*

Ключові слова: *розпізнавання іменованих сутностей, емоційна тональність, визначення емоційної тональності, Stanza, VADER*