Theoretical and applied foundations for improving fake news detection systems based on the use of CNN neural networks

Mikolaj Karpinski[1,†], Dmytro Borovyk[2,†], Serhii Lienkov[3,†], Oleh Borovyk[4,*,†], Oleksandr Seliukov[5,†], and Iryna Panibrat[6,†]

[1] *Institute of Security and Computer Science University of the National Education Commission, 2 Podchorazych str, Krakow, 30-084 Poland*

[2] *Khmelnytskyi National University, Institute str., 11, Khmelnytskyi, 29016, Ukraine*

[3] *Military Institute, Taras Shevchenko National University of Kyiv, Zdanovska str. 81a, Kyiv, 03189, Ukraine*

[4] *Administration of the State Border Guard Service of Ukraine, Volodymyrska str., 26, Kyiv, 01001, Ukraine*

[5] *School of Aerospace Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, P.R. China; (O.S.)*

[6] *Bohdan Khmelnytskyi National Academy of the State Border Guard Service of Ukraine, Shevchenko str., 46, Khmelnytskyi, 29000, Ukraine*

Abstract

The article addresses the problem of developing an information system for fake news detection, presents the design of its architecture and describes its software implementation. It is substantiated that improving the effectiveness of fake news detection can be achieved by enhancing the method based on a multilayer convolutional neural network (CNN). The proposed enhancement involves the addition of a dropout layer, increasing the kernel size and modifying the activation function. The use of the PolitiFact and LIAR datasets for neural network training is justified. An information system implementing the proposed method has been developed. The effectiveness of the proposed method has been evaluated. The use of the TensorFlow classification model and Logistic Regression as baseline models for comparison is justified. The results demonstrate that the proposed method is generally more effective than the existing methods considered in this research.

Keywords

online social networks, fake news detection, artificial intelligence, neural network, information technologies, method, algorithm

1. Introduction

In the current context of countering the armed aggression of the Russian Federation, a critical issue is the detection of fake news deliberately disseminated on social media with the aim of destabilizing public opinion, spreading panic and inciting fear among the population. The spread of fake news on the Internet occurs at a faster rate than that of verified information, as people are naturally drawn to novel or sensational content and tend to share it without verifying its authenticity. Fake news affects individuals' daily lives, manipulates their thoughts and emotions, alters their beliefs, and may lead to poor decision-making. The primary motives behind the

PhD Workshop on Artificial Intelligence in Computer Science at 9th International Conference on Computational Linguistics and Intelligent Systems (CoLInS-2025), May 15–16, 2025, Kharkiv, Ukraine

[*] Corresponding author.

[†] These authors contributed equally.

mikolaj.karpinski@uken.krakow.pl (M. Karpinski); db07757@gmail.com (D. Borovyk); lenkov_s@ukr.net (S. Lienkov); bov_nadpsu@ukr.net (O. Borovyk); selukov.ov@hnuba.edu.ua (O. Seliukov); irynaborovyk2017@ukr.net (I. Panibrat)

0000-0002-8846-332X (M. Karpinski); 0000-0003-4715-0658 (D. Borovyk); 0000-0001-7689-239X (S. Lienkov); 0000-0003-3691-662X (O. Borovyk); 0000-0001-7979-3434 (O. Seliukov); 0000-0002-3209-9119 (I. Panibrat)

dissemination of fake news include financial gain, the incitement of hatred based on extremist motives, manipulation of public consciousness for political purposes, and the formation of biased opinions during electoral campaigns, among others [1].

Many individuals face difficulties distinguishing fake news from real news, regardless of gender, age or level of education. Identifying fake news is challenging because, as scientific studies indicate, the human ability to differentiate between true and false information is relatively limited, estimated at approximately 54% [2]. Since fake news has become a global challenge and a serious threat to democracy, the economy, and peaceful coexistence, various stakeholders including civil society organizations, journalists, politicians, and researchers are working to reduce the associated risks.

Therefore, the problem of fake news dissemination via online social media (OSM) is now a global issue, and the development of effective countermeasures is an urgent task. Addressing this challenge requires the development of models capable of detecting fake news and limiting its spread. Today, advanced information technologies, particularly artificial neural networks are actively employed to tackle the problem of fake news detection. These technologies enable fake news detection systems to automatically process vast amounts of information and identify potentially false content.

Artificial intelligence contributes to more accurate and efficient identification of fake news. The application of AI in this domain is a crucial step toward ensuring societal stability and alleviating fear and panic among the population.

2. Related works

At present, various validated approaches to fake news detection exist. One prominent approach is based on the application of different machine learning (ML) and deep learning (DL) algorithms. Another relies on sentiment analysis of news content and the examination of emotions expressed in user comments. Several additional approaches also merit attention, further analysis, and investigation, each demonstrating a varying degree of effectiveness depending on the dataset. Classical ML algorithms typically include logistic regression (LR), support vector machines (SVM), decision trees (DT), naive Bayes (NB), random forest (RF), XGBoost (XGB), and combinations thereof. Higher-level ML algorithms encompass convolutional neural networks (CNN), bidirectional long short-term memory networks (BiLSTM), bidirectional gated recurrent units (BiGRU), hybrid models such as CNN-BiLSTM and CNN-BiGRU, as well as ensemble approaches based on these techniques. Deep learning-based models include BERTbase and RoBERTabase. In [3], the authors provide a review of ML-based fake news detection approaches using two scenarios for word representation methods – statistical and context-independent. Furthermore, [1] presents a comparative evaluation of eight advanced ML models, including CNN, BiLSTM, BiGRU, CNN-BiLSTM, CNN-BiGRU, various hybrid models with two types of text representation (context-independent and context-aware embeddings), BERTbase and RoBERTabase.

Many studies on fake news detection in online social media (OSM) are based on one or several key features such as content, network diffusion, or user behavior [4, 5]. The analysis of user comments to assess attitudes toward news items can play a significant role in detecting fake news [6, 7] and can provide insights into the credibility of published news content [8, 9].

In [10], it is argued that user comments possess high discriminative value in detecting fake news, where sentiment [11] or emotion expression [12] plays a decisive role. According to [13], users tend to express emotions such as fear, disgust and surprise in response to fake news, whereas reactions to real news are more likely to involve anticipation, sadness, joy and trust. However, the authors of that study did not explore the extent to which emotions can effectively identify fake news. As noted in [14], novelty may be a critical component of fake news, significantly enhancing its potential for dissemination and acceptance within society. Most existing studies utilizing sentiment analysis focus on the emotional signals present in the content of fake news [15]. It is also common for users to use emojis instead of textual comments to convey their reactions to specific news items on online social media (OSM) platforms [16, 17].

In this context, deep learning (DL) techniques substantially contribute to the classification, prediction, and analysis of textual content [18], owing to their ability to learn effectively [18, 19] and to detect features and complex patterns [20]. Studies [21, 22] demonstrate that incorporating sentiment- and emotion-based features significantly improves the accuracy of fake news detection for most deep learning models when compared to using textual features alone.

Moreover, the authors of these studies suggest that sentiment analysis of news content and emotional analysis of user comments can be leveraged by social media platforms to combat the spread of fake news. However, implementing this approach presents challenges, particularly when dealing with imbalanced datasets. The authors' comparative analysis of alternative fake news detection approaches led to the conclusion that the methods discussed above are effective and promising, especially in terms of their potential to inform the development of new models with high detection accuracy across diverse datasets.

Therefore, the purpose of the article is to improve the method and develop a fake news detection system based on the optimization of the neural network architecture.

3. Materials and methods

To achieve the stated objective, it is considered appropriate to carry out the following tasks: enhancement of the fake news detection method; analysis of statistical indicators for evaluating the quality of fake news detection; software implementation of the fake news detection method based on neural network technologies; and evaluation of the effectiveness of the proposed fake news detection method.

3.1. Enhancement of the Fake News Detection Method

The proposed method for detecting fake news is based on neural network technologies, specifically a multilayer convolutional CNN neural network.

Based on the analysis of the aforementioned detection methods, it is proposed to use a five-layer convolutional neural network for implementing the author's method, along with an optimized network architecture. Figure 1 presents the initial structure of the neural network.

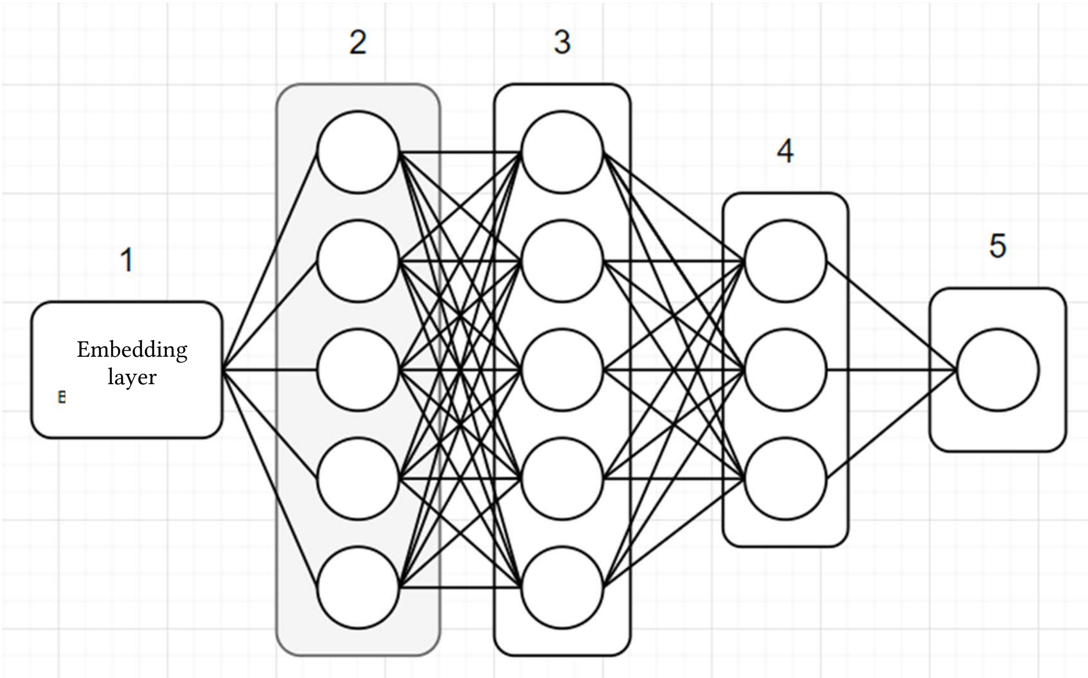

Figure 1: Initial Structure of the Neural Network.

This neural network consists of five layers:

1. Embedding Layer: this layer transforms the input textual data into a dense vector representation. It utilizes pre-trained word embeddings and sets the weights as non-trainable, meaning the embedding weights remain unchanged during model training. The input dimension of this layer corresponds to the vocabulary size of the analyzed text.
2. One-dimensional Convolutional Layer (Conv1D): A one-dimensional convolutional layer with 64 filters and a kernel size of 5. It is used to apply convolution to the input data. The ReLU (Rectified Linear Unit) activation function is used at this layer.
3. One-Dimensional Max Pooling Layer (MaxPooling1D): This layer is used to extract the most significant features from the input data (i.e., the output of the previous layer) and to reduce the dimensionality of the data.
4. LSTM Layer: This is a long short-term memory (LSTM) layer with 64 units. LSTM is a type of recurrent layer capable of capturing sequential information. In this model, it is used to process and model the textual data.
5. Dense Layer: This is a fully connected perceptron layer in which all neurons are connected to the neurons of the previous layer. It uses a sigmoid activation function and is responsible for producing the final binary classification output. In other words, this layer directly answers the question: “Is the analyzed news item true or fake?”

The described model is compiled using binary cross-entropy as the loss function and the Adam optimizer. It is designed for binary classification and the sigmoid activation function in the final layer enables the model to output probabilities for binary labels.

Upon analyzing the performance of the model, it was found that it achieved relatively high accuracy on the training datasets (91,2%). However, when the model was tasked with classifying previously unseen texts (i.e., data not included in the training set), its accuracy decreased to 86,6%. This revealed the necessity of optimizing the neural network architecture to improve the reliability of news classification on out-of-sample data.

The improved neural network architecture is illustrated in Figure 2.

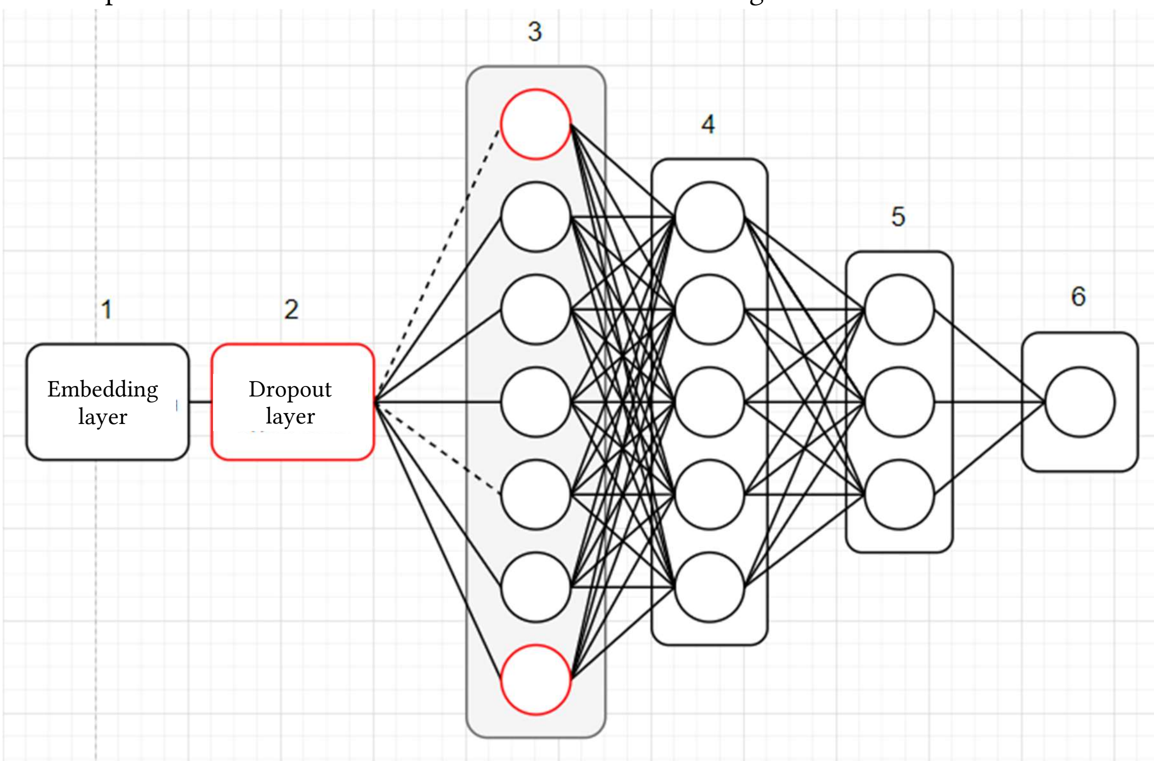

Figure 2: Improved Neural Network Architecture.

The improved architecture includes the following modification:

1. Dropout Layer: this is a regularization layer that helps prevent overfitting by randomly deactivating a portion of the input units during training. Experimental results indicated that the optimal dropout rate is 25% of the total number of input units. This means that during training, the neural network will ignore 25% of the input units (i.e., input words) for each news item. As a result, instead of learning weights that fit only the training dataset, the neural network learns to generalize to similar, previously unseen data. This layer significantly enhances the network's performance when processing news content not encountered during training.
2. Increased Kernel Size in the Conv1D Layer: The kernel size in the one-dimensional convolutional layer was increased from 5 to 7, enabling the model to more effectively extract key features of the analyzed objects and to filter out less significant details.
3. Changed Activation Function in the MaxPooling1D Layer: The activation function of the one-dimensional max pooling layer was changed from the sigmoid function to the ReLU (Rectified Linear Unit) function, as illustrated in Figure 3.

The ReLU (Rectified Linear Unit) activation function is a nonlinear function widely used in neural networks, particularly in deep architectures. Its key characteristics include:

nonlinearity: ReLU is a nonlinear function, which enables neural networks to model complex relationships and solve nonlinear tasks. Without nonlinear activation functions, neural networks lose their capacity to learn and generalize complex patterns.

simplicity and efficiency: ReLU has a simple mathematical structure and is computationally efficient making it well-suited for large-scale deep learning models.

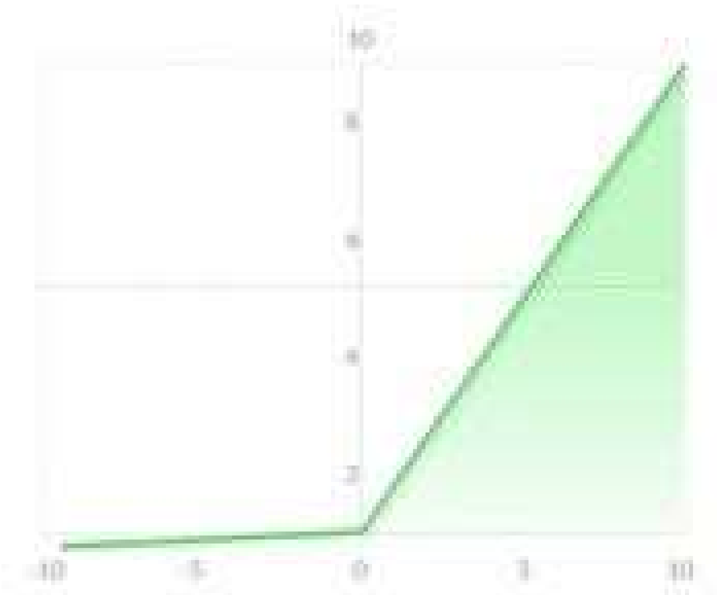

Figure 3: ReLU Activation Function.

The proposed neural network architecture combines embedding and convolutional layers to capture local textual features, followed by an LSTM layer to capture long-term dependencies. These features are then processed by a fully connected (dense) layer to produce the final binary classification output.

The addition of a dropout layer to the existing model significantly improved the network's performance on previously unseen news items. With this layer, the model independently learns weights for the vector representations of each news article. Increasing the kernel size in the convolutional layer allowed the model to detect more substantial features while discarding minor details in the input text. However, this modification also introduced a drawback that increased risk of overfitting. Therefore, combining the dropout layer with the enlarged convolutional kernel size results in better feature and weight extraction while simultaneously reducing the likelihood of overfitting.

3.2. Statistical Metrics for Evaluating Fake News Detection Performance

To evaluate the performance of neural network classifiers in detecting fake news, the following statistical metrics are proposed: accuracy (A), precision (P), recall (R), F1-score is an area under the ROC-curve (AUC), as well as Type I and Type II errors.

Accuracy (A) is a measure of the classifier's ability to correctly classify information as either fake or real. Accuracy (A) can be calculated as follows:

$$A_{ccuracy} = \frac{TP + TN}{TP + TN + FP + FN}, \tag{1}$$

where TP, TN, FP, FN - represent true positives, true negatives, false positives, and false negatives, respectively.

Precision (P) is a measure of the classifier's exactness, where a low value indicates a high number of false positive results. Precision (P) is calculated as the number of true positive predictions divided by the total number of predicted positive instances, and is given by the formula

$$P_{recision} = \frac{TP}{TP + FP}. \tag{2}$$

Recall (R) is a measure of the classifier's completeness; for example, a low recall value indicates a high number of false negative results. It is calculated as the number of true positives divided by the sum of true positives and false negatives:

$$R_{ecall} = \frac{TP}{TP + FN}. \tag{3}$$

The $F1$-score ($F1$) is calculated as the weighted harmonic mean of the classifier's precision and recall:

$$F1 = \frac{2 \cdot P_{recision} \cdot R_{ecall}}{P_{recision} + R_{ecall}} = \frac{2 \cdot TP}{2 \cdot TP + FP + FN}. \tag{4}$$

The Area Under the ROC Curve (AUC) is a metric used to compare learning algorithms and to construct optimal learning models. An AUC value close to 1 indicates a strong system capable of accurately distinguishing between real and fake news, while an AUC value close to 0 indicates a weak system (i.e., one that classifies all fake news as real and vice versa).

The AUC can be calculated using the following expression:

$$AUC = \frac{1 - FPR + TPR}{2}. \tag{5}$$

The True Positive Rate (TPR) refers to the percentage of positive instances that are correctly classified. In contrast, the False Positive Rate (FPR) is the proportion of negative instances that are incorrectly classified as positive, relative to all actual negative instances.

Type I and Type II errors are concepts from mathematical statistics and its applied domains.

4. Experiment

To implement the proposed fake news detection method and assess the credibility of information sources, it is necessary to develop a corresponding application that would allow users to test the method. During the research process, a decision was made to develop the application in a web-based environment. This decision is justified by the fact that a web environment enables efficient processing of large volumes of data through the use of distributed systems and cloud-based solutions, ensuring high performance and processing speed. Web-based applications are also easy to distribute and update, which facilitates the rapid deployment of improved models and ensures users have access to the latest updates without the need to reinstall the application.

The software implementation of the proposed method can be realized as a web application, in which the core and most critical component is the neural network-based fake news detection method.

For the development of the application's backend, the PHP programming language and its Laravel framework were selected. Laravel is a powerful tool for implementing the logic layer of web applications. It provides robust support for working with various databases and fully implements the core principles of object-oriented programming as well as the SOLID principles.

Thus, the use of the Laravel framework allows developers to efficiently build high-quality and user-friendly web applications using advanced web development technologies.

Given the choice of Laravel for the backend, Vue.js was a natural choice for the frontend of the application. According to the official documentation, Laravel and Vue.js are highly compatible and integrate seamlessly with each other. Vue.js is a modern JavaScript framework for building web applications that employs state-of-the-art technologies for compiling and bundling frontend components. Laravel, in turn, provides built-in support for Vite, a modern frontend build tool specifically designed for compiling Vue.js applications.

To implement the styling component of the web application, the Tailwind CSS framework was chosen. Tailwind CSS is a modern utility-first CSS framework for building responsive web interfaces. It provides flexible and fully customizable styling options through the use of utility classes. Instead of writing custom CSS rules manually, developers can apply predefined style classes offered by Tailwind, which streamlines the development process.

For the implementation of the fake news detection method itself, the Python programming language was used.

Overall, this technology stack demonstrates strong synergy and interoperability. The functional architecture of the application is illustrated in Figure 4.

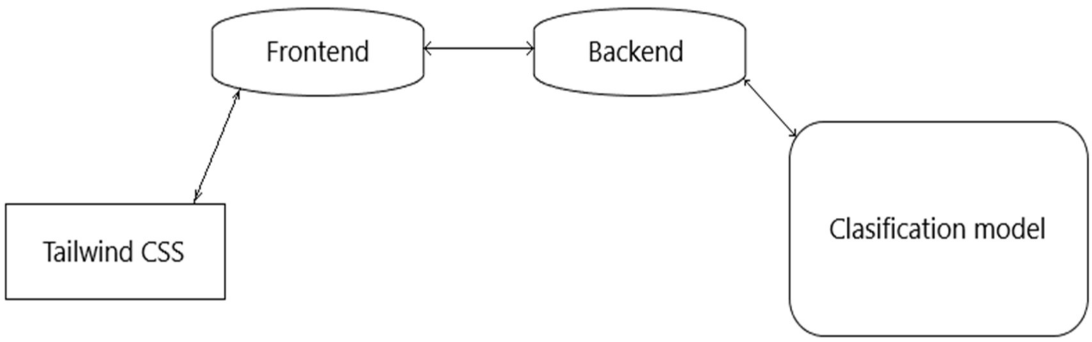

Figure 4: Functional Architecture of the Application.

To design and visualize the structure of the information system, UML diagrams were used.

To model the sequence of interaction with the information system, both an activity diagram and a sequence diagram were developed.

The activity diagram is particularly useful for visualizing the sequence of actions, decision points, and data exchange between various system components. This diagram is presented in Figure 5.

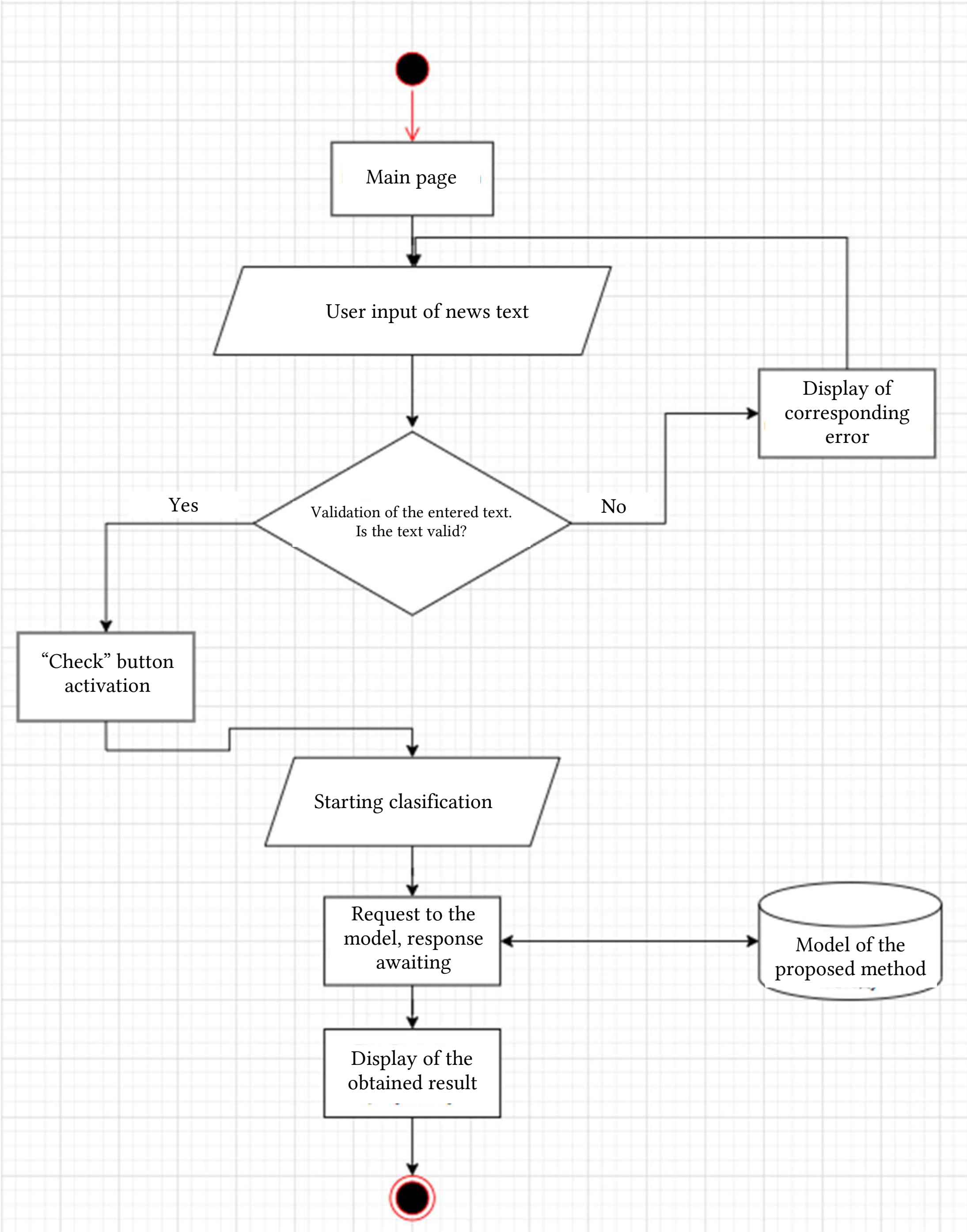

Figure 5: UML Activity Diagram.

The sequence diagram helps to visualize the sequence of events and method calls between different objects which facilitates understanding of the interactions among system components in a specific context. The corresponding diagram is shown in Figure 6.

To represent the software structure of the fake news detection model, a class diagram was developed (Figure 7).

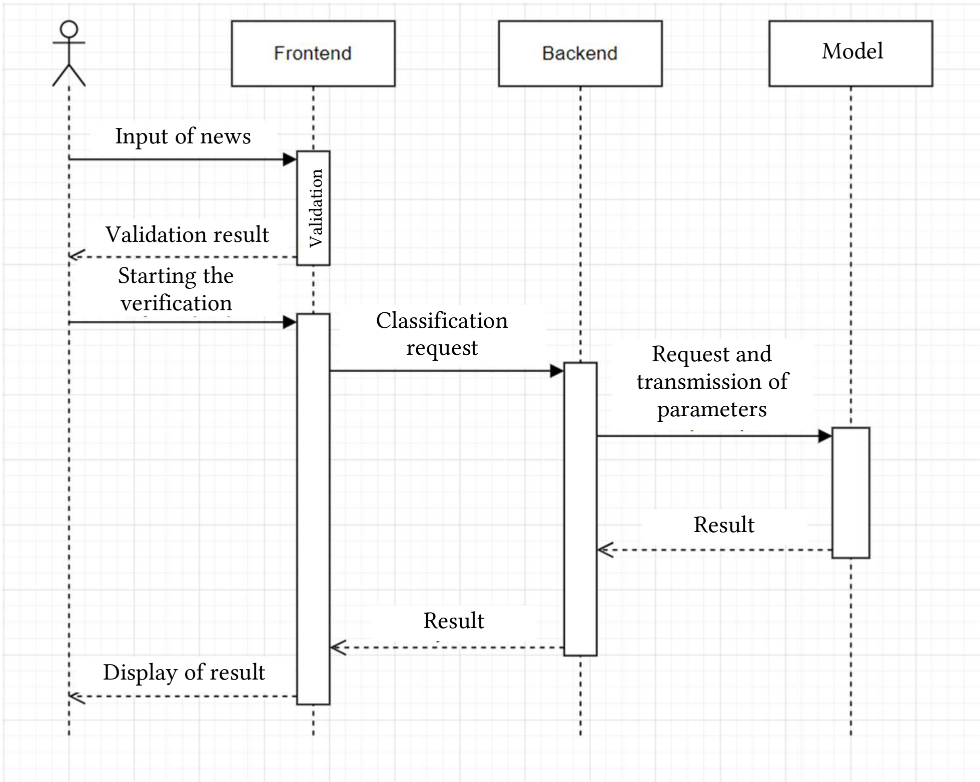

Figure 6: UML Sequence Diagram.

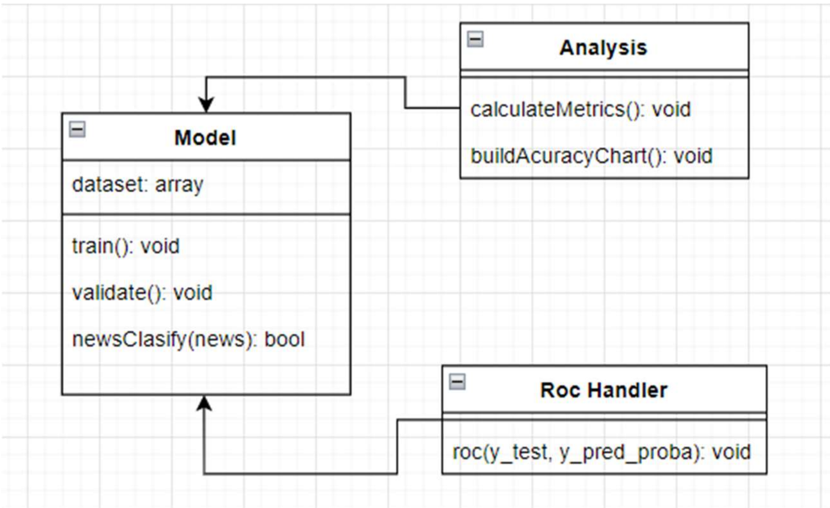

Figure 7: UML Class Diagram.

5. Discussion

An important task is to evaluate the effectiveness of the proposed fake news detection method. To accomplish this, it is advisable to perform a comparative analysis. As models for comparison, the proposed method with an enhanced architecture and existing baseline models should be considered.Among the existing models, the TensorFlow classification model and the Logistic Regression model are selected.

The rationale for comparing the proposed model with the TensorFlow classification model lies in the fact that the latter serves as a prototype for the author's solution. Meanwhile, the relevance of analyzing the Logistic Regression model is based on the following:

Logistic Regression is a machine learning method used to solve binary classification tasks, where the model predicts the probability that an object belongs to a certain class. It employs a logistic (sigmoid) function to calculate probabilities and a logarithmic loss function to evaluate model accuracy.

To facilitate comparative evaluation (experiments) and avoid the need to adapt each model to the system parameters of the experimental computing environment, it is proposed to use the Anaconda software platform.

An essential stage of the method's operation is the training of the respective model. For training the model of the proposed method, the "PolitiFact" and "LIAR" datasets were used [22-25].

PolitiFact is a fact-checking organization that evaluates the accuracy of statements made by politicians and other public sources of information in the United States. Its results are frequently used for analyzing political discourse and assessing the credibility of such statements. In addition, the organization provides public access to its dataset, also named PolitiFact.

This dataset consists of a collection of political news articles. The total size of the dataset is 6235 news items. It includes the following fields: Number, Title, Text, and Label.

Number – the serial number of the news item.

Title – the headline of the news. Typically, the headline alone is insufficient for accurate classification, as it does not provide the neural network with enough information to determine weights and features. Therefore, it is more appropriate to use the next field, Text, for model training.

Text – the full body of the news article, containing a detailed description.

Label – this field has two values: "*Real*" and "*Fake*", which are used to classify the news items.

The dataset was split into two parts – training and testing, according to a specified percentage ratio. The training set was used to train the neural network, while the test set was used to validate the model's performance. The training set consisted of 4988 news items, and the test set consisted of 1247 news items.

Additionally, for more comprehensive validation of the model performance, the LIAR dataset was employed. The LIAR dataset is a collection of textual data that includes statements subjected to fact-checking by PolitiFact. The dataset contains texts annotated with varying levels of truthfulness, indicating whether a statement is reliable or misleading.

Each statement may also be assigned to a specific category, such as politics, economy, health, and others. The dataset structure (i.e., its set of columns) is consistent with the structure of the PolitiFact dataset.

A detailed description of the data from the analyzed datasets is presented in Table 1.

Table 1

Detailed Description of the Analyzed Datasets

Data Set	Number of Records	Number of Verified (Real) News Articles	Number of Fake News	Dataset Labeling
PolitiFact	6235	3741	2494	Number, Title, Text, Label
LIAR	5912	2792	3120	Number, Title, Text, Label

Thus, several datasets with identical annotation schemes were used for training and validating the model.

From a functional perspective, the LIAR dataset was employed to ensure that the addition of a new layer to the neural network architecture and the increase in kernel size of the one-dimensional convolutional layer would enable the network to identify the key features of news articles not only present in the training set but also in previously unseen news.

To evaluate the effectiveness of the proposed method, a series of experiments was conducted, and the results were compared across the analyzed models (i.e., the proposed model, the TensorFlow classification model, and Logistic Regression) using the described datasets (PolitiFact and LIAR).

Training Results on the "PolitiFact" Dataset

Since the values of evaluation metrics may vary between runs, six experimental runs of the models were performed on the test dataset to compute the average overall accuracy. The following formula was used to calculate the average overall accuracy:

$$A = \frac{\sum_{i=1}^{N} x_i}{N}, \tag{6}$$

where N – total number of experiments; x_i – the value of the corresponding indicator; i – sequence number.

Having obtained the average accuracy, it is possible to calculate the standard deviation using the following formula:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - A)^2}{N}}. \tag{7}$$

Thus, the value of the standard deviation corresponds to the error of the mean, which can be expressed as ± a specific value. In other words: Overall accuracy = Mean accuracy ± Standard deviation.

Experiment № 1 Results

Figure 8 presents a graph illustrating the accuracy trend of the proposed method over training steps. The overall accuracy achieved by the proposed method on this dataset amounted to 93,32%.

Table 2 displays the results of all evaluated models on various performance metrics for the input dataset used in Experiment 1.

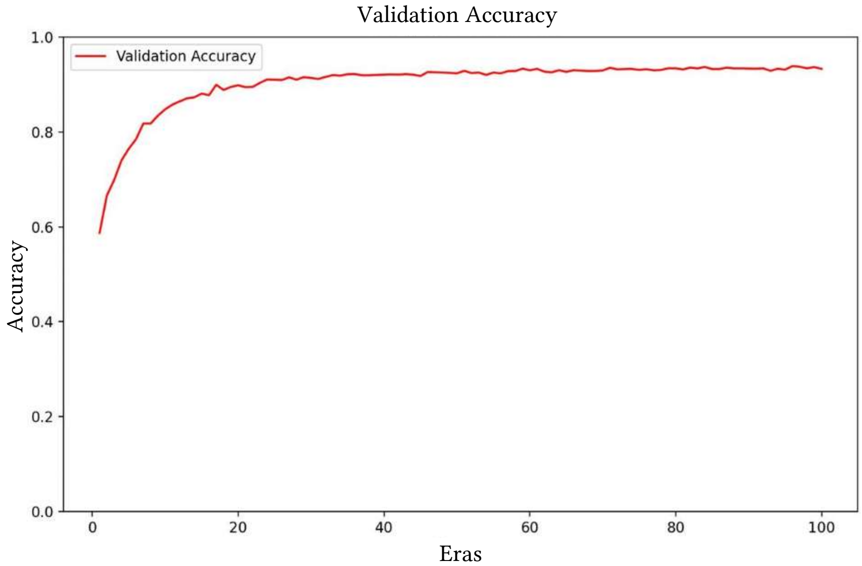

Figure 8: Overall Accuracy on the PolitiFact Training Dataset.

Table 2

Performance Metrics of the Investigated Methods on the PolitiFact Dataset

Method	Overall Accuracy	F1-score	recall	precision
The proposed model	93,32%	0,926	0,919	0,93
TensortFlow classification model	89,37%	0,893	0,88	0,896
LogisticRegression	91,4%	0,914	0,924	0,916

The increase in metric values can be observed in Table 3.

Table 3
Improvement in Metric Values of the Proposed Method Compared to Existing Methods on the PolitiFact Dataset

Pairs of Methods under Comparison	Overall Accuracy	F1-score	recall	precision
The proposed model - TensortFlow classification model	3,95%	0,033	0,039	0,034
The proposed model - LogisticRegression	1,920%	0,012	-0,005	0,014

As shown in Table 3, the proposed method outperforms the existing TensorFlow classification model across all evaluation metrics. It also demonstrates superior performance over the Logistic Regression model in all metrics except for recall.

A similar evaluation was conducted in Experiments 2 through 6.

Based on the aggregated results from Experiments 1 to 6, it can be concluded that the average accuracy of the proposed method is 93,22% (according to Equation (6)). Considering Equation (7), the standard deviation was calculated to be 0,99%. These results indicate that the proposed method performs consistently and effectively on the PolitiFact training dataset.

However, to ensure more robust validation of the method’s performance, it is advisable to evaluate it further using the “LIAR” dataset.

Experiment №. 1

Figure 9 presents the accuracy trend of the proposed method over training steps. The overall accuracy of the proposed method on this dataset reached 91,36%.

Table 4 summarizes the performance metrics of the evaluated models on the input dataset used in Experiment 1.

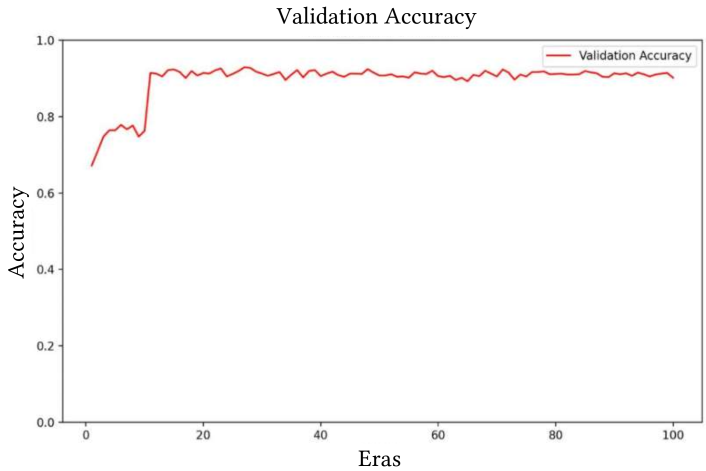

Figure 9: Overall Accuracy on the “LIAR” Training Dataset.

Table 4
Performance Metrics of the Evaluated Methods on the “LIAR” Dataset

Method	Overall Accuracy	F1-score	recall	precision
The proposed model	91,36%	0,913	0,902	0,915
TensortFlow classification model	89,11%	0,891	0,889	0,892
LogisticRegression	91,18%	0,911	0,908	0,915

The increase in performance metric values can be observed in Table 5.

Table 5
Improvement in Metric Values of the Proposed Method Compared to Existing Methods on the "LIAR" Dataset

Pairs of Methods under Comparison	Overall Accuracy	F1-score	recall	precision
The proposed model - TensortFlow classification model	2,25%	0,022	0,013	0,023
The proposed model - LogisticRegression	0,18%	0,002	-0,006	0

As shown in Table 5, the proposed method demonstrates superior performance across all metrics compared to the existing TensorFlow classification model. Additionally, the proposed method outperforms the Logistic Regression model in all metrics except for the recall metric.

A similar evaluation was carried out in Experiments 2-6.

Based on the aggregate results from Experiments 1-6, it can be concluded that the average accuracy of the proposed method is 91,57% (in accordance with formula (6)). Taking into account formula (7), the standard deviation was found to be 0,78%. This leads to the conclusion that the proposed method demonstrates a high level of effectiveness when applied to the "LIAR" dataset.

At the same time, on the new "LIAR" dataset, the overall accuracy of the proposed method is slightly lower than on the training dataset "PolitiFact." However, similar performance degradation is also observed for the existing models examined in this study.

Therefore, the proposed method generally demonstrates higher effectiveness compared to the existing approaches under investigation. Consequently, it can be applied both for verifying the credibility of news and for their real-time classification.

6. Conclusion

Based on the results of the conducted study, the following conclusions can be drawn:

- a promising approach to improving the effectiveness of fake news detection in online social media is to enhance the detection method based on a multi-layer convolutional neural network (CNN);
- the essence of the proposed enhancement lies in adding a dropout layer, increasing the kernel size, and replacing the activation function of the one-dimensional max-pooling layer from a sigmoid function to the ReLU function.
- the rationale for using the "PolitiFact" and "LIAR" datasets for training the neural network has been provided;
- an information system has been developed that implements the proposed method. The backend part of the application was developed using the PHP programming language and the Laravel framework. For the frontend, the Vue.js framework was used. The Tailwind CSS framework was chosen to implement the styling of the web application. The fake news detection method itself was implemented using the Python programming language and such libraries as TensorFlow, NumPy, and Scikit-learn for neural network processing.
- the effectiveness of the proposed fake news detection method has been evaluated. The compared models included the proposed method with the improved architecture and existing models: the TensorFlow classification model and Logistic Regression. According to the experimental results, the average overall accuracy of the proposed method on the "PolitiFact" dataset is 93,22%, with a standard deviation of 0,99%. On the "LIAR" dataset, the average overall accuracy is 91,57%, and the standard deviation is 0,78%. The proposed method demonstrates higher efficiency compared to the studied

existing models and can thus be applied both for assessing the credibility of news and for real-time classification.

The direction of future research is seen in further improving the effectiveness of the proposed method.

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

[1] Oliveira, N.R.; Medeiros, D.S.; Mattos, D.M. A sensitive stylistic approach to identify fake news on social networking. IEEE Signal Process. Lett. 2020, 27, 1250–1254. [CrossRef]

[2] Alghamdi, O.; Lin, Y.; Luo, S.: A comparative study of machine learning and deep learning techniques for fake news detection. Information 2022, 13, 576, 28 p.

[3] Guo, M.; Xu, Z.; Liu, L.; Guo, M.; Zhang, Y. An Adaptive Deep Transfer Learning Model for Rumor Detection without Sufficient Identified Rumors. Math. Probl. Eng. 2020, 7562567. [CrossRef]

[4] Kim, Y.; Kim, H.K.; Kim, H.; Hong, J.B. Do Many Models Make Light Work Evaluating Ensemble Solutions for Improved Rumor Detection. IEEE Access 2020, 8, 150709–150724. [CrossRef]

[5] Barmak O. V., Borovyk D. O., Borovyk O. V., Skrypnyk T. K. Possible directions for enhancing existing approaches to fake news detection include leveraging the potential of machine and deep learning algorithms, analyzing the sentiment of news content and examining emotional cues in user comments. Collective Scientific Works of the Military Institute of Taras Shevchenko National University of Kyiv. Issue № 80. Kyiv: Military Institute of KNU, 2024, 98-117.

[6] Tian, L.; Zhang, X.; Wang, Y.; Liu, H. Early detection of rumours on twitter via stance transfer learning. In Advances in Information Retrieval: 42nd European Conference on IR Research, ECIR 2020, Lisbon, Portugal, 14–17 April 2020, Proceedings, Part I 42; Springer: Cham, Switzerland, 2020; Volume 12035, p. 575.

[7] Liu, Y.; Wu, Y.-F.B. Fned: A deep network for fake news early detection on social media. ACM Trans. Inf. Syst. (TOIS) 2020, 38, 1–33. [CrossRef]

[8] Lin, L.; Chen, Z. Social rumor detection based on multilayer transformer encoding blocks. Concurr. Comput. Pract. Exp. 2021, 33, 6083. [CrossRef]

[9] Albahar, M. A hybrid model for fake news detection: Leveraging news content and user comments in fake news. IET Inf. Secur. 2021, 15, 169–177. [CrossRef]

[10] Alonso, M.A.; Vilares, D.; Gómez-Rodríguez, C.; Vilares, J. Sentiment analysis for fake news detection. Electronics 2021, 10, 1348. [CrossRef]

[11] Ghanem, B.; Rosso, P.; Rangel, F. An emotional analysis of false information in social media and news articles. ACM Trans. Internet Technol. (TOIT) 2020, 20, 1–18. [CrossRef]

[12] Vosoughi, S.; Roy, D.; Aral, S. The spread of true and false news online. Science 2018, 359, 1146–1151. [CrossRef]

[13] Kumari, R.; Ashok, N.; Ghosal, T.; Ekbal, A. What the fake? Probing misinformation detection standing on the shoulder of novelty and emotion. Inf. Process. Manag. 2022, 59, 102740. [CrossRef]

[14] Zhang, X.; Cao, J.; Li, X.; Sheng, Q.; Zhong, L.; Shu, K. Mining dual emotion for fake news detection. In Proceedings of the WWW '21: The Web Conference 2021, Ljubljana, Slovenia, 19–23 April 2021; 3465–3476.

[15] Mikolaj Karpinski, Liudmyla Borovyk, Serhii Lienkov, Vitalii Savchenko, Iryna Panibrat, Vadym Sobko Theoretical and Applied Bases of Creating a System for Automatic Phraseological Units Classification in English Texts Based on the Artificial Intelligence Technologies Usage. Proceedings of the Computational Linguistics Workshop (CLW-CoLInS-

2025) at the 9th International Conference on Computational Linguistics and Intelligent Systems (CoLInS-2025) – Vol-3976, May 15–16, 2025, Kharkiv, Ukraine. ISSN 1613-0073; 219-233. https://ceur-ws.org/Vol-3976/

[16] Imran, A.S.; Daudpota, S.M.; Kastrati, Z.; Batra, R. Cross-cultural polarity and emotion detection using sentiment analysis and deep learning on COVID-19 related tweets. IEEE Access 2020, 8, 181074–181090. [CrossRef]

[17] Pota, M.; Ventura, M.; Catelli, R.; Esposito, M. An effective BERT-based pipeline for Twitter sentiment analysis: A case study in Italian. Sensors 2020, 21, 133. [CrossRef]

[18] Islam, M.R.; Liu, S.; Wang, X.; Xu, G. Deep learning for misinformation detection on online social networks: A survey and new perspectives. Soc. Netw. Anal. Min. 2020, 10, 82. [CrossRef] [PubMed]

[19] Borovyk D. O., Barmak O. V. An improved Method for fake news detection based on the use of CNN neural network. Bulletin of Khmelnytskyi National University. Technical Sciences. № 5 (Vol. 2), 2023. Khmelnytskyi: Publishing and Printing Center of KhNU, 2023, 19-24.

[20] Bahad, P.; Saxena, P.; Kamal, R. Fake News Detection using Bi-directional LSTM-Recurrent Neural Network. Procedia Comput. Sci. 2019, 165, 74–82. [CrossRef]

[21] Subramani, S.; Wang, H.; Vu, H.Q.; Li, G. Domestic violence crisis identification from facebook posts based on deep learning. IEEE Access 2018, 6, 54075–54085. [CrossRef]

[22] https://github.com/KaiDMML/FakeNewsNet

[23] https://paperswithcode.com/dataset/liar

[24] https://www.v7labs.com/blog/f1-score-guide

[25] https://en.wikipedia.org/wiki/Precision_and_recal