Computer Engineering and Applications Vol. 13, No. 3, October 2024

Fake News Detection Using Optimized Convolutional Neural Network and Bidirectional Long Short-Term Memory

Winda Kurnia Sari[1], Iman Saladin B. Azhar[2], Zaqqi Yamani[3], Yesinta Florensia[4]

[1,3]*Department of Information System, Faculty of Computer Science, Universitas Sriwijaya*
[2]*Department of Computer Engineering, Faculty of Computer Science, Universitas Sriwijaya*
[4]*Department of Informatics Engineering, Faculty of Computer Science, Universitas Sriwijaya*
**imansaladin@unsri.ac.id*

ABSTRACT

The spread of fake news in the digital age threatens the integrity of online information, influences public opinion, and creates confusion. This study developed and tested a fake news detection model using an enhanced CNN-BiLSTM architecture with GloVe word embedding techniques. The WELFake dataset comprising 72,000 samples was used, with training and testing data ratios of 90:10, 80:20, and 70:30. Preprocessing involved GloVe 100-dimensional word embedding, tokenization, and stopword removal. The CNN-BiLSTM model was optimized with hyperparameter tuning, achieving an accuracy of 96%. A larger training data ratio demonstrated better performance. Results indicate the effectiveness of this model in distinguishing fake news from real news. This study shows that the CNN-BiLSTM architecture with GloVe embedding can achieve high accuracy in fake news detection, with recommendations for further research to explore preprocessing techniques and alternative model architectures for further improvement.

Keywords: Fake News, CNN-BiLSTM, Word Embedding.

1. INTRODUCTION

In recent years, the rapid advancement of online social media has facilitated interactions and information exchange among individuals, allowing them to connect and stay updated with current news easily. However, behind this convenience, much of the information circulating on social media is dubious and even intentionally designed to deceive, such as fake news, hoaxes, clickbait, and satirical news [1]. The spread of such information can have negative and potentially harmful impacts across various sectors, from politics to public health to maintain user trust, social networks must address fake news on their platforms, even though this may result in a decrease in the number of users. Avoiding fake news is not an easy task, so it is crucial to refer to research conducted on fake news detection [2],[3].

Various researchers are currently striving to detect fake news with the aid of machine learning, which has proven effective in this task. Different algorithms are applied to enhance accuracy in identifying fake news. A study by [4] highlights the prevalence of the fake news problem in today's vast Internet environment. Using the

WELFake and Real and Fake News Classification datasets from Kaggle, this research compares five machine learning algorithms for fake news classification: Naïve Bayes, decision tree, random forest, support vector machine, and K-nearest neighbor. The results show that decision tree (DT) and random forest (RF) achieved the highest accuracy, reaching 93.92% and 91.18%, respectively.

According to the study by [5], the WELFake model for fake news detection achieved an accuracy of 96.73% on the WELFake dataset, surpassing models such as BERT and CNN with accuracy improvements of 1.31% and 4.25%, respectively. This model also enhanced accuracy by up to 10% on the McIntire and BuzzFeed datasets and outperformed Word2vec methods by 1.73%. The research plans to further develop the model by incorporating knowledge graphs and user credibility. The work by [6] addresses the issue of misinformation spread on social media platforms, where users often rely on these networks for information. The study highlights the dangers of accepting rumors and false information as truth. By employing machine learning techniques and natural language processing, this research aims to identify fake news with high accuracy. The proposed method, which utilizes support vector machines and the WELFake dataset, has been compared to existing models and demonstrates impressive accuracy of up to 93.6% in detecting misinformation.

Another important study by [7] proposes a fake news detection approach using a deep learning model that integrates word embedding techniques with Convolutional Neural Networks (CNN) and Bidirectional Long Short-Term Memory (BiLSTM). This model, trained on the unbiased WELFake dataset, demonstrated that the combination of Word2Vec CBOW and Word2Vec Skip-gram models with CNN on BiLSTM layers achieved an accuracy of up to 97%.

By leveraging the strengths of CNN and BiLSTM for fake news detection, the model can be trained to effectively capture both spatial and temporal features of the text. CNNs are capable of extracting local patterns in the text, such as sequences of words and important phrases, while BiLSTM can understand long-term context by considering word sequences in both forward and backward directions. This combination allows the model to identify complex linguistic features and filter out fake news with high accuracy. The implementation of this method on the WELFake dataset demonstrates that this hybrid approach excels in detecting fake news, offering a more effective solution to the challenges of misinformation in the digital age.

2. MATERIAL AND METHODS

2.1 PROPOSED METHODOLOGY

This study proposes a methodology that includes a literature review of the past 5 years and the use of a public dataset consisting of 72,000 samples for the data collection phase. In the data preprocessing stage, the first step involves embedding words using GloVe (Global Vectors for Word Representation) with 100-dimensional vectors. This is followed by splitting the data into training and testing sets, tokenization, and stopwords removal. Once data preprocessing is complete, a CNN-BiLSTM model is constructed with parameter tuning. Subsequently, the model is evaluated, and result analysis is conducted to understand the model's performance in detecting fake news and identifying the factors influencing its performance.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

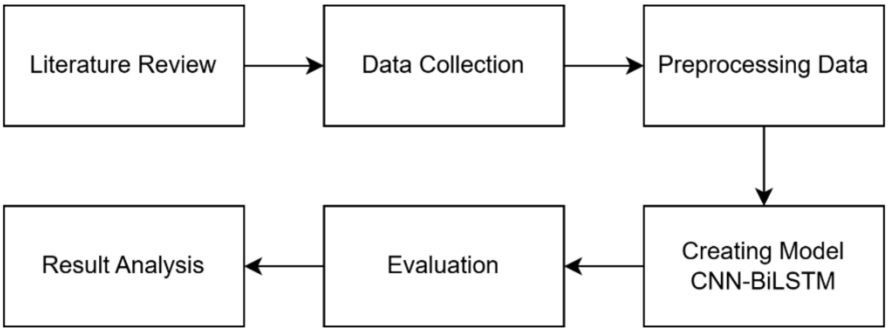


FIGURE 1. Proposed Methodology

2.2 DATASET

A detailed WELFake dataset was created by merging four prominent news sources—Kaggle, McIntire, Reuters, and BuzzFeed Political. Designed to enhance classifier accuracy and improve machine learning training, this dataset consists of 72,134 news articles, with 35,028 labeled as real and 37,106 as fake. Although the original CSV file included 78,098 entries, only 72,134 entries were utilized in the data frame. The dataset comprises four columns: Serial Number, Title, Text, and Label, with 0 indicating fake news and 1 indicating real news [5].

Table 1 presents the WELFake open dataset, showing the 72,134 news articles categorized into 35,028 real and 37,106 fake entries. Table 2 illustrates the even distribution of fake and real news across the four feature categories within the WELFake dataset.

2.3 WORD EMBEDDINGS

Word embedding systems and language structures vary significantly across different languages, necessitating the creation of tailored word embedding models for each specific language [8]. Various types of word embeddings, such as GloVe (Global Vectors for Word Representation), Word2Vec [9], and FastText [10], provide different methods for representing words as vectors.

GloVe is an unsupervised learning method that generates vector representations of words by analyzing global co-occurrence statistics from a text corpus. The model captures interesting linear relationships in the word vector space and is trained on word co-occurrence data derived from a single pass through the corpus, which can be computationally demanding for large datasets but is efficient in subsequent training stages. The GloVe model, as demonstrated by [11], utilizes word co-occurrence matrices to represent word relationships and is commonly trained on extensive corpora such as Wikipedia 2014 and Gigaword 5. For this research, a GloVe model was employed with a 6 billion word corpus and 100-dimensional vectors to enhance the accuracy of word representations, which resulted in generating 400,000 word vectors.

2.4 DATA PREPROCESSING

2.3.1. TOKENIZATION

Tokenization, which involves dividing text into smaller components such as words, is an essential first step in numerous NLP tasks, including machine translation, sentiment analysis, and information retrieval. Although tokenization is relatively fast compared to other steps, improving its efficiency is vital as it can reduce overall inference latency. Enhanced tokenization can significantly impact mobile NLP applications, where minimizing latency is crucial. Additionally, faster tokenization systems can greatly save computational resources for large web services like Google, Facebook, and Twitter, which handle billions of queries and web pages. Efficient tokenization not only improves performance but also helps in reducing power consumption and environmental impact [12].

TABLE 1.
Dataset Specifications

Dataset	Real	Fake
Kaggle	10387	10413
McIntire	3171	3164
Reuters	21417	23481
BuzzFeed Politicial	53	48
WELFake dataset	35028	37106

TABLE 2.
Distribution of fake and real news in WELFake

Category	Real %	Fake %
Short sentences	60.9	39.1
Readability index	51.7	48.3
Subjectivity	45.4	54.6
Number of articles	53.9	46.1
WELFake dataset	48.5	51.45

2.3.2. STOP WORDS REMOVAL

In Natural Language Processing (NLP), stop words refer to frequently used words that contribute little meaning to the text and are generally excluded during preprocessing. Common examples include "are," "of," "the," and "at" [13]. Removing these words is standard practice as they have minimal influence on tasks such as sentiment analysis, thereby enhancing both the efficiency and precision of the analytical process.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

2.5 CNN-BILSTM ARCHITECTURE

Convolutional Neural Networks (CNNs), a distinct type of artificial neural network recognized for their ability to effectively identify features in various positions. This model has effectively tackled various issues in image processing and has been successfully applied to natural language tasks like sentiment analysis, question answering, and text summarization. Its unique architecture is crafted to optimize learning efficiency. A CNN functions as a multi-layered network where each layer's output becomes the input for the following layer. Typically, it includes an input layer, multiple hidden layers, and an output layer [14].

BiLSTM and LSTM are specialized types of Recurrent Neural Networks (RNNs) tailored for sequential data processing. While LSTM effectively addresses the vanishing gradient problem found in standard RNNs, BiLSTM improves performance even further by processing data in both forward and backward directions. This makes it especially efficient for tasks involving text and other sequential data [15]. For instance, in text analysis, features are typically extracted using a convolutional layer before being processed by a BiLSTM layer. Details regarding the parameters and training of the CNN–BiLSTM architecture can be found in Table 3.

TABLE 3.
Proposed CNN-BiLSTM Architecture

Layer type	Description
Embedding Layer	Converts words into embedding vectors using GloVe.
BiLSTM Layer (1st)	Processes sequential data in both directions (forward and backward) with 64 units.
Conv1D layer	Extracts local features from the output of the first BiLSTM with 16 filters and a kernel size of 3.
MaxPooling1D layer	Reduces the dimensionality of the data with a pool size of 2.
BiLSTM Layer (2nd)	Processes sequential data in both directions with 16 units.
BiLSTM Layer (3rd)	Processes sequential data in both directions with 32 units, returning the output as a single vector.
Dense layer	Connects all neurons with 64 units and ReLU activation.
Dropout layer	Reduces overfitting with a dropout rate of 0.5.
Output layer	The last layer with 1 unit and sigmoid activation for binary classification.

3. RESULTS AND DISCUSSION

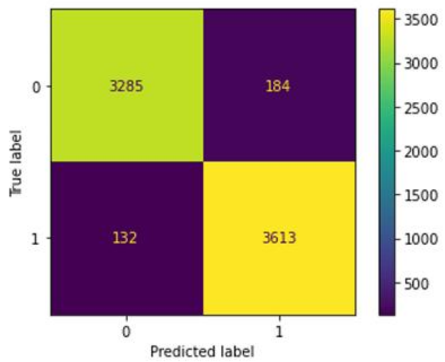
After training and testing the data on the WELFake dataset using the CNN-BiLSTM architecture and tuning its hyperparameters, a satisfactory accuracy result was achieved. This study divided the training and testing data into ratios of 90:10, 80:20, and 70:30, and it was concluded that a larger data ratio results in better performance in this case. To evaluate the learning process, the test data consisted of

50,493 samples. The study also utilized the `ReduceLROnPlateau` callback, which dynamically reduces the learning rate when validation accuracy does not improve after several epochs. This strategy helps the model achieve faster convergence and prevent overfitting by adjusting the learning rate as needed during training.

Figure 2 shows the confusion matrix and model classification report for the 90:10 ratio, with an accuracy of 96%. Additionally, figures 3 and 4 illustrate the loss and accuracy curves for both the training and validation data at a 90:10 dataset ratio. Furthermore, the accuracy comparison results for the three data ratios, which have been trained and tested with consistent parameters and preprocessing, can be found in Table 4. Overall, an ROC Curve result of 0.99 indicates that your model performs exceptionally well and has a strong predictor in distinguishing between classes within the dataset. The result can be seen in Figure 5.

TABLE 4.
Accuracy Comparison Results

Ratio Data	Accuracy %
90:10	96
80:20	95
70:30	95




```
Model classification report :
              precision    recall  f1-score   support

           0       0.96      0.95      0.95      3469
           1       0.95      0.96      0.96      3745

    accuracy                           0.96      7214
   macro avg       0.96      0.96      0.96      7214
weighted avg       0.96      0.96      0.96      7214
```

FIGURE 2. The result of the confusion matrix for the 90:10 data split

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

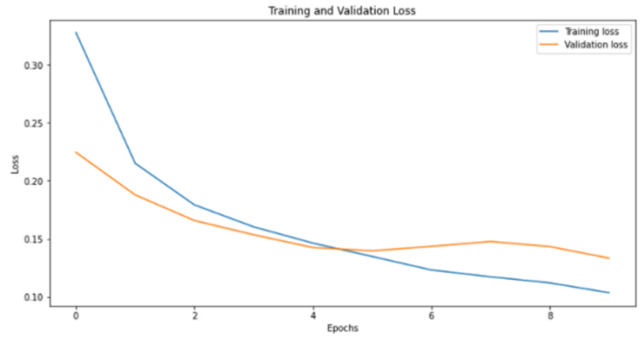


FIGURE 3. Training dan Validation Loss

FIGURE 4. Training dan Validation Accuracy

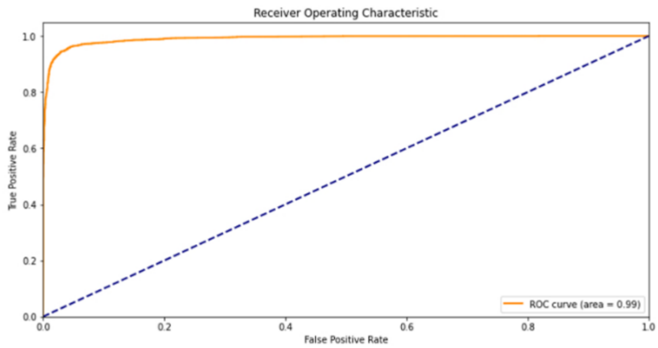


FIGURE 5. ROC Curve

4. CONCLUSION

The result shows that the CNN-BiLSTM architecture, when applied to the WELFake dataset, effectively achieves high performance in detecting fake news. The model's performance is significantly influenced by the ratio of training to test data, with larger ratios leading to improved accuracy. Specifically, using a 90:10 training-to-test ratio yielded a notable accuracy of 96%, indicating the model's robust ability to differentiate between real and fake news articles. The application of GloVe embeddings with 100-dimensional vectors, resulting in a total of 400,000-word vectors, enhanced the accuracy of word representations and contributed to the model's superior performance.

Overall, the study confirms that the CNN-BiLSTM model is effective for fake news detection and highlights the importance of using adequate data ratios, advanced embedding techniques, and dynamic learning rate adjustments to optimize model performance.

REFERENCES

[1] Z. Khanam, B. N. Alwasel, H. Sirafi, and M. Rashid, “Fake News Detection Using Machine Learning Approaches.” IOP conference series: materials science and engineering (Vol. 1099, No. 1, p. 012040), 2021.

[2] R. Katarya and M. Massoudi, “Recognizing Fake News in Social Media with Deep Learning: A Systematic Review,” in *2020 4th International Conference on Computer, Communication and Signal Processing (ICCCSP)*, 2020, pp. 1–4.

[3] X. Zhang and A. A. Ghorbani, “An overview of online fake news: Characterization, detection, and discussion,” *Inf. Process. Manag.*, vol. 57, no. 2, p. 102025, 2020.

[4] S. Sahu, P. Bansal, and R. Kumari, “Enhancing Information Integrity: Machine Learning Methods for Fake News Detection BT - Fourth Congress on Intelligent Systems,” 2024, pp. 247–257.

[5] P. K. Verma, P. Agrawal, I. Amorim, and R. Prodan, “WELFake: Word Embedding Over Linguistic Features for Fake News Detection,” *IEEE Trans. Comput. Soc. Syst.*, vol. 8, no. 4, pp. 881–893, 2021.

[6] K. V Mahijith, A. U. Sunil, C. A., A. Sreelathan, and M. V Thomas, “A Novel Fake News Detection Approach Using Machine Learning,” in *2022 Third International Conference on Intelligent Computing Instrumentation and Control Technologies (ICICICT)*, 2022, pp. 448–454.

[7] M.-A. Ouassil, B. Cherradi, S. Hamida, M. Errami, O. El Gannour, and A. Raihani, “A Fake News Detection System based on Combination of Word Embedded Techniques and Hybrid Deep Learning Model.” (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 13, No. 10, 2022.

[8] S. Wang, W. Zhou, and C. Jiang, “A survey of word embeddings based on deep learning,” *Computing*, vol. 102, no. 3, pp. 717–740, 2020.

[9] S. J. Johnson, M. R. Murty, and I. Navakanth, “A detailed review on word embedding techniques with emphasis on word2vec,” *Multimed. Tools Appl.*, vol. 83, no. 13, pp. 37979–38007, 2024.

[10] N. Badri, F. Kboubi, and A. H. Chaibi, “Combining FastText and Glove Word Embedding for Offensive and Hate speech Text Detection,” *Procedia Comput. Sci.*, vol. 207, pp. 769–778, 2022.

[11] J. Pennington and R. Socher, “Glove: Global vectors for word representation,” in *AES: Journal of the Audio Engineering Society*, 1971, vol. 19, no. 5, pp. 417–425.

[12] X. Song, A. Salcianu, Y. Song, D. Dopson, and D. Zhou, “Fast WordPiece Tokenization Xinying.” arXiv preprint arXiv:2012.15524, 2020.

[13] N. S. M. Nafis and S. Awang, “An Enhanced Hybrid Feature Selection Technique Using Term Frequency-Inverse Document Frequency and Support Vector Machine-Recursive Feature Elimination for Sentiment Classification,”

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

IEEE Access, vol. 9, pp. 52177–52192, 2021.

[14] M. Rhanoui, M. Mikram, S. Yousfi, and S. Barzali, “A CNN-BiLSTM Model for Document-Level Sentiment Analysis.” Machine Learning and Knowledge Extraction, 1(3), pp. 832–847, 2019.

[15] M. F. Aslan, M. F. Unlersen, K. Sabanci, and A. Durdu, “CNN-based transfer learning–BiLSTM network: A novel approach for COVID-19 infection detection.” Applied Soft Computing, 98, 106912, 2021.