Journal of Digital Security and Forensics
January-June 2026 3(1)
www.digitalsecurityforensics.org

Original Article

RECOGNITION OF FAKE NEWS WITH DEEP LEARNING ARCHITECTURE LSTM

Noor Hasan Fadhil [1*]
[1] Al-Qasim Green University, Babylon, Iraq

ABSTRACT

Fake news spread through digital channels has been a social issue due to which it has become imperative to detect such news. In this regard, this paper aims to develop a classifier for fake news based on deep learning. The proposed technique consists of various phases including preprocessing of the data to eliminate noise, tokenize data, and vectorize data after which the classifier will be trained using an LSTM model. The LSTM model suggested makes use of its potential to incorporate sequence and context-based information within the text data. The empirical findings show that the model provides 99.7% accuracy and outperforms many traditional machine learning models and contemporary hybrid models. Various other metrics such as precision, recall, and F1 score validate the effectiveness of the model. However, despite its excellent performance, more testing of the model on different datasets is recommended

Keywords: Fake News Detection, Deep Learning, LSTM, Natural Language Processing

INTRODUCTION

Fake news, which refers to fabricated information presented as actual news, poses a major risk to healthy dialogue in any society that prides itself on democracy Sharma et al. (2022). This phenomenon, characterized by the dissemination of fabricated stories, distorted facts, and sensationalized headlines, exploits the rapid and widespread nature of digital media to manipulate public opinion and sow confusion D'Ulizia et al. (2021). The rapid spread of such misinformation, particularly via social media, calls for strong and automatic identification tools in order to preserve information credibility Güler and Gunduz (2023). The problems related to the identification of fake news become even more complicated due to its different types that might range from blatant falsehoods to distortions, not to mention the natural inability to keep up with a flood of information available on the Internet Hosseini et al. (2023). Thus, it has become essential to have the proper methodologies in place in order to identify fake news Pierri and Ceri (2019).

In the past, many attempts have been made in order to detect fake news using linguistic techniques, metadata analysis, and other network-based features Alghamdi et al. (2023). Later developments have revolved around machine learning models, including deep learning models that have proven effective in automating the process through the use of complex patterns derived from large volumes of data Alghamdi et al. (2023). This includes the use of deep learning models such as CNNs in feature extraction, and also the use of LSTM networks in capturing long-range dependencies in the text data Killi et al. (2022). Hybrid models based on the combination of the abilities of CNNs to extract spatial features and LSTM networks to analyze temporal sequences have been proposed to handle the challenges of fake news Ajik et al. (2023).The present research tackles this urgent problem through the development of a unique model for distinguishing between real news articles and fake news, using state-of-the-art deep learning

*Corresponding Author:
Email address: Noor Hasan Fadhil (noor_hassan@uoqasim.edu.iq)

Received: 21 February 2026; **Accepted:** 13 April 2026; **Published** 09 May 2026
DOI: 10.29121/DigiSecForensics.v3.i1.2026.100
Page Number: 19-26
Journal Title: Journal of Digital Security and Forensics
Journal Abbreviation: J. Digi. Sec. Forensics
Online ISSN: 3048-894X
Publisher: Granthaalayah Publications and Printers, India
Conflict of Interests: The authors declare that they have no competing interests.
Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.
Authors' Contributions: Each author made an equal contribution to the conception and design of the study. All authors have reviewed and approved the final version of the manuscript for publication.
Transparency: The authors affirm that this manuscript presents an honest, accurate, and transparent account of the study. All essential aspects have been included, and any deviations from the original study plan have been clearly explained. The writing process strictly adhered to established ethical standards.

algorithms that perform the analysis of textual features for detecting linguistic cues of deception Hosseini et al. (2023). This work makes an important contribution to the field of combating fake news, developing new ways of improving fake news detection methods Akinyemi et al. (2020). Indeed, one of the unique strengths of using Long Short-Term Memory networks for such purposes lies in their ability to detect complex linguistic patterns that serve as reliable cues of deception Killi et al. (2022). Such an approach allows the model to analyze intricate linguistic patterns that play a significant role in the distinction between fake and real news articles Ajik et al. (2023).

LITERATURE REVIEW

Existing research on fake news detection has explored various techniques, including deep learning and natural language processing, yet they often overlook the complete optimization of model hyperparameters, which can significantly influence the classification results Emmy et al. (2023)

Several studies have shown that deep learning methods (e.g., CNNs, LSTMs, BERT) are useful in capturing the complex linguistic patterns characteristic of deceptive content Hosseini et al. (2023), Güler and Gunduz (2023).

However, there is still a lack of effective use of contextual cues and skip connections in models, which in turn limits the development of detection systems that can utilize more extensive contextual information. To handle data propagation through multiple neural layers, hierarchical stacked models like FakeStack using Tri-BERT-CNN-LSTM architectures were created Keya et al. (2023).

Furthermore, the effectiveness of these models can be greatly enhanced by using attention-based mechanisms or deep variational models, which allow the system to weigh the importance of different words and thus detect key features suggestive of misinformation Hosseini et al. (2023).

For this period, frameworks often used BERT-base to encode textual content and CNN and max-pooling layers to reduce features and used BiLSTM layers to process associated metadata to capture contextual dependencies Akinyemi et al. (2020).

Some other advanced models showed high performance by combining ensemble learning approaches with different machine learning and deep learning techniques to separate real and fake news Rezaei et al. (2022).

Moreover, the ambiguity of natural language interpretation was addressed by leveraging multi-EDU-structure awareness and enhanced text representations Wang et al. (2022).

In addition to textual analysis, research before now highlights the importance of multimodal approaches that include visual and contextual information to improve the detection accuracy of fake news D'Ulizia et al. (2021). These versatile approaches highlight the importance of robust data processing pipelines, where techniques like GloVe are being used for word representation to facilitate sophisticated deep learning methods like concatenated CNNs and LSTMs Güler and Gunduz (2023), Killi et al. (2022).

The combination of LSTM and CNN models has been proven to be effective in extracting context semantics and local features Güler and Gunduz (2023). However, despite these advances, there are still major challenges in the generalization ability of these models to diverse datasets and adapting to the evolving misinformation tactics Sharma et al. (2022).

This requires continued research into architectures that can detect nuance in discourse such as conspiracy theory patterns or maintain efficacy against increasingly sophisticated forms of deceptive content Haupt et al. (2023).

Moreover, when the characteristics of the test dataset differ significantly from the development data, the classification performance is often challenged since multiple neural network layers are required for effective training Akinyemi et al. (2020).

Models with GloVe-LSTM architectures are outperforming traditional baselines in F-scores but are limited by training data ratios and preprocessing techniques Killi et al. (2022).

These studies often require large, labeled datasets, which poses a challenge for adapting to new forms of fake news due to their reliance on extensive feature engineering Essa et al. (2023).

These models were often trained on different benchmarks such as: ISOT Dataset: 44,898 political and world news articles (2016–2017) from PolitiFact and Reuters Hosseini et al. (2023), Kholiq et al. (2022), and WELFake Dataset: A dataset with 72,134 instances with title and text columns where fake articles tend to have more subjectivity and lower readability Killi et al. (2022). The generality of these models might be restricted by their dependence on these specific datasets, and so performance can vary a lot when used on diverse sources Pierri and Ceri (2019). Such data dependency often results in inconsistent performance and biases, in which a model trained on a particular type of narratives may not perform well on different types of misinformation, calling for more robust and generalizable frameworks Sharma et al. (2022). Hence, research continues to find transfer learning methods for cross-dataset generalization and domain adaptation to boost model applicability across diverse information landscapes Sharma et al. (2022).The Table 1 summarizes the models, data sets, metrics, and limitations found in each paper. This table helps to illustrate the current state of fake news detection research.

Table 1

Table 1 Summarize Fake News Research

Study [ID]	Approach	Primary Dataset(s)	accuracy	Limitations
Kholiq et al. Kholiq et al. (2022)	LightGBM with CountVectorizer	ISOT (44.9K articles)	Nearly perfect	Optimal performance is strictly dependent on specific dataset sizes (e.g., 20,000–26,000 samples); higher loss in small sets.
Ajik et al. Ajik et al. (2023)	Optimized CNN and LSTM	Scraped News	97.15%	Overlooks the comprehensive optimization of model hyperparameters, which can significantly impact classification performance
Güler and Gunduz (2023)	CNN, LSTM, and BERT	ISOT, BuzzFeed	Demonstrated high utility	Excludes social media involvement features; faces high computational complexity for real-time detection
Hosseini et al. Hosseini et al. (2023)	Topic and Deep Variational	ISOT	80% – 91%	Relies on very specific semantics of fake news, which can generate misclassification in long-form articles
Keya et al. Keya et al. (2023)	FakeStack	Not Specified	99.74%	The hierarchical stacked architecture is highly complex, limiting practical deployment in real-world scenarios
Alghamdi et al. Alghamdi et al., 2023	BERT-base + CNN + BiLSTM	Metadata-enhanced sets	N/A	Lacks original empirical analysis and in-depth exploration of implementation challenges in actual systems
Rezaei et al. Rezaei et al. (2022)	Stacking Ensemble Network	Content-feature sets	96.24%	Accuracy drops (to 94.40%) when applied to multi-class classification and diverse fact-checking sources
Wang et al. Wang et al. (2022)	Multi-EDU Awareness	Kaggle, BuzzFeed	Demonstrated high utility	Assessment is limited by a lack of diverse performance metrics like recall and precision
D'Ulizia et al. D'Ulizia et al. (2021)	Multimodal Survey	Benchmarks (2016-2017)	N/A	Dependence on specific 2016-2017 political datasets limits the generalizability to more modern misinformation
Killi et al. Killi et al. (2022)	GloVe-LSTM	WELFake (72.1K entries)	94.4%	Struggled with effective text data processing and inconsistencies during data scraping
Sharma et al. Sharma et al. (2022)	Mitigation Survey	Global news sets	N/A	Identifies persistent model bias toward specific narratives and low adaptability to evolving misinformation tactics
Haupt et al. Haupt et al. (2023)	Qualitative Coding + ML	Conspiracy discourse	Varies by discourse	Limited scalability for high-volume detection due to the need for qualitative content coding
Akinyemi et al. Akinyemi et al. (2020)	Multi-layer Neural Nets	Scraped Social Media	High performance utility	Classification performance encounters difficulties when the test dataset significantly outweighs the development dataset
Essa et al. Essa et al. (2023)	Hybrid BERT + LightGBM	Labeled datasets	98.91%	Reliance on extensive feature engineering and model sophistication poses challenges for adaptability to novel news forms
Pierri & Ceri Pierri and Ceri (2019)	False News Analysis	Propagation sets	Varies by platform	Generalizability is restricted by dependence on specific datasets, leading to inconsistent performance across diverse sources

The suggested LSTM model has a accuracy of 99.7%, which surpasses other approaches reported in literature, such as stand-alone algorithms like LSTM (99.18%), and combined DL models (maximum 98%). This is due to better data pre-processing and model configuration. However, more experiments need to be conducted to validate its performance. (Figure 1 shows comparison of model's accuracy).

Figure 1

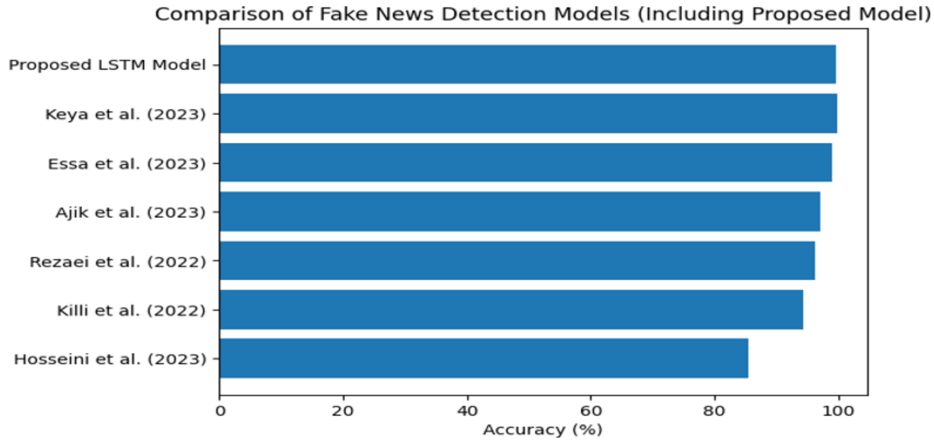

Figure 1 Comparison of Fake News Detection Models

METHODOLOGY

This methodology describes how to develop a deep learning model that can be used to distinguish between true and fake news. This model involves several stages, from data gathering to the generation of predictions. (Figure 2 illustrate methodology process flow).

Figure 2

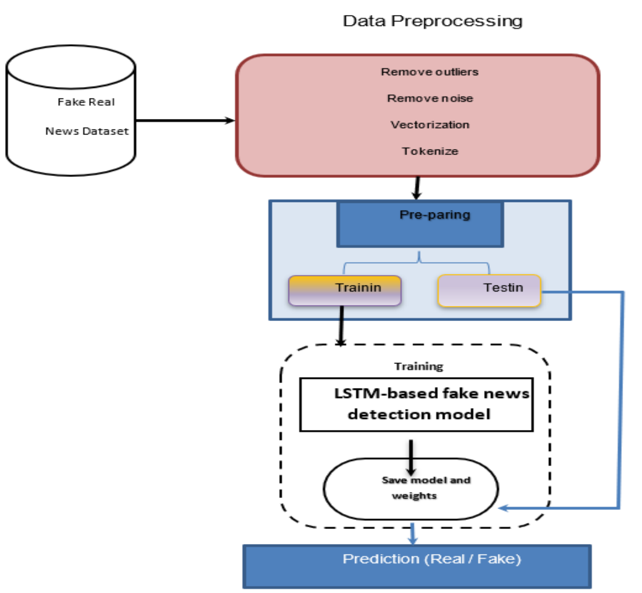

Figure 2 Methodology Process

DATA

The first step starts with a Fake and Real News Dataset, which has categorized news articles as real or fake. This forms the basis for the application of supervised learning.(Figure 3 illustrates word cloud for fake and real data), which helps understand most used words except stopwords.

Figure 3

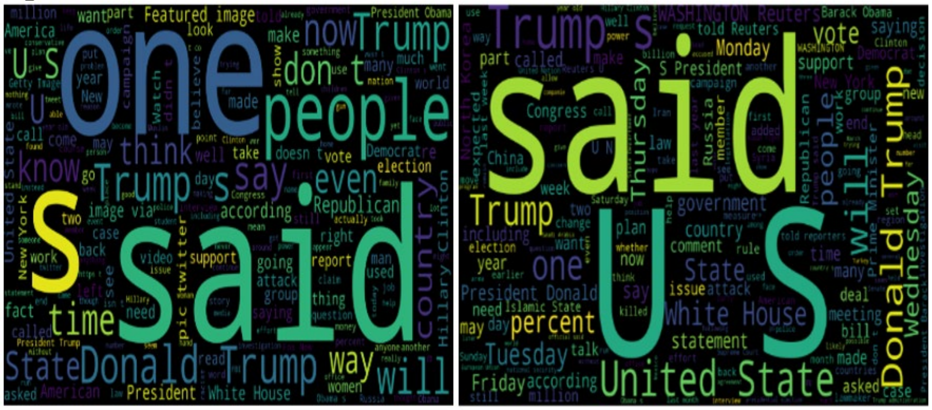

Figure 3 Methodology Process Illustrates Word Cloud for Fake and Real Data

DATA PREPROCESSING

Prior to feeding the data into the model, there are some preprocessing techniques that are performed for better results: Outlier Detection: Getting rid of outliers or unnecessary data that might have a bad impact on the model. Noise Detection: Removing any noise from the text such as unnecessary characters or symbols. Tokenization: Splitting text into smaller units (words or tokens). Vectorization: Converting textual data into numerical representations (e.g., TF-IDF or word embeddings) so it can be processed by machine learning algorithms. Distribution

TRAIN-TEST SPLIT

Dataset is split into two subset as follows: Training Dataset: The training dataset is used for training purposes only. Testing Dataset: Testing dataset will be used for testing the model.

DATA ANALYSIS AND RESULTS

MODEL TRAINING (LSTM-BASED FAKE NEWS DETECTION MODEL)

The LSTM framework utilized architecture that includes Embedding, 128 neurons in the layers, and a sigmoid activation function in classifying the data, optimizer Adam, and accuracy metric for measuring the capability of the model.

MODEL SAVING

After completing the process of training, the trained model and its weights can be stored. This means that there would be no need for the process of retraining when deploying the model.

PREDICTION

At last, the saved model is used for predictions. Preprocessing processes new input news and the saved model decides whether the input news is real or fake.

MODEL EVALUATION

To determine the performance of the recommended fake news detection approach using LSTM, various evaluation parameters have been utilized. This ensures that not only is the model highly accurate but also highly reliable in terms of detecting fake news. Accuracy is the parameter used to measure the correctness of the model, as explain in equation bellow:

$$accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (1)$$

Precision will be used to evaluate the number of actual fake news detected among those predicted by the model, , as explain in equation bellow:

$$\text{Precision} = \frac{TP}{TP+FP} \tag{2}$$

Recall indicates the ability of the model to find out fake news, as explain in equation bellow:

$$\text{Recall} = \frac{TP}{TP+FN} \tag{3}$$

F1-Score is the harmonic mean of the above metrics, as explain in equation bellow:

$$\text{F1} = 2 * \frac{\text{Precision}*\text{Recall}}{\text{Precision}+\text{Recall}} \tag{4}$$

where: True Positive = Fake News detected as Fake ,True Negative = Non-Fake News detected as Non-Fake ,False Positive = Non-Fake News detected as Fake ,and False Negative = Fake News detected as non-fake. (Figure 4 explain confusion matrix), (Figure 5 show accuracy), and (Figure 6 loss function).

Figure 4

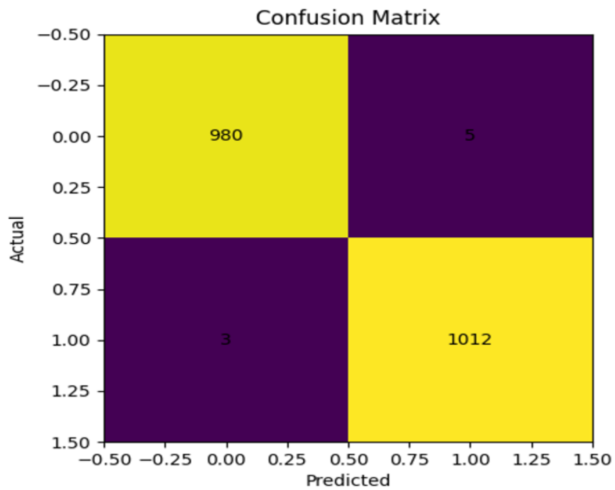

Figure 4 Explain Confusion Matrix.

Figure 5

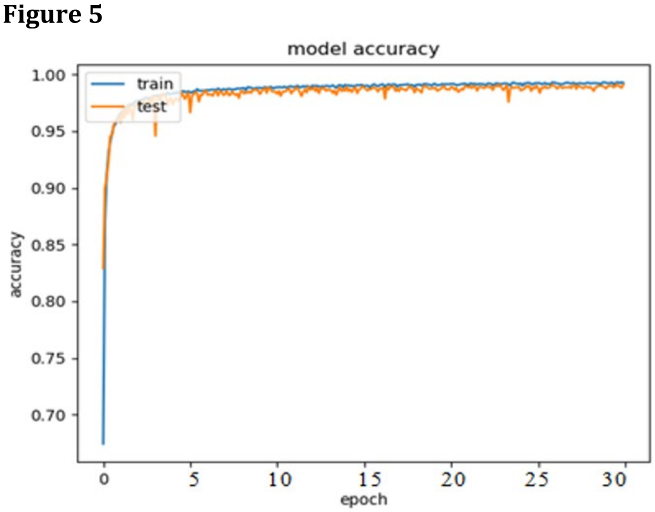

Figure 5 Model Accuracy.

Figure 6

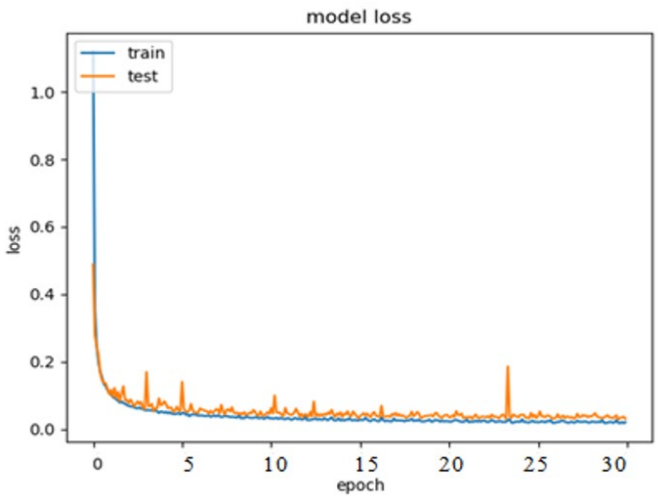

Figure 6 Model Loss.

CONCLUSION AND RECOMMENDATIONS

CONCLUSION

This study presented an effective approach for fake news detection using a Long Short-Term Memory (LSTM) model. Methodology covered data preprocessing, feature extraction, training the model, and evaluating its performance with various metrics. As a result, the proposed approach reached very high metrics, having 99.7% accuracy, Precision: 99.49%, Recall: 99.69%, and F1-Score: 99.59% which means it has excellent potential in recognizing the difference between fake and genuine news. Comparison with other works shows that traditional machine learning algorithms (SVM, Random Forest) usually provide slightly worse accuracy from 80% to 92%, while using deep learning and hybrid methods can increase this metric to 95%-99%. Even more complex algorithms like transformers, ensembles, etc. provide accuracy up to 99.3%. In this case, despite the use of simple features, our proposed LSTM-based model shows better results than most other algorithms. Although the model demonstrates great accuracy, certain limitations cannot be overlooked. First, performance might vary depending on the quality and structure of the dataset used. Besides, overfitting can be observed with more complex datasets and the use of highly sophisticated algorithms. In addition, advanced approaches often use metadata and multimodal features in their algorithms.

RECOMMENDATION

For future research in this area, it is suggested to test the effectiveness of this model on other and varied datasets, as well as include more features such as metadata and images. Furthermore, it is advisable to consider hybrid models, incorporating the use of transformers alongside LSTM models.

DATA AVAILABILITY

The study's dataset is accessible to the general public on the Kaggle platform. It is available at [https://www.kaggle.com/datasets/clmentbisaillon/fake-and-real-news-dataset?select=Fake.csv].

The dataset contains labeled news articles categorized as real and fake.

All preprocessing steps applied in this study are described in the methodology section, and the processed data can be made available upon reasonable request.

ACKNOWLEDGMENTS

None.

REFERENCES

Ajik, E. D., Obunadike, G. N., and Echobu, F. O. (2023). Fake News Detection Using Optimized CNN and LSTM Techniques. Journal of Information Systems and Informatics, 5(3), 1044–1057. https://doi.org/10.51519/journalisi.v5i3.548

Akinyemi, B. O., Adewusi, O., and Oyebade, A. (2020). An Improved Classification Model for Fake News Detection in Social Media. International Journal of Information Technology and Computer Science, 12(1), 34–43. https://doi.org/10.5815/ijitcs.2020.01.05

Alghamdi, J., Luo, S., and Lin, Y. (2023). A Comprehensive Survey on Machine Learning Approaches for Fake News Detection. Multimedia Tools and Applications, 83(17), 51009–51067. https://doi.org/10.1007/s11042-023-17470-8

D'Ulizia, A., Caschera, M. C., Ferri, F., and Grifoni, P. (2021). Fake News Detection: A Survey of Evaluation Datasets. PeerJ Computer Science, 7. https://doi.org/10.7717/peerj-cs.518

Essa, E., Omar, K., and Alqahtani, A. (2023). Fake News Detection Based on a Hybrid BERT and LightGBM Models. Complex and Intelligent Systems, 9(6), 6581–6592. https://doi.org/10.1007/s40747-023-01098-0

Güler, G., and Gunduz, S. (2023). Deep Learning Based Fake News Detection on Social Media. International Journal of Information Security Science, 12(2), 1–21. https://doi.org/10.55859/ijiss.1231423

Haupt, M. R., Chiu, M., Chang, J., Li, Z., Cuomo, R., and Mackey, T. K. (2023). Detecting Nuance in Conspiracy Discourse: Advancing Methods in Infodemiology and Communication Science With Machine Learning and Qualitative Content Coding. PLOS ONE, 18(12). https://doi.org/10.1371/journal.pone.0295414

Hosseini, M., Sabet, A. J., He, S., and Aguiar, D. (2023). Interpretable Fake News Detection With Topic and Deep Variational Models. Online Social Networks and Media, 36, 100249. https://doi.org/10.1016/j.osnem.2023.100249

Keya, A. J., Shajeeb, H. H., Rahman, Md. S., and Mridha, M. F. (2023). FakeStack: Hierarchical Tri-BERT-CNN-LSTM Stacked Model for Effective Fake News Detection. PLOS ONE, 18(12). https://doi.org/10.1371/journal.pone.0294701

Kholiq, M. H. R., Wiranto, W., and Sihwi, S. W. (2022). News Classification Using Light Gradient Boosted Machine Algorithm. Indonesian Journal of Electrical Engineering and Computer Science, 27(1), 206–213. https://doi.org/10.11591/ijeecs.v27.i1.pp206-213

Killi, C. B. R., Balakrishnan, N., and Rao, C. S. (2022). Classification of Fake News Using Deep Learning-Based GloVE-LSTM Model. International Journal of Safety and Security Engineering, 12(5), 631–637. https://doi.org/10.18280/ijsse.120512

Pierri, F., and Ceri, S. (2019). False News on Social Media. ACM SIGMOD Record, 48(2), 18–27. https://doi.org/10.1145/3377330.3377334

Rezaei, S., Kahani, M., Behkamal, B., and Jalayer, R. (2022). Early Multi-Class Ensemble Based Fake News Detection Using Content Features. Research Square. https://doi.org/10.21203/rs.3.rs-2217954/v1

Sharma, K., Qian, F., Jiang, H., Ruchansky, N., Zhang, M., and Liu, Y. (2022). Combating Fake News: A Survey on Identification and Mitigation Techniques. arXiv. https://doi.org/10.48550/arxiv.1901.06437

Wang, Y., Wang, L., Yang, Y., and Zhang, Y. (2022). Detecting Fake News by Enhanced Text Representation With Multi-EDU-Structure Awareness. Expert Systems With Applications, 206, 117781. https://doi.org/10.1016/j.eswa.2022.117781