Machine Learning and Transformers Model for Accurate Fake News Classification

Omar Sedqi Kareem

Department of Public Health, College of Health and Medical Technology - Shekhan, Duhok Polytechnic University, Duhok, Iraq

Abstract: *The rapid spread of misinformation across digital platforms poses significant threats to social trust, policymaking, and public well-being, highlighting the necessity of accurate and scalable detection frameworks. This study presents a comparative evaluation of traditional machine learning classifiers, ensemble methods, deep neural architectures, and Transformer-based models for fake news detection across three widely used benchmark datasets: ISOT, Jruvika, and Gossipcop. A rigorous preprocessing pipeline was applied, including text normalization, tokenization, stopword removal, and lemmatization, to ensure high-quality feature representations. Models were assessed using accuracy, precision, recall, F1-score, and AUC, offering a comprehensive performance profile. Results show that while ensemble methods, particularly XGBoost, achieved strong outcomes (e.g., 99.86% accuracy and 99.75% F1-score on ISOT), and the Residual Neural Network improved generalization by capturing complex linguistic dependencies, (F1 = 97.67% on Jruvika), the DistilBERT-based model consistently achieved state-of-the-art performance, with 99.92% accuracy and 99.93% F1-score on ISOT, 99.50% accuracy and 99.47% F1-score on Jruvika, and 85.37% accuracy with 90.73% F1-score on Gossipcop. These findings confirm the superiority of contextualized Transformer embeddings for robust, scalable, and efficient detection of fake news.*

Keywords: Fake News Detection, Residual Neural Network, Ensemble Learning, Text Classification, DistilBERT.

1. INTRODUCTION

The concept of fake news generally refers to the dissemination of false or misleading information through informal or unofficial communication platforms, such as social media networks (e.g., Twitter, Instagram, Facebook). Within this framework, phenomena such as disinformation, misinformation, clickbait, and other deceptive content represent different manifestations of fake news that increasingly shape and distort the information landscape on the internet [1]. To combat this trend, many fact-checking organizations are investing extensive resources into addressing misinformation. However, the fact-checking approach, which relies predominantly on domain experts to manually verify the accuracy of news articles, has proven ineffective against the volume of fake news and the speed at which it is produced. Therefore, it is of utmost importance to develop accurate automated fake news detection systems [2].

The first approaches to develop automated detection of fake news are based on the hand engineering of large feature sets on a multitude of factors, the content of the news, linguistic features, i.e., syntax, grammar, and usage of word choice (language); profiles of users; and the propagation paths of news sources. These feature sets are then used to train Machine Learning (ML) classifiers to determine the degree of truthfulness of the content [3]. Several recent attempts have integrated graph theory with deep learning approaches, with these researchers primarily concerned with addressing the limitations of traditional deep learning models, which largely consist of their ability to exploit relational and structural information from the text provided. A number of recent studies presented several variants of a graph neural network approach to cope with the challenges of traditional techniques in deep learning [4].

An extensive literature review was conducted to identify a suitable approach for applied models of ML in news, such as hate/fake news. Previous studies have explored various techniques, and their advantages and disadvantages have been discussed. Significantly, Reddy et. al. [5] proposed a stacked ensemble of CNN, BiLSTM, and MLP models trained with GloVe and BERT embeddings. It achieved commendable performance with F1-scores as high as 96.03% and each model does not achieve as high a performance on other datasets. Its strengths include the use of transformers and word embeddings with ensemble learning. However, the limitations are notable: it relies on the dataset size, focuses on binary text classification (with only two classes), and neglects the use of multimodal embeddings and explainability. Jouharet. et. al. [6] examined a limited number of machine-learning models on the ISOT dataset. Among these models, XGBoost had the best performance across accuracy, precision, recall, F1-score, and ROC-AUC, exhibiting strong generalization with little evidence of overfitting. While the study employed several classifiers and a diverse range of evaluation metrics, all the results were obtained from a single textual dataset, and there was no hyperparameter tuning or sophisticated use of embeddings. The study [7] employed DistilBERT on the Jruvika dataset, attaining 97.7% accuracy and F1-scores of 0.97–0.98, while demonstrating faster inference and lower computational demands compared to BERT. Although it shows strong efficiency and performance, its reliance on a single dataset raises concerns of overfitting and limited generalizability to more complex NLP tasks. Folino et al. [1] proposed a semi-supervised ensemble of lightweight BERT models that incorporates self-training, pseudo-labeling, and uncertainty-based selection, achieving an F1-score of 85 on benchmark datasets. The approach is efficient with limited labeled data but may falter with weak initial classifiers, suggesting the need for integration with few-shot or active learning strategies. This paper [8] presented ABERT, a parameter-efficient BERT variant for fake news detection that cuts trainable parameters by 67.7% while maintaining performance, achieving 85.79% accuracy on GossipCop and 91.98% on PolitiFact. It offers scalability and reduced computational cost, though limited interpretability and the absence of social context remain challenges. This paper makes the following key contributions:

1. **Comprehensive Evaluation:** A systematic comparison of classical machine learning models, ensemble methods, a Residual Neural Network, and a DistilBERT-based Transformer across three benchmark datasets (ISOT, Jruvika, Gossipcop).
2. **Robust Preprocessing Pipeline:** Implementation of advanced text preprocessing (normalization, tokenization, lemmatization) to ensure reliable and consistent feature representations.
3. **Empirical Findings:** Demonstration that while ensemble methods such as **XGBoost** achieve strong results (**99.86% accuracy on ISOT**), the **DistilBERT-based model consistently sets a new benchmark**, achieving **99.92% accuracy on ISOT, 99.50% on Jruvika, and 85.37% on Gossipcop**.
4. **Significance:** The study establishes that **contextualized Transformer embeddings outperform traditional and neural baselines**, ensuring both accuracy and robustness on diverse datasets, including noisy real-world corpora.

The paper is organized as follows: Section 2 introduces the datasets and preprocessing pipeline and presents the proposed methodology and models; Section 3 outlines the experimental setup and evaluation metrics and discusses the results and comparison with prior studies; and Section4 concludes with key findings and future directions.

2. Materials and Methods

2.1. Dataset

This study employs three benchmark datasets for fake news detection: ISOT, Jruvika, and Gossipcop. The ISOT [9] dataset consists of 44,898 articles, including 23,481 fake and 21,417 true news reports, each with title, text, subject, and date information. The Jruvika [10] dataset contains 4,009 labeled instances of real and fake news, widely used in recent deep learning studies such as DistilBERT-based detection. The Gossipcop [11] dataset comprises 22,140 news articles, with 5,323 labeled as fake and 16,817 labeled as true, providing a larger and more diverse source of misinformation and fact-based reporting. While ISOT is balanced across classes, Jruvika is relatively small, and Gossipcop exhibits a noticeable imbalance, with a higher proportion of true articles than fake ones. These datasets collectively provide a robust basis for evaluating fake news detection models across different scales and class distributions, as shown in Table 1.

Table 1. Fake News Datasets Summary

Dataset	Total Articles	Fake	TRUE
ISOT	44898	23481	21417
Jruvika	4009	2,137	1,872
Gossipcop	22140	5323	16817

2.2. Preprocessing Methodology

Before training fake news detection models, rigorous data preprocessing was performed to ensure quality, reduce noise, and enhance model generalization. The initial step involved removing non-textual attributes, specifically the date and subject columns, as they do not contribute semantically to news veracity classification and could introduce bias [12]. Next, we conducted a data integrity check by identifying null values and duplicate entries. While no nulls were present, 5,793 duplicates were detected; these were eliminated to prevent overfitting and data leakage, yielding a final cleaned corpus of 39,105 articles. To unify the textual input, we created a new statement column by concatenating the title and text fields:

$$S=T \parallel B \tag{1}$$

where T is the news title, B is the body, and ‖ denotes string concatenation.

Following dataset restructuring, the text was normalized through a multi-stage linguistic cleaning pipeline. First, all tokens were converted to lowercase, ensuring uniformity and reducing vocabulary sparsity (e.g., Trump vs trump)[13]. Next, punctuation and numeric characters were stripped using regex-based cleaning, preserving only meaningful words. The clean text was then tokenized, segmenting each statement into individual words $\{w1,w2,\ldots,wn\}$. To reduce redundancy, we removed stopwords (e.g., the, and, is), which are frequent but carry limited discriminative value. Finally, lemmatization was applied using the WordNet lemmatizer to reduce words to their base form, e.g., running → run, better → good, thereby reducing inflectional variation [14]. The full transformation can be formalized as:

$$S{*}=\text{Lemmatize (RemoveStop (Tokenize (Clean(S))))} \tag{2}$$

where Clean(·) denotes lowercasing and punctuation/number removal, Tokenize(·) splits into tokens, RemoveStop(·) filters out stopwords, and Lemmatize(·) maps each token to its canonical lemma.

Finally, the processed dataset was split into a 70% training subset and a 30% testing subset using stratified sampling to preserve the class distribution. This ensures balanced representation of fake (0) and true (1) classes across splits, reducing sampling bias and improving the robustness of model evaluation.

2.3. Machine Learning Algorithm

To evaluate the effectiveness of various supervised learning techniques for detecting fake news, several classical and ensemble-based classifiers were considered. These models differ in their learning principles, interpretability, and ability to handle non-linear patterns. Below, a brief explanation is provided for each classifier, highlighting its main contribution to the detection task.

2.3.1. Decision Tree: A Decision Tree is a non-parametric model that learns decision rules from features in the form of a tree structure. At each internal node, the dataset is split based on the feature that maximizes information gain or minimizes Gini impurity, resulting in a leaf node that represents the predicted class. Its key advantage is interpretability, as the decision path can be easily traced. However, a single decision tree is prone to overfitting [15].

2.3.2. Random Forest: The Random Forest is an ensemble method that builds multiple decision trees on randomly sampled subsets of the training data and aggregates their predictions through majority voting. This reduces overfitting and variance compared to a single decision tree, making it a more robust approach for text classification tasks. Its ability to handle high-dimensional data makes it effective in fake news detection [16].

2.3.3. Gradient Boosting: Gradient Boosting constructs a strong classifier by sequentially adding weak learners (usually shallow decision trees), where each new tree corrects the errors of the previous ensemble by optimizing a gradient descent-based loss function. This stepwise refinement makes gradient boosting powerful for capturing complex patterns, though it can be computationally expensive and sensitive to hyperparameters.

2.3.4. Passive Aggressive Classifier: The Passive Aggressive Classifier is particularly suited for online learning. It updates its weights only when it misclassifies a sample (an aggressive step), while leaving the weights unchanged when predictions are correct (a passive step).

This makes it efficient for large-scale, real-time text streams, such as news feeds. Its margin-based updates enable it to adapt quickly, although stability may be compromised with noisy data.

2.3.5. XGBoost: XGBoost is an advanced implementation of gradient boosting optimized for speed and performance. It introduces regularization to prevent overfitting and utilizes parallelized tree construction, efficiently handling missing values and sparse features. XGBoost is widely used in competitions and research due to its high accuracy and scalability, making it a strong candidate for detecting fake news.

2.3.6. Logistic Regression: Logistic Regression is a linear model that estimates the probability of a class label using the logistic (sigmoid) function. Despite its simplicity, it performs remarkably well in high-dimensional text classification tasks when combined with feature extraction methods such as TF–IDF. Its interpretability and efficiency make it a strong baseline for fake news detection, although it may underperform in cases with complex, non-linear relationships, as shown in Figure 1.

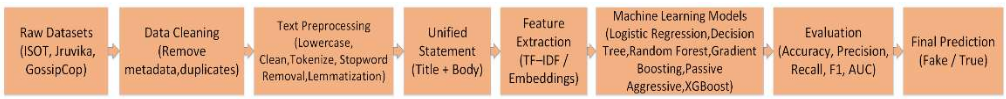

Figure 1. Proposed Fake News Detection Pipeline.

2.4. Residual Neural Network

The proposed model is a Residual Neural Network (ResNN) specifically designed for binary text classification in fake news detection. The input consists of high-dimensional vectors generated from feature extraction methods such as TF–IDF or embeddings. The network begins with an initial dense projection of 512 neurons, followed by batch normalization, ReLU activation, and dropout regularization to stabilize learning and prevent overfitting. Two residual blocks are incorporated, where skip connections reintroduce the input into the transformed output, thereby mitigating vanishing gradients and enabling deeper representations. After residual learning, the architecture transitions into bottleneck layers (256 and 128 neurons), compressing features into a more compact and discriminative representation. Each hidden layer is coupled with normalization, activation, and dropout to improve robustness. Finally, a softmax output layer with two neurons produces probabilities for fake and true news, making the model suitable for categorical cross-entropy optimization, as shown in Figure 2.

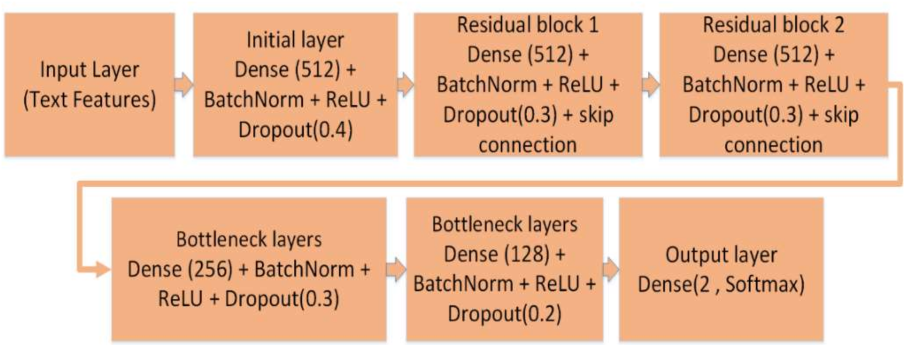

Figure 2. Residual Neural Network (ResNN) for Fake News Detection.

The network is trained using the Adam optimizer with a decaying learning rate to strike a balance between convergence speed and stability. To prevent overfitting, early stopping is applied by monitoring validation loss with a patience of six epochs, restoring the best weights once training halts. Class labels are one-hot encoded, aligning with the softmax output format to ensure probability-based classification. The model is trained with a batch size of 64 for computational efficiency, and evaluation metrics include accuracy, precision, recall, and AUC, providing a holistic assessment of classification performance. This integrated design — combining residual learning, normalization, dropout regularization, and adaptive training strategies — enables the ResNN to effectively capture semantic patterns while maintaining generalization, making it a strong candidate for detecting fake news across diverse datasets.

2.5. DistilBERT-base-uncased Model

This study employs DistilBERT-base-uncased, a lightweight Transformer with approximately 66M parameters, for binary fake news detection. DistilBERT retains BERT's bidirectional self-attention but compresses the architecture through knowledge distillation, yielding faster training and inference while maintaining strong representational power. A two-neuron classification head is added for probability estimation over fake and true labels, making the model efficient and effective for large-scale text classification. Text is tokenized using the WordPiece tokenizer, truncated to 256 tokens, and dynamically padded to minimize computational overhead. Labels are standardized and mapped to class IDs to ensure compatibility with evaluation metrics. Training is performed using the Trainer API, with a batch size of 16, a learning rate of 2×10^{-5}, a weight decay of 0.01, and up to five epochs, incorporating early stopping. Model selection is guided by the F1-score, ensuring balanced performance between fake and true classes. The proposed setup combines efficiency and robustness by leveraging a distilled Transformer with task-specific refinements, including dynamic padding, warmup scheduling, mixed precision, and F1-driven early stopping. Unlike heavier Transformer models, this configuration emphasizes practical deployment by reducing resource costs without compromising accuracy, offering a novel balance of performance and efficiency for misinformation detection tasks.

3. Results and Discussion

The proposed fake news detection models were evaluated on ISOT, Jruvika, and Gossipcop datasets using accuracy, precision, recall, F1-score, and AUC. Experiments were conducted on a system equipped with an Intel Core i5 CPU, 7th Generation, and 8GB of RAM, implemented in Python using Transformers, TensorFlow/Keras, and Scikit-Learn. Classification outputs were divided into Fake News and True News, and results were compared across classical machine learning classifiers, a Residual Neural Network, and a DistilBERT-based Transformer.

3.1. Evaluation and Analysis of Results

On the ISOT dataset, all models achieved high accuracy (>98%), with XGBoost (99.68%) and Gradient Boosting (99.55%) outperforming classical baselines. The Residual Neural Network performed strongly (99.02%), but the Transformer (DistilBERT) achieved near-perfect results (99.92% accuracy, F1 = 99.93, AUC = 1.0), confirming its superiority in capturing contextual features, as shown in Table 2.

Table 2. Performance Metrics of Fake News Detection Models on ISOT Dataset.

Algorithm	Accuracy %	Precision %	Recall %	F1-scor %	ROC AUC %
Decision Tree	99.44	99.33	99.65	99.49	99.42
Random Forest	98.39	97.84	99.24	98.54	98.31
Gradient Boosting	99.55	99.40	99.78	99.59	99.53
Passive Aggressive Classifier	99.36	99.16	99.67	99.41	99.32
XGBoost	99.68	99.57	99.84	99.71	99.66
Logistic Regression	98.48	98.18	99.04	98.61	98.42
RNN	99.02	98.78	99.43	99.10	99.86
DistilBERT	99.92	99.95	99.91	99.93	1.00

For the Jruvika dataset, performance differences were more pronounced. While Decision Tree underperformed (94.48%), ensemble methods and Passive Aggressive Classifier (98.16%) improved outcomes. Both NN (97.91%) and XGBoost (97.66%) were competitive, but DistilBERT again dominated (99.50% accuracy, F1 = 99.47, AUC ≈ 1.0), as shown in Table 3.

Table 3. Performance Metrics of Fake News Detection Models on Jruvika Dataset.

Algorithm	Accuracy %	Precision %	Recall %	F1-scor %	ROC AUC %
Decision Tree	94.48	94.31	93.25	93.78	94.36
Random Forest	96.99	94.30	99.25	96.71	97.21
Gradient Boosting	97.32	96.13	97.94	97.03	97.38
Passive Aggressive Classifier	98.16	97.58	98.31	97.94	98.18
XGBoost	97.66	96.85	97.94	97.39	97.68
Logistic Regression	96.57	95.06	97.37	96.20	96.65
RNN	97.91	97.04	98.31	97.67	99.81
DistilBERT	99.50	99.47	99.47	99.47	99.98

The Gossipcop dataset proved more challenging, with lower overall scores. Classical ensembles such as Random Forest (90.27% F1) and Logistic Regression (90.61% F1) performed well, while Gradient Boosting achieved very high recall (98.33%) at the cost of precision. The NN showed balanced performance (89.85% F1, AUC = 85.81), but DistilBERT achieved the best overall balance (85.37% accuracy, 90.73% F1, AUC = 87.33), highlighting its robustness on noisy data, as shown in Table 4.

Table 4. Performance Metrics of Fake News Detection Models on Gossipcop Dataset.

Algorithm	Accuracy %	Precision %	Recall %	F1-scor %	ROC AUC %
Decision Tree	79.05	85.65	86.99	86.32	70.48
Random Forest	84.44	85.94	95.06	90.27	72.98
Gradient Boosting	81.87	81.58	98.33	89.17	64.10
Passive Aggressive Classifier	78.68	85.52	86.58	86.05	70.15
XGBoost	83.64	84.22	96.55	89.97	69.72
Logistic Regression	84.76	85.11	96.88	90.61	71.67
RNN	83.67	85.08	95.20	89.85	85.81
DistilBERT	85.37	87.47	94.23	90.73	87.33

In summary, while ensemble methods provided strong baselines and NN improved generalization, the Transformer-based model consistently delivered state-of-the-art results, confirming its effectiveness for reliable fake news detection.

3.2. DISCUSSION

To validate the robustness of the proposed framework, its performance was compared with that of several state-of-the-art studies that applied machine learning and deep learning approaches for fake news detection across multiple benchmark datasets. Table 5 summarizes the findings. Recent research Reddy et. al. [5], [6] proposed a logistic regression–based stacked ensemble evaluated on ISOT, PolitiFact, and Gossipcop, achieving an F1-score of 96.03% but with inconsistent precision and recall. Another study Jouharet. et. al. [6] applied XGBoost to the ISOT dataset, achieving an impressive 99.86% accuracy and 99.75% F1-score, demonstrating the strength of boosted ensembles for structured text representations. A more recent investigation Amandeep et. al. [7] employed DistilBERT, a base-cased distilled version of SQuAD, on the Jruvika dataset, reporting 97.7% accuracy and a 97% F1-score. While competitive, these results indicate that lighter Transformers require task-specific fine-tuning to reach optimal performance. Similarly, a semi-supervised BERT-based approach Folino et al. [1] applied to PolitiFact and Gossipcop obtained 84% accuracy and 85% F1-score, but performance degraded on noisy datasets, illustrating the limitations of semi-supervised fine-tuning in diverse domains. Alghamdi et. al. [8] introduced ABERT, a lightweight BERT adaptation, which achieved 85.79% accuracy and 84.34% F1-score on Gossipcop, reflecting gains in efficiency but at the cost of reduced generalization.

In contrast, the proposed model demonstrated state-of-the-art performance across all three datasets. On ISOT, it achieved 99.92% accuracy and 99.93% F1-score, surpassing both classical and ensemble baselines. On Jruvika, it reached 99.50% accuracy and 99.47% F1-score, outperforming the previously reported DistilBERT model by nearly two percentage points. On the more challenging Gossipcop dataset, it delivered 85.37% accuracy, 94.23% recall, and 90.73% F1-score, representing a substantial improvement over ABERT and semi-supervised BERT. These results highlight the novelty of the proposed framework in balancing high accuracy on structured datasets and robust recall on noisy real-world data, a property critical for misinformation detection in practice.

Table 5. Comparison of the Proposed Model with Existing Studies on Fake News Detection.

Ref.	Year	Dataset	Model	Performance %			
				Accuracy	Precision	Recall	F1-scor
Reddy et. al. [5], [6]	2024	ISOT PolitiFact Gossipcop	Ensemble Approach best one (Logistic Regression Based Stacked Ensemble)	-	-	-	96.03 61.11 74.24
Jouharet. et. al. [6]	2024	ISOT	Machine Learning Algorithm (XGBoost)	99.86	99.76	99.24	99.75
Amandeep et. al. [7]	2025	Jruvika	DistilBERT base-cased-distilled-squad	97.7	-	-	97
Folino et al. [1]	2025	PolitiFact, Gossipcop	semi-supervised DL (BERT)	84	-	-	85
Alghamdi et. al. [8]	2025	PolitiFact, Gossipcop, AI	ABERT (Gossipcop dataset)	85.79	85.82	85.79	84.34
Proposed Model	2025	ISOT Jruvika Gossipcop		99.92 99.50 85.37	99.95 99.47 87.47	99.91 99.47 94.23	99.93 99.47 90.73

4. Conclusion

This work conducted an in-depth comparative study of machine learning, ensemble, deep learning, and Transformer-based approaches for fake news detection, demonstrating the relative strengths and limitations of each. Classical models such as Decision Tree and Logistic Regression offered interpretability but yielded lower performance, with accuracies typically below 95%. Ensemble methods, particularly XGBoost, significantly improved predictive robustness, achieving 99.86% accuracy on ISOT. The Residual Neural Network further enhanced representational capacity, obtaining 99.10% F1 on ISOT and 97.67% on Jruvika, thereby confirming the utility of residual connections in handling complex linguistic structures. Nevertheless, the DistilBERT-based model emerged as the most effective, delivering near-perfect results on ISOT (99.92% accuracy, 99.93% F1) and Jruvika (99.50% accuracy, 99.47% F1), while also achieving superior balance on the more challenging Gossipcop dataset (85.37% accuracy, 90.73% F1, 94.23% recall). These outcomes establish the proposed framework as a new benchmark in fake news detection, combining state-of-the-art accuracy with robustness across diverse datasets. Future work may explore domain-adaptive pretraining and explainable AI extensions to improve generalization and interpretability in real-world deployments further.

Acknowledgments

Duhok Polytechnic University supported this research. The author gratefully acknowledges the guidance of the President of Duhok Polytechnic University and the continuous assistance provided by the university throughout this study.

REFERENCES

[1] *G. Folino, M. Guarascio, L. Pontieri, and P. Zicari, “Discovering ensembles of small language models out of scarcely labelled data for fake news detection,” Appl Soft Comput, vol. 171, p. 112794, 2025.*

[2] C.-O. Truică, E.-S. Apostol, M. Marogel, and A. Paschke, “GETAE: graph information enhanced deep neural network ensemble architecture for fake news detection,” Expert Syst Appl, vol. 275, p. 126984, 2025.

[3] T. M. H. Gedara, V. Loia, and S. Tomasiello, “A fuzzy-based multimodal approach for interpretable fake news detection,” Appl Soft Comput, p. 113277, 2025.

[4] A. Sahi et al., “SGDM-GRU: Spectral graph deep learning based Gated Recurrent Unit model for accurate fake news detection,” Expert Syst Appl, vol. 281, p. 127572, 2025.

[5] J. Reddy, S. Mundra, and A. Mundra, “Ensembling Deep Learning Models for Fake News Classification,” Procedia Comput Sci, vol. 235, pp. 2766–2774, 2024.

[6] J. Jouhar, A. Pratap, N. Tijo, and M. Mony, “Fake news detection using

python and machine learning," Procedia Comput Sci, vol. 233, pp. 763–771, 2024.

[7] S. Suresh, "Transforming Fake News Detection: Leveraging DistilBERT Models for Enhanced Accuracy," Procedia Comput Sci, vol. 260, pp. 283–290, 2025.

[8] J. Alghamdi, Y. Lin, and S. Luo, "ABERT: Adapting BERT model for efficient detection of human and AI-generated fake news," International Journal of Information Management Data Insights, vol. 5, no. 2, p. 100353, 2025.

[9] H. Ahmed, I. Traore, and S. Saad, "Detection of online fake news using n-gram analysis and machine learning techniques," in International conference on intelligent, secure, and dependable systems in distributed and cloud environments, Springer, 2017, pp. 127–138.

[10] A. Yadav, S. Gaba, H. Khan, I. Budhiraja, A. Singh, and K. K. Singh, "Etma: Efficient transformer-based multilevel attention framework for multimodal fake news detection," IEEE Trans Comput Soc Syst, vol. 11, no. 4, pp. 5015–5027, 2023.

[11] K. Shu, D. Mahudeswaran, S. Wang, D. Lee, and H. Liu, "Fakenewsnet: A data repository with news content, social context, and spatiotemporal information for studying fake news on social media," Big Data, vol. 8, no. 3, pp. 171–188, 2020.

[12] D. A. Mura et al., "Is it fake or not? A comprehensive approach for multimodal fake news detection," Online Soc Netw Media, vol. 47, p. 100314, 2025.

[13] M. Abdullah, Z. Hongying, A. Javed, O. Mamyrbayev, F. Caraffini, and H. Eshkiki, "A joint learning framework for fake news detection," Displays, p. 103154, 2025.

[14] M. A. B. Al-Tarawneh, O. Al-irr, K. S. Al-Maaitah, H. Kanj, and W. H. F. Aly, "Enhancing fake news detection with word embedding: A machine learning and deep learning approach," Computers, vol. 13, no. 9, p. 239, 2024.

[15] S. Das, R. Kumari, and R. K. Singh, "Detection of Fake News by Convolutional Neural Networks and Recurrent Neural Networks," Procedia Comput Sci, vol. 258, pp. 2278–2289, 2025.

[16] M. rao Kondamudi and S. ranjan Sahoo, "Integrating Explainable AI with Enhanced Ensemble Models for Accurate and Transparent Fake News Detection in OSN's," Procedia Comput Sci, vol. 258, pp. 1081–1090, 2025.