AOI :10.100.234513.0226
13th international conference on Microelectronics, Circuits and Systems, Micro2026

A Comparative Study of Machine Learning, Deep Learning, and Transformer-Based Models for Fake News Detection

Gurudev Kumar, Chaman Praveen, Dr. C. Jayaprakash, Dr. Om Prakash

SCSE, Galgotias University, Greate Noida, India
Corresponding Author: Gurudev.23SCSE1010335@galgotiasuniversity.ac.in

ABSTRACT

The BERT model shows a steady decrease in both training and validation loss, as well as an improvement in accuracy. This shows that the model is converging and learning the context effectively, without any signs of overfitting. The digital news media and social media have significantly changed the manner in which people consume information. Although the digital news media has improved the dissemination of information, it has also resulted in an increase in the dissemination of false information. False information is disseminated rapidly due to its sensational headlines and the fact that it is disseminated without verification. Proposed in this paper is a comprehensive text-based fake news detection system using machine learning and deep learning techniques. Several traditional machine learning classifiers are experimented with TF-IDF feature extraction techniques, including Naive Bayes, Logistic Regression, Support Vector Machines, Random Forest, and SGD Classifier. In addition to this, several deep learning models like Convolutional Neural Networks (CNN), Bidirectional LSTM (BiLSTM), and a fine-tuned BERT transformer model are also employed for semantic feature extraction of the text. The performance of all models is compared using standard evaluation metrics such as accuracy, precision, recall, and F1-score. The experimental results show that the transformer models are performing better than the traditional models because of their understanding of the context, and the traditional models are performing well with less computational complexity. The importance of the combination of classical and modern approaches in NLP for effective misinformation detection is the focus of this paper.

I. INTRODUCTION

In the contemporary digital age, information sharing strategies for news delivery primarily vary from conventional sources to digital media and information networking sites where news can spread at an instantaneous rate [1]. The swift accessibility of information via digital media enables news to spread to millions of people within minutes; nonetheless, it can equally allow uncontrolled spreading of misleading news [2]. Many people receive information via digital media and subsequently forward it without ascertaining its reliability, resulting in widespread unscathed Spread of misleading news [1], [2].

First of all, fake news has serious implications for society with adverse effects such as political manipulation [3], misinformation in case of public health crises [4], and eroded public trust [5]. Such news is characterized by the promotional use of language that is emotive in nature, along with misleading headings of news articles, such that it becomes very difficult for individuals or society in general to identify what is real news or what is fabricated news [6]. Due to such issues with traditional methods of fact-checking, which is neither efficient nor feasible in coping with such high volumes of digital information [7], fake news has emerged as an essential requirement of contemporary information systems [5].

From a computational point of view, machine learning and natural language processing methodologies provide remarkably active solutions for text analysis in large volumes [1]. Classical machine learning models rely on the statistical representation of text data, for instance, TF-IDF, in order to capture linguistically informative patterns, while deep learning models capture semantic and sequential dependencies in texts [6]. However, very recently, the performance of self-attention-based transformer model architectures like BERT is providing better results with a more accurate detection of contextual relationships than any other model [8] [7]. Therefore, in this paper, the performance of traditional machine learning model-based, deep learning model-based, and transformer model-based text-based fake news detection approaches is being investigated in search of a better balance of their efficiency, scalability, and accuracy [9] [3].

- Using a text-based approach, an evaluation of classical machine learning, deep learning models, transformers, and fake news detection is proposed.
- A united architecture where it combines TF-IDF machine learning classifiers with both sequence-based deep learning architectures and transformers built upon the BERT architecture.
- Superior levels of achievement can also be achieved using self-attention-based BERT, as compared to ML and DL.
- It attained an accuracy of 96% and 99% accuracy.

ISBN: 978-81-985770-0-9

The rest of the paper is organized as follows: Section 2 summarizes related work for Fake News Detection techniques. Section 3 followed by datasets and media processing. Section 4 section delves into the proposed model. The results and analysis are discussed in Section 5. Finally, the conclusion and future direction are in Section 6.

II. LITERATURE REVIEW

The problem of detecting fake news has gained popularity in recent times, especially with the alarming rate of misinformation dissemination on Internet media and social media sites. Generally, previous studies relied on conventional machine learning architectures such as Naive Bayes, Support Vector Machines, Logistic Regression, and Random Forest with handcrafted features like Term Frequency–Inverse Document Frequency, n-grams, sentiment features, etc. All these approaches were proven to be robust in terms of efficiency but failed to capture intrinsic semantic meaning and contextual dependencies in text [1] [2].

Recently, studies have indicated the potential of attention mechanism in collaboration with deep learning models. Attention mechanism is able to focus the attention of the model on more semantically important words. This has also helped to improve the contextual meaning of the given data, thereby facilitating more precise classification. Architectures such as multi-head attention, co-attention, and graph-aware attention networks have achieved higher performance in the classification of fake news compared to traditional CNN–LSTM architecture in uni-modal as well as multi-modal fake news detection scenarios [10] [11].

Detection of fake news in multimodal has also emerged as one of the critical directions, in which textual information is combined with visual and social context features. Various studies have shown that learning from multimodal techniques improves accuracy in event detection compared to text-only approaches by capturing complementary information from different modalities [12] [13] [14]. However, there are a number of obstacles associated with these methods, which have been identified due to the considerations of data availability and domain as well as generalization problems.

Research on fake news detection has moved ahead with the advent of transformer architectures and the concept of large language models. It is true that models following the transformer architecture, particularly the adoption of models like BERT, deliver highly effective results in detecting fake news by focusing on the self-attention mechanism and the overall concept of embeddings. It is important to note here that benchmark tests prove beyond any doubt the supremacy of models following the above mechanisms while comparing them with other models following machine learning as well as traditional deep learning approaches [5] [8] [6] [7]. Domain-specific models, for example, models that address COVID-19 fake news, show really worth when put to use in real-life scenarios [4].

In general, all the surveyed literature on event extraction shows a clear evolution from feature-engineered machine learning models to deep learning and transformer-based approaches, focused on contextual and semantic understandings. For the traditional models, the superiority is seen due to their efficiency and interpretability. However, for the state-of-the-art models, the transformer-based models remain superior. The need for this comparative analysis of machine learning, deep learning, and transformer-based models is well covered by this work [9] [15] [3].

III. DATASET AND PREPROCESSING

Data in this study is based on a Kaggle collection that is accessible to the public, containing labeled articles of fake news and real news. A data set was formed by combining two data files, namely “fake news” and “genuine news.” Articles in the dataset were labeled as either 0 or 1, depending on the nature of the news.

A. Preprocessing and Feature Engineering

Text processing techniques were employed to clean the input texts. This entails the following steps: conversion of texts to lowercase, removal of unnecessary symbols from the texts, and proper tokenization of the texts. By filtering out the stop words from the texts, unnecessary noise could be removed from the texts. Next, we’re going to extract some text, where we’re actually taking the first step in doing so. This step gets the text data ready for feature extraction; in other words, it prepares the data in order to bring out features. While in traditional ‘machine learning’ methods, ‘TF-IDF’ vectorization of text into numerical vectors is a standard practice.

In unigrams and bigrams, the specific words and even minute suggestions included in short sentences were emphasized by focusing on specific words while minimizing the effect of common stop words. Once the deep learning techniques were processed, the text was tokenized and arranged in sequences with a specific length. By using word embedding, the relationships were learned.

IV. PROPOSED METHODOLOGY

The paper proposes the flexible approach toward recognizing fake news in text data using the combination of machine learning and deep learning models. The process is achieved in a stepwise manner, where the steps involve data collection, data preprocessing, feature selection, model training, and lastly, model evaluation. During preprocessing, the text data is normalized first, where the text is converted to lower case, unwanted characters removed, and tokenization is performed on

ISBN: 978-81-985770-0-9

the data. Additionally, the stop words in the data are removed in order to reduce data noise, thereby enhancing the data preprocessing step to improve the effectiveness of the model. The standard evaluation parameters are applied for measuring the effectiveness of this model.

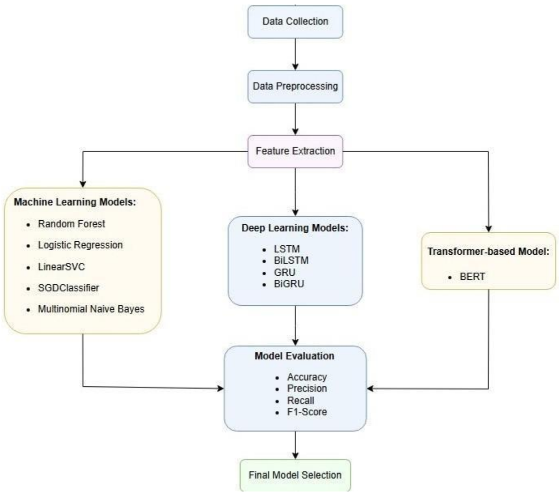

Fig. 1. System Model

A. Model Training

Multiple models were trained independently to compare their effectiveness like:

Machine Learning Models:

- **Multinomial Naive Bayes:** Naive Bayes with a twist in the form of a multinomial model is a traditional favorite in text classification since it takes advantage of available information in term frequencies. It is well-suited to work with TF-IDF and to operate well on high-dimensional sparse data, and therefore it was the natural choice for the final classification model of misinformation detection for its superior performance, simplicity, and computational efficiency.
- **Logistic Regression:** Logistic Regression is a linear predictor used to predict the likelihood of a certain case belonging to a certain class. Also, it can handle vast text data and produce more understandable results if evaluation is run.
- **Linear Support Vector Classifier:** Similar to Logistic Regression, Linear SVC is also a 'margin-based classifier', as it tries to increase the gap, also referred to as 'margin', between different classes. It usually tends to perform better with text data, as most text tends to have a very high number of elements in the feature space, and it goes very well with 'TF-IDF' representations.
- **Random Forest Classifier:** Random Forest, on the other hand, uses the bagging principle with multiple decision trees that provide the capability for enhanced prediction accuracy, reduction of overfitting, the discovery of non-linear variables, and robustness over multiple types of data.
- **SGD Classifier:** SGD refers to Stochastic Gradient Descent Classifier. This is a very lightweight classifier that is optimized for optimization tasks and can efficiently handle big data. Particular uniqueness is felt when it forms the stack of other linear classifiers like the SVM or logistic regression, whose parameters are optimized over by stochastic gradient descent.

Deep Learning Models:

- **LSTM:** A recurrent neural network for which a framework was created, one which utilizes dependencies. It deals directly with each and every part of a text sequence, using memory and gate systems.
- **Bidirectional LSTM:** Recurrent Neural Network - Reads text in both directions, both left to right as well as the other way around, in order to recognize the pattern in the text itself.

Transformer-Based Models:

- **BERT:** A transformer-based language model that learns deep contextual embeddings using self-attention. The model has the capability to deliver superior performance by focusing on the semantic-contextual interpretation of the text itself. In this research, the BERT model, similar to the previous model in the transformer group, has been fine-tuned to access the contextual embeddings provided in the pre-trained model.

Each of the models was tested on the held-out, where the accuracy of each model was evaluated by computing the accuracy, precision, recall, and F1-score. These three measures collectively offer a comprehensive feel of the comprehensiveness with which the model separates the real and the fake news. After experimenting with the various alternatives, the final classifier model chosen is the transformer-based model, which is a BERT model. This model performed superiorly in terms of accuracy, precision, recall, and F1-score compared to the performance of the other models. Its power lies in its semantic capacity in understanding the context of news articles, helping forward the accurate differentiation of news as either real or fake in nature. We have saved, together with its tokenizer, the fine-tuned model of BERT. In the case of the actual deployment of the model, the news text inputted will be tokenized using the pre-trained tokenizer to generate the appropriate predictions on the real vs. fake news using the pre-trained model.

V. RESULTS AND ANALYSIS

The experimental results show that there are obvious differences in performance among the models tested. The traditional machine learning models based on TF-IDF features performed well as baselines, and Logistic Regression and Linear SVC were always among the best models. The training accuracy of the Random Forest classifier is close to perfect, and the validation accuracy rises steadily with the increase in the size of the training

ISBN: 978-81-985770-0-9

data. The small gap between the training and validation curves indicates that the classifier has good generalization capability.

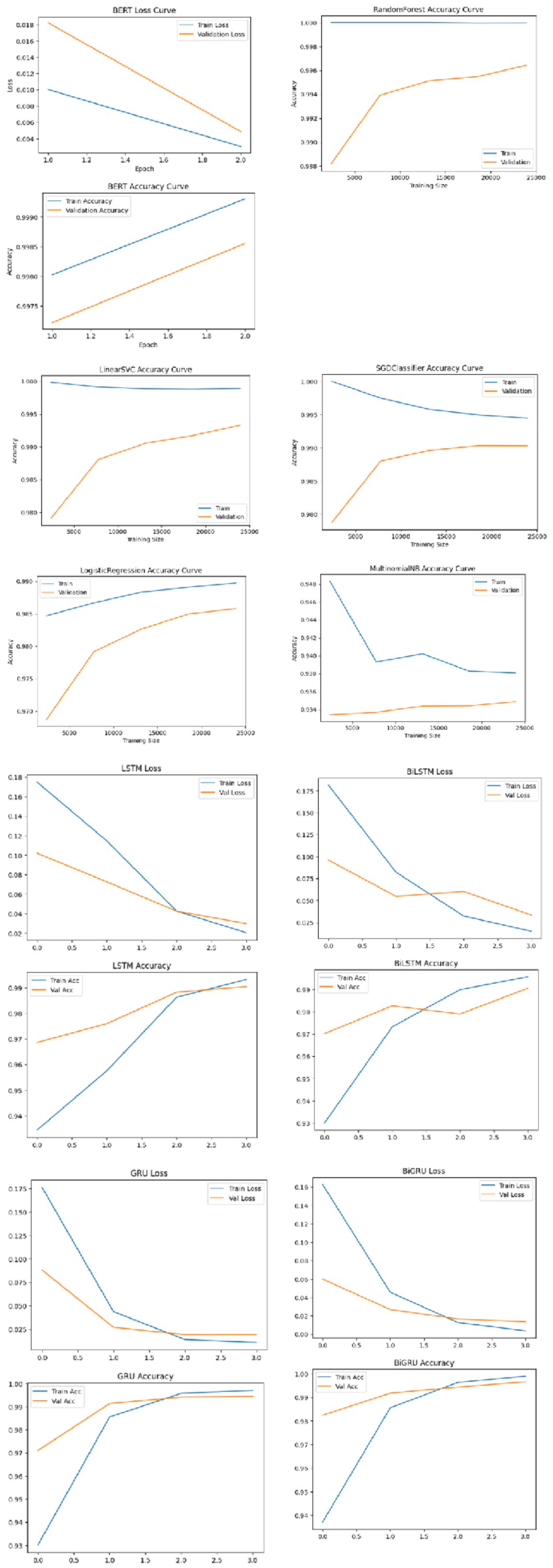

Fig. 2. Results

The Linear SVM model holds a high level of training accuracy and a gradual increase in validation accuracy with the increase in the training data. This shows that the model performs well on high-dimensional TF-IDF features and has a good generalization capability. The training accuracy of the SGD Classifier slightly decreases as the size of the data increases, but the validation accuracy increases and becomes stable. This is a sign of good regularization and the lack of overfitting when working with a lot of text data. Logistic Regression shows a steady rise in both training and validation accuracy as the size of the training data increases. The fact that the two curves are very close to each other indicates that it is an efficient algorithm for large-scale text classification problems. Multinomial Naive Bayes has relatively poor accuracy compared to other classifiers, and there is little improvement in accuracy as the training set size increases. This is because the model is based on assumptions about word frequencies.

The models of deep learning showed improved performance through the acquisition of semantic representations. The BiLSTM model outperformed CNN-based models through the inclusion of the context of news articles. The loss and accuracy of the LSTM model are decreasing and increasing steadily with each epoch. This indicates that the model has the ability to learn the dependencies in the news text. BiLSTM learns faster and has a higher validation accuracy than LSTM since BiLSTM is capable of learning context from both the forward and backward passes. The GRU model decreases its loss rate in the early stages, implying that it converges faster than the LSTM model, which has higher computational complexity.

The BERT model performed the best on all evaluation metrics. With reference to the ability of a BERT model to understand contexts and connections, it is possible to classify complicated news contents accurately, though at a computational cost. The comparative analysis shows that although transformer-based models have better accuracy, traditional models and deep learning models are still useful because of their efficiency and resource requirements.

A side-by-side examination of all the models that we tested shows that there are clear differences in the performance of the models in terms of how accurately a model can address the aspect of “fake news” in the text that is provided. Some models, such as the traditional machine learning models that make use of TF-IDF, such as the logistic regression, the SVC, and the SGD classifier, exhibited a strong performance in terms of reliability, which shows the level of efficiency that these models possess in terms of the classification of the nature of the text that is inputted, as this kind of classification is a big dimension in this case.

Both TextCNN and BiLSTM excel in the deep learning environment as they utilize the meaning and word order of the text best. In fact, BiLSTM outperforms all the

ISBN: 978-81-985770-0-9

convolution-based methods. On the other hand, BERT steadily decreases training and validation losses and hence accuracy, which means it learns very well without overfitting. This indicates a strong transformer-based approach in fake news detection. All in all, the BERT-based model is the leader in all metrics, thanks to its better contextual representations for providing more accurate classifications than other traditional and deep learning models.

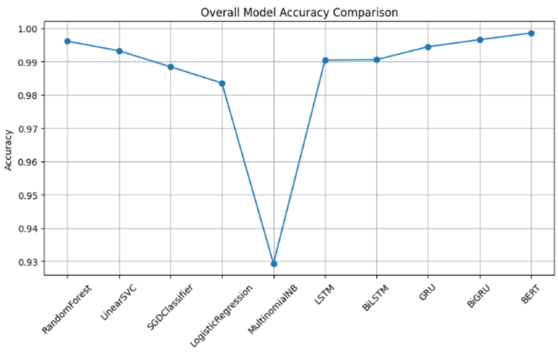

Fig. 3. Model accuracy comparison

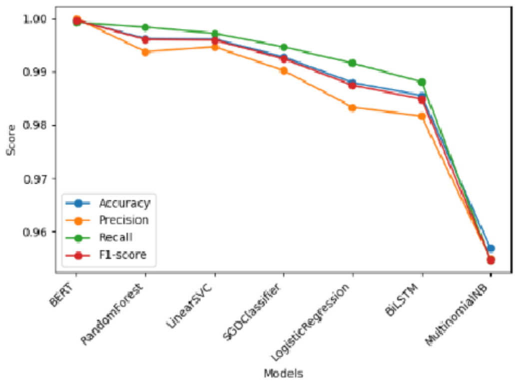

Fig. 4. Comparative Performance of ML/DL models

TABLE I:
PERFORMANCE COMPARISON OF MODELS

Model	Accuracy	Precision	Recall	F1-Score
BERT	0.998552	1.000000	0.996956	0.998475
BiGRU	0.996548	0.994863	0.997892	0.996376
RandomForest	0.996102	0.997884	0.993911	0.995893
GRU	0.994432	0.994145	0.994145	0.994145
LinearSVC	0.993207	0.992281	0.993443	0.992861
BiLSTM	0.990535	0.992006	0.988056	0.990027
LSTM	0.990423	0.983811	0.996253	0.989993
SGDClassifier	0.988419	0.985322	0.990398	0.987853
LogisticRegression	0.983519	0.979749	0.985714	0.982722
MultinomialNB	0.929287	0.926944	0.924122	0.925531

The chosen model, which was meant for the detection of fake news, was discovered to actually be a BERT-based transformer. With all the models, the former proved superior not just in terms of accuracy, but also F1-score. BERT excels in grasping the context and interpretation of text in news articles—in instances where misleading articles mostly rely on nuance and understanding of such content. While it does require a bit of computing power, its reliability cannot be overstated, and it would be the best bet for ensuring reliable detection of fake news.

VI. CONCLUSION

The paper demonstrates the effectiveness in detecting fake news using text-based machine learning and deep learning models. In this manner, the paper provides a comprehensive comparative analysis using a wide variety of models, including some cutting-edge techniques in the realm of fake news detection. From the findings, it can be stated that BERT-inspired models head the chart, all thanks to their ability to comprehend the essence of language. It is also important to note herein that traditional machine learning approaches, which make used of TF-IDF features, also yield a high level of accuracy while having much lower computational demands.

On the whole, the piece reminds us of the importance of integrating different techniques to develop fake news detection systems with access to scale and strength. The model discussed above may be incorporated with news verification platforms, social media analysis, and content management platforms. Moving forward, if researchers wish to take misinformation detection a step further, it might be worth exploring in terms of multimodal data, explorable AI techniques, and making it work in real-time.

REFERENCES

[1] A. Iqbal, K. Shahzad, S. A. Khan, and M. S. Chaudhry, "The relationship of artificial intelligence (ai) with fake news detection (fnd): a systematic literature review," Global Knowledge, Memory and Communication, vol. 74, no. 5-6, pp. 1617–1637, 10 2023. [Online]. Available: https://doi.org/10.1108/GKMC-07-2023-0264

[2] 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, mar 2021. [Online]. Available: https://doi.org/10.1088/1757-899X/1099/1/012040

[3] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake news detection on social media: A data mining perspective," arXiv preprint arXiv:1708.01967, 2017.

[4] Y. Wang, Y. Zhang, X. Li, and X. Yu, "Covid-19 fake news detection using bert-based models," arXiv preprint arXiv:2109.14816, 2021.

ISBN: 978-81-985770-0-9

[5] M. F. Mridha, A. J. Keya, M. A. Hamid, M. M. Monowar, and M. S. Rahman, “A comprehensive review on fake news detection with deep learning,” IEEE Access, vol. 9, pp. 156 151–156 170, 2021.

[6] J. Alghamdi, Y. Lin, and S. Luo, “A comparative study of machine learning and deep learning techniques for fake news detection,” Information, vol. 13, no. 4, pp. 1–18, 2022.

[7] “A benchmark study of machine learning models for online fake news detection,” Machine Learning with Applications, vol. 4, p. 100032, 2021.

[8] J. Yi, Z. Xu, T. Huang, and P. Yu, Challenges and Innovations in LLM-Powered Fake News Detection: A Synthesis of Approaches and Future Directions. New York, NY, USA: Association for Computing Machinery, 2025, p. 87–93. [Online]. Available: https://doi.org/10.1145/3728725.3728739

[9] M. M. Lawal and A. Abdulrauf, “Fake news detection using bi-lstm architecture: A deep learning approach on the isot dataset,” Journal of Computing Theories and Applications, vol. 3, no. 2, pp. 104–114, 2025.

[10] O. Prakash and R. Kumar, “Fake news detection in social networks using attention mechanism,” in Proceedings of the International Conference on Cognitive and Intelligent Computing: ICCIC 2021, Volume 2. Springer, 2023, pp. 453–462.

[11] X. Xu, P. Yu, Z. Xu, and J. Wang, “A hybrid attention framework for fake news detection with large language models,” in 2025 5th International Conference on Neural Networks, Information and Communication Engineering (NNICE), 2025, pp. 587–590.

[12] O. Prakash and R. Kumar, “A unified generalization enabled ml architecture for manipulated multi-modal social media,” Multimedia Tools and Applications, vol. 83, no. 8, pp. 22 749–22 771, 2024.

[13] M. Nasser, N. I. Arshad, A. Ali, H. Alhussian, F. Saeed, A. Da’u, and I. Nafea, “A systematic review of multimodal fake news detection on social media using deep learning models,” Results in Engineering, vol. 26, p. 104752, 2025.

[14] S. Alonso-Bartolome and I. Segura-Bedmar, “Multimodal fake news detection,” arXiv preprint arXiv:2106.05215, 2021.

[15] K. M. Fouad, S. F. Sabbeh, and W. Medhat, “Arabic fake news detection using deep learning.” Computers, Materials & Continua, vol. 71, no. 2, 2022.

Cite this article as:

Gurudev Kumar, Om Prakash and et. al. " A Comparative Study of Machine Learning, Deep Learning, and Transformer-Based Models for Fake News Detection", Proceedings of 13th international conference on Microelectronics, Circuits and Systems, Micro2026.
Displayed as online on 27th June 2026.

Link: http://actsoft.org/science/micro2026-pro/415-micro2026.pdf

ISBN: 978-81-985770-0-9