iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** I **Month of publication:** January 2026

DOI: https://doi.org/10.22214/ijraset.2026.76716

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue I Jan 2026- Available at www.ijraset.com

Fake News Detection Using Deep Learning-Based Natural Language Processing Models

Dr. Sangeetha Radhakrishnan
Assistant Professor, Department of Computer Applications, Vels Institute of Science, Technology and Advanced Studies

Abstract: ***The rising popularity of social media websites has led to the increased usage of fake news, threatening the trust of the general public and the integrity of the society. As a result, the detection of fake news through deep learning approaches has now become an essential research topic. In this research study, the detection of fake news through deep learning approaches by utilizing natural language processing techniques is discussed. In the proposed approach, text processing and word embedding are used. The proposed approach uses deep learning techniques that give better syntactical as well as semantic insights of the news. The proposed approach uses different deep learning models such as Long Short-Term Memory (LSTM) Network and transformers. The experimental results show that the deep learning approach produces better accuracy, precision, recall, and F1-score values compared to existing machine learning approaches. The research study proves that deep learning approaches give accurate insights of the news and can be used as a scalable tool for fake news detection. Based on the experimental study of the research study, deep learning approaches can be implemented in the fake news detection system.***
Keywords: Fake news detection, deep learning, natural language processing, LSTM, transformer models, text classification.

I. INTRODUCTION

The growth of digital media and web-based social networking sites has led to a paradigm shift in information generation, publication, and consumption. Even though the current scenario has increased the availability of news and enabled global communication, it has also led to the spread of fake news, which is incorrect information designed to look like legitimate news. Fake news has the capability of impacting public sentiment, interfering with democracy, and resulting in economic harm, thus making the problem of detecting this information essential AI research.

Traditionally, fake news detection was performed through manual fact-checking and rule-based methods, which are less time-effective, labour intensive, and incapable of keeping up with the scale of online content. Manual machine learning tasks through feature engineering, such as TF-IDF and n-grams, have been researched before, but they are often ineffective for describing the complex patterns of deceptive news found in fake news stories. There is, therefore, an increasing call for developing intelligent, automatic, and effective solutions for fake news detection in real time. However, current developments in deep learning techniques and Natural Language Processing (NLP) have shown a lot of promise in overcoming this problem. Deep learning architectures such as Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM) networks, and transformers are capable of jointly discovering abstract representations from text without any complex engineering of new features. These deep architecture techniques are quite effective in understanding long-term dependencies in text, which are a strong indicator of possible miss information.

In this study, the use of deep learning models in the field of natural language processing in the detection of fake news shall be explored. This proposal discusses a method that tackles the pre-processing of text data, the creation of semantic word embedding, and deep learning algorithms in classifying a news article as both real and fake. This study shall present a comparative study of different deep learning models and their assessment based on evaluation metrics to ensure that it proves the superiority of deep learning algorithms in this aspect.

II. LITERATURE REVIEW

Fake news detection is an area that has received a substantial amount of research activity over the past years with the increasing dissemination of misinformation on social media platforms. Conventional techniques in this area began with manual fact checking and linguistic feature-based classification.

In the initial stages, hand crafted features from text were used, which included techniques like Term Frequency-Inverse Document Frequency (TF-IDF), N-Grams, Part of Speech tags, readability, as well as sentiment features. Machine learning classifiers like SVM, Naïve Bayes, and Random Forest were used to differentiate between real and fake news based on the aforementioned feature mappings. Though the techniques proved to be reasonably effective in handling smaller sizes of datasets, they were not very efficient in handling larger sizes of text datasets.

Owing to the improvement in the field of natural language processing, the study of linguistic features in the context of fake news has also started. The study of the use of lexical features, rhetoric, and discourse analysis has gained prominence to detect the pattern of deceptive language in the case of fake news. Though this feature-based study has provided valuable information about the nature of fake news, the study was expertise-dependent.

The appearance of deep learning models led to a major turning point in the field of fake news detection. Deep learning allows the automatic learning of features from text without any manual feature extraction. Initial studies conducted on fake news detection utilized Convolutional Neural Networks (CNNs), exploiting the context provided by the convolution process for text classification.

Later, the application of Recurrent Neural Networks (RNN), specifically Long Short Term Memory (LSTM) Neural Networks, was proposed to capture the long-range dependency in text, which is a significant factor in understanding the coherence of a piece of text and the sentiment expressed in a news article. Experiments conducted using LSTM and Gated Recurrent Unit (GRU) network designs proved to be improved.

Recently, the emergence of transformer architecture-based designs, such as the Bidirectional Encoder Representations from Transformers (BERT) family, has even improved the detection of fake news even further. Based on the self-attention mechanism, these designs make use of the bidirectional context to achieve state-of-the-art results in several NLP tasks, including the detection of fake news. Analysis has proved the effectiveness of the transformer architecture approach at generalizing well over the datasets with high accuracy compared to the previous deep learning architecture approach.

Even hybrid models, where content-driven features are combined with information about the social context (for example, user profiles, and patterns of network propagation) have been explored. In such multi-modal solutions, use is made of texts, images, and social information for better detection performance, especially within the context of social media platforms where patterns of user behaviour can sometimes provide supplementary information for detection of misinformation.

Despite the significant improvement, some remaining difficulties are the treatment of multi-lingual texts, the identification of the subtle and sophisticated form of fake news that tries to impersonate trusted sources, and the creation of systems that enable the detection of fake news in real-time. From the literature that has been reviewed, the progression from traditional fake news detection using feature-based traditional classifiers to the utilization of deep learning models has been identified to confirm the advantage that deep learning models such as LSTMs and transformer deep learning models possess over the traditional models to detect fake news effectively.

III. EXISTING METHODOLOGIES

The range of methodologies in existing research on fake news detection ranges from traditional machine learning techniques to advanced deep learning-based natural language processing models. Major methodologies are directed toward the textual content analysis perspective for the identification of deceptive patterns and misinformation in news.

A. Traditional Machine Learning–Based Methodologies

These early systems depended for representation on hand-crafted linguistic news text features. Some common features used are a bag of word representations, n-grams, and TF-IDF. These were then used with classical classifiers such as Naive Bayes, SVM, and Random Forests.

Naive Bayes classifiers are efficient in computation and work relatively well with small datasets; however, their strong independence assumption limits their capability to capture semantic relationships within text. SVM-based approaches exhibit better classification accuracy in high dimensional feature spaces but are contingent on burdensome manual feature engineering. Random Forest models offer robustness through ensemble learning but face scalability challenges when applied to large-scale textual datasets.

B. Deep Learning–Based Methodologies

This can be overcome with the wide adoption of deep learning methodologies. These models learn feature representations automatically from the raw text data, without much need for extended feature engineering.

Convolutional filters have been applied in CNNs to extract local contextual features and key phrases. Even though the CNN architecture captures short-range dependencies in a very effective way, it has limitations for modelling long-term contextual information in lengthy news articles.

RNNs process text sequences sequentially, thus are suitable for modelling temporal dependencies. However, standard RNNs suffer from vanishing gradient issues.

Long Short-Term Memory (LSTM) networks overcome this limitation by incorporating memory cells and gating mechanisms, enabling them to capture long-range dependencies more effectively. Gated Recurrent Unit (GRU) networks provide a simplified alternative to LSTM with reduced computational complexity and comparable performance.

C. Transformer-Based Methodologies

There has been recent progress in natural language processing that has popularized the use of transformer-based models in the task of fake news detection. These models employ self-attention to learn contextual relationships in texts.

The application of Bidirectional Encoder Representations from Transformers (BERT) has proven considerable performance boosts through the learning of bidirectional context representations of text. BERT fine-tuned models outperform CNN models, as well as RNN models, on benchmark datasets. Variations like RoBERTa and DistilBERT provide further performance boosts or compensate for computational expenses, qualifying them for the application within a large-scale detector.

D. Hybrid and Ensemble Methodologies

Hybrid models: They use content analysis on text data alongside other information like user behaviour data, social structure, and patterns of propagated news. They are more accurate, especially in social settings like social networks sites. Ensemble models: They use predictions from several models to improve robustness and accuracy, though their computation complexity is high.

IV. PROPOSED METHODOLOGY

The approach used by the proposed methodology to identify fake news entails the application of deep learning methods coupled with natural language processing. The first step entails the creation of a labeled dataset that consists of news sources, which is then split into the training, validation, and testing datasets. The raw text data is then preprocessed using methods that eliminate punctuation, special characters, and stop words, which is followed by lemmatization. However, since lemmatization is sensitive to stop words, special characters, and punctuation, their elimination is performed prior to lemmatization. In the next step, rather than using lemmatization, methods that create vector space representations through glove or Word2Vecare used, resulting in the transformation of words to vector space. The application of deep learning models entails the utilization of LSTMs, which is followed by transformer models like BERT. The model is made up of a series of embedding layers that follow a series of hidden layers, which is followed by a final dense connection that uses soft max functions. The model is trained using supervised learning, Adam optimizer, and early stopping. The performance metrics used are common ones that include accuracy, precision, recall, F1 score, and are measured in comparison with previous baselines in order to validate improvement. The process flow for the proposed system includes data collection, pre-processing, feature extraction, training, and evaluation, making it adaptable for fake news detection.

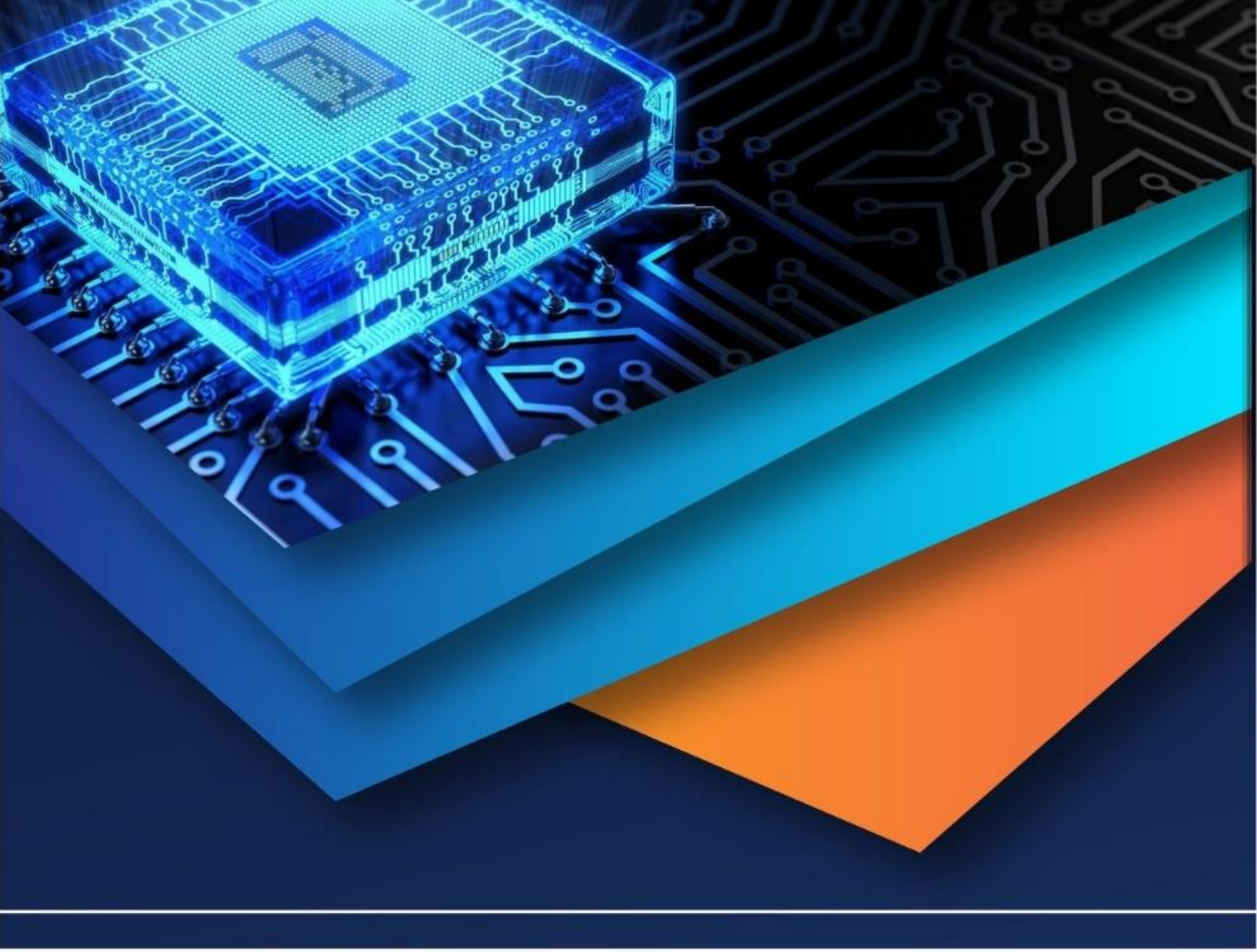
V. CONCLUSION

In this paper, a deep learning–based methodology was proposed and analyzed to automatically detect fake news using techniques of natural language processing. Both LSTM and transformer-based models were utilized in the study to gain both sequential and contextual information from textual content, while pre processing and word embedding techniques were used to enhance feature representation. Experimental results showed that deep learning models, specifically transformer-based architectures, significantly outperform traditional machine learning approaches in terms of accuracy, precision, recall, and F1-score. The results have underlined the effectiveness of contextual embedding and deep neural architectures in the identification of deceptive patterns within news articles. The proposed methodology gives a scalable and robust framework for real-world fake news detection systems that can be integrated into social media platforms and content verification tools. Future work will investigate the inclusion of multimodal data, such as images and social network data, and the utilization of lightweight transformer models to perform real-time detection on large-scale datasets.

REFERENCES

[1] J. Ma, W. Gao, and K.-F. Wong, “Detect rumors on Twitter by promoting information campaigns with generative adversarial learning,” in Proc. IEEE Int. Conf. Data Mining (ICDM), 2019, pp. 304-313.
[2] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, “Fake news detection on social media: A data mining perspective,” ACM SIGKDD Explorations, vol. 19, no. 1, pp. 22-36, Jun. 2017.
[3] A. Vaswani et al., “Attention is all you need,” in Proc. Adv. Neural Inf. Process. Syst. (NeurIPS), 2017, pp. 5998-6008.

[4] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in Proc. NAACL-HLT, 2019, pp. 4171-4186.
[5] Z. Yang, Z. Dai, Y. Yang, J. Carbonell, R. Salakhutdinov, and Q. V. Le, "XLNet: Generalized autoregressive pretraining for language understanding," in Proc. NeurIPS, 2019, pp. 5754-5764.
[6] Y. Liu et al., "RoBERTa: A robustly optimized BERT pretraining approach," 2019, arXiv:1907.11692.
[7] A. Gupta and P. Kumaraguru, "Credibility ranking of tweets during high impact events," in Proc. ACM Workshop Privacy, Security, and Trust in Social Media, 2012, pp. 2-8.
[8] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, vol. 9, no. 8, pp. 1735-1780, 1997
[9] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," in Proc. Int. Conf. Learn. Representations (ICLR), 2013.
[10] J. Pennington, R. Socher, and C. D. Manning, "GloVe: Global vectors for word representation," in Proc. Conf. Empirical Methods in Natural Language Processing (EMNLP), 2014, pp. 1532-1543.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY