International Journal of Current Trends in Engineering and Technology
www.ijctet.org, ISSN: 2395-3152
Volume: 10, Issue: 05 (SEP-OCT, 2024)

Fake News Detection Using Machine Learning: A Comparative Analysis of Algorithms

Ansh Sharma, Aakash, Anushka Singh, Adarsh Singh
SCSET, Bennett University, Greater Noida, India
e22cseu1322@bennett.edu.in, e22cseu1324@bennett.edu.in, e22cseu0808@bennett.edu.in, e22cseu1323@bennett.edu.in

Abstract-With the rapid proliferation of fake news, distinguishing between real and fabricated information has become a critical challenge in the digital age. Fake news not only manipulates public opinion but also poses a significant threat to social stability. This study explores the application of machine learning techniques for the automatic detection of fake news articles. A labelled dataset comprising both genuine and fake news is utilized, incorporating preprocessing techniques such as data cleaning, text normalization, and feature extraction. The research implements and evaluates multiple machine learning models, including Logistic Regression, Naïve Bayes, and Support Vector Machines (SVM), comparing their performance based on standard evaluation metrics such as accuracy, precision, recall, and F1-score. The experimental results demonstrate that machine learning models significantly enhance fake news detection, with DistilBERT achieving a state-of-the-art accuracy of 99.90%, outperforming traditional approaches. The study underscores the effectiveness of Natural Language Processing (NLP)-based techniques in improving classification accuracy. The findings contribute to the growing body of literature on automated misinformation detection and highlight potential avenues for further research, such as integrating metadata-based features, ensemble learning approaches, and real-time detection systems. These insights serve as a foundation for developing more robust and scalable solutions for combating misinformation in digital media.

Keywords- Fake News Detection, Machine Learning, Natural Language Processing (NLP), Transformer Models, Support Vector Machines (SVM), Deep Learning

1. Introduction

The rapid evolution of digital communication, facilitated by the internet and social media, has revolutionized the dissemination of information on a global scale. However, this transformation has also exposed a significant drawback—the widespread circulation of false information, commonly referred to as "fake news." Fake news consists of deliberately fabricated stories or misinformation designed to mislead the public by masquerading as legitimate news. Its impact has been profoundly detrimental, influencing political discourse, inciting violence, and spreading misinformation on critical issues such as public health and safety. Given the vast volume of online information, distinguishing between authentic and fabricated content has become increasingly challenging. The urgency of fake news detection has never been greater. The ability to automatically identify and filter false content is crucial for maintaining the integrity of information and preserving public trust. Effective detection mechanisms can mitigate the detrimental effects of misinformation, prevent the dissemination of deceptive content, and promote credible and verified information sources. This issue is particularly relevant in the context of media literacy, where individuals must critically assess the quality and reliability of the information they consume. This study aims to address the issue of fake news detection by leveraging machine learning techniques. Our methodology involves multiple key stages, starting with data preprocessing, which involves cleaning and transforming raw news text into a structured format suitable for analysis. We implement and compare several machine learning models, including Logistic Regression, Naïve Bayes, and Support Vector Machines (SVM), to classify news articles as either real or fake. The performance of these models is evaluated using standard metrics such as accuracy, precision, recall, and F1-score.

Additionally, we employ Natural Language Processing (NLP) techniques to extract meaningful textual features that enhance the classification process. The primary objective of this research is to develop a robust and efficient fake news detection system capable of achieving high accuracy in identifying fabricated news articles. By improving detection capabilities, this study contributes to the broader effort of combating online misinformation, reinforcing public trust in digital media, and enhancing the reliability of information ecosystems.

The remainder of this paper is structured as follows. Section 2 is a Literature Review that provides a comprehensive review of previous research on fake news detection, highlighting various machine learning and deep learning approaches. Section 3 is a Methodology that discusses the data preprocessing techniques, feature extraction methods, and machine learning models used in this study. Section 4 has Evaluation Metrics that define the performance measures, including accuracy, precision, recall, and F1-score, used to assess model effectiveness. Section 5 presents the results and analysis of the experimental results and compares the performance of different models. Finally, Section 6 is a Conclusion and Future Work that summarizes the key findings of the research and outlines potential future directions for enhancing fake news detection systems.

2. Literature Review

In recent years, fake news detection has emerged as a critical area of research across academic and technological domains. Scholars from diverse disciplines, including computer science, linguistics, and communication studies, have explored methods to detect and filter misinformation automatically. The widespread use of social media platforms has further underscored the necessity of reliable fake news detection systems, as false information can propagate rapidly and influence public perception on a massive scale.

Evolution of Fake News Detection Approaches

Traditional rule-based and statistical approaches were initially employed for fake news detection. However, these methods often fail to capture the linguistic complexity and contextual nuances of deceptive content. As a result, researchers have increasingly turned to machine learning (ML) and deep learning (DL) approaches, which have demonstrated significantly improved accuracy in detecting fabricated news.

Hybrid Models for Fake News Detection

One notable advancement in the field was proposed by Ruchansky et al. (2017), who developed a hybrid model combining content-based and network-based features. Their approach incorporated text mining techniques, such as word embeddings and sentiment analysis, alongside network-based features, including user credibility and source reliability. This multi-faceted approach significantly enhanced the accuracy of fake news detection by integrating multiple sources of evidence. Their experimental findings demonstrated that leveraging both textual and contextual signals improves model performance compared to traditional classification techniques.

Deep Learning for Contextual Understanding

Further advancements in fake news detection have been driven by deep learning techniques, which excel in capturing complex linguistic patterns. Khan et al. (2020) investigated the application of Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) for fake news classification. Their study highlighted the significance of contextual and linguistic features, demonstrating that deep learning models outperform traditional machine learning approaches, such as Logistic Regression and Naïve Bayes, in identifying deceptive content. The improved performance of deep learning models can be attributed to their ability to extract high-dimensional semantic representations from text, making them more effective in distinguishing fake news from authentic articles.

Comparative Analysis of Machine Learning Models

Several studies have compared the performance of different machine learning classifiers in fake news detection. Pugoy et al. (2019) conducted an extensive evaluation of Naïve Bayes, Random Forest (RF), and Support Vector Machines (SVM) on fake news datasets. Their findings revealed that SVM when combined with Term Frequency-Inverse Document Frequency (TF-IDF) features, achieved superior precision and recall compared to other models. Their research further emphasized the importance of feature selection and extraction techniques in enhancing classification performance.

Role of NLP in Fake News Detection

Natural Language Processing (NLP) has played a pivotal role in advancing fake news detection by improving feature extraction techniques. Zhao et al. (2018) demonstrated how NLP-based approaches, including part-of-speech tagging and dependency parsing, can uncover deep linguistic and semantic structures in news articles. Their study achieved notable improvements in feature extraction, leading to higher classification accuracy than earlier models. This highlights the growing importance of NLP techniques in tackling the complexity of human language and their effectiveness in identifying deceptive content. The evolution of fake news detection has transitioned from rule-based methods to advanced machine learning and deep learning approaches. Studies have demonstrated that hybrid models, deep learning architectures, and NLP techniques significantly enhance classification accuracy. The comparative analysis of different machine learning algorithms underscores the importance of feature engineering and contextual analysis in identifying deceptive content. Future research directions could focus on integrating metadata features, ensemble learning strategies, and real-time detection systems to improve fake news classification further.

3. Methodology

The methodology for this research comprises three major phases: data preprocessing, machine learning model training, and performance evaluation. These steps ensure that the dataset is properly processed, models are effectively trained, and their performance is rigorously assessed using standard evaluation metrics. The overall workflow of the fake news detection pipeline is illustrated in Figure 1.

3.1 Data Preprocessing

Data preprocessing is a crucial step in preparing raw text for machine learning algorithms by removing noise and enhancing feature representation. The preprocessing techniques employed in this process include data cleaning, dataset combination, tokenization, and feature extraction. Data cleaning involves removing punctuation, special characters, and stopwords such as “and,” “the,” and “is” to eliminate irrelevant information. Additionally, all text was converted to lowercase to maintain uniformity and reduce redundancy. For dataset combination, two separate datasets were merged into a single dataset, with binary labels assigned—1 for real news and 0 for fake news. Tokenization was then applied to convert text into individual tokens, enhancing feature representation and facilitating model training. Finally, feature extraction was performed using DistilBERT, a transformer-based model, to convert textual data into numerical vectors suitable for machine learning models. This deep learning-based technique captures semantic and contextual information, leading to improved classification accuracy.

3.2 Machine Learning Models Implemented

To evaluate the effectiveness of machine learning approaches in fake news detection, several models were implemented, including Logistic Regression, Naïve Bayes, Support Vector Machines (SVM), and DistilBERT. Logistic Regression served as a baseline classification model, applying a linear decision boundary for binary classification. It is computationally efficient and interpretable, making it a suitable choice for text classification tasks.

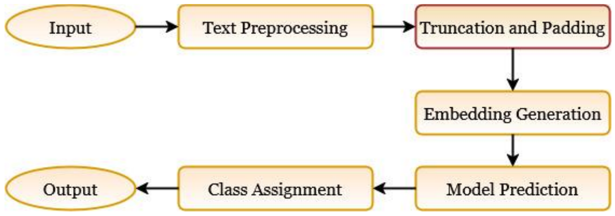

Figure 1 illustrates the workflow of the fake news detection pipeline.

Naïve Bayes, a probabilistic classifier, was employed due to its assumption of word independence, allowing it to handle high-dimensional sparse data effectively. Support Vector Machines (SVM) were utilized as a margin-based classifier that identifies the optimal decision boundary between real and fake news, demonstrating strong performance in high-dimensional feature spaces, particularly for text-based data representations. Additionally, DistilBERT, a lightweight

International Journal of Current Trends in Engineering and Technology
www.ijctet.org, ISSN: 2395-3152
Volume: 10, Issue: 05 (SEP-OCT, 2024)

transformer-based model, was implemented to optimize text classification tasks. This deep learning model captures complex contextual dependencies in textual data, offering superior performance compared to traditional machine learning models in detecting deceptive content.

3.3 Model Training and Optimization

Model training was conducted using the AdamW optimizer, which was chosen for its efficiency in handling large-scale text data. To enhance the training process, learning rate scheduling was applied to adjust the convergence speed, ensuring optimal dynamic performance. Additionally, gradient clipping was implemented to prevent exploding gradients, thereby maintaining model stability throughout training. This comprehensive approach not only facilitates effective feature extraction but also enhances model robustness and overall performance in detecting fake news articles.

4. Evaluation Metrics

To assess the effectiveness of the machine learning models in fake news detection, we utilized standard evaluation metrics. These metrics provide quantitative measures of model performance and ensure reliable comparisons across different classification algorithms.

Accuracy: Accuracy measures the proportion of correctly classified articles among the total number of samples. It is defined as:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

In the context of fake news detection, True Positives (TP) refers to fake news articles that are correctly classified as fake. At the same time, True Negatives (TN) represent real news articles that are accurately identified as real. Conversely, False Positives (FP) occur when real news articles are incorrectly classified as fake, leading to misclassification. Similarly, False Negatives (FN) arise when fake news articles are mistakenly identified as real, which can significantly impact the model's reliability in detecting deceptive content.

Precision: Precision quantifies the proportion of correctly classified fake news articles among all articles labelled as fake. It is defined as:

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

A high precision value indicates that the model minimizes false positives, ensuring that articles classified as fake are indeed fake.

Recall: Recall, also known as sensitivity, measures the ability of the model to identify actual fake news articles correctly. It is given by:

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

A high recall score indicates that the model effectively identifies most of the fake news instances, reducing the number of false negatives.

F1-Score: The F1-score provides a harmonic mean between precision and recall, ensuring a balanced evaluation of the model's performance:

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

The F1 score is particularly useful when the dataset is imbalanced, as it accounts for both false positives and false negatives in classification.

Confusion Matrix: The confusion matrix, as shown in Table 1, provides a comprehensive overview of the model's classification performance by analyzing the distribution of true positives, true negatives, false positives, and false negatives.

Table 1 Confusion Matrix for Fake and Real News Classification

Actual / Predicted	Fake News (Predicted)	Real News (Predicted)
Fake News (Actual)	TP	FN
Real News (Actual)	FP	TN

The confusion matrix enables a detailed performance evaluation, highlighting misclassifications and helping to refine the model for improved accuracy.

5. Dataset Details

For this study, we utilized a publicly available Fake and Real News Dataset from Kaggle, consisting of 21,417 real news articles sourced from reputable media organizations and 23,502 fake news articles deliberately crafted to mislead readers. The dataset includes three key features: Title, which represents the headline of the article; Text, which contains the full content; and Subject, which categorizes the article into domains such as politics, world news, and technology. To ensure a balanced training process and mitigate potential biases, dataset proportions were adjusted, and stratified sampling was applied. This preprocessing step guarantees that both real and fake news articles are adequately represented during model training, thereby improving the robustness and generalizability of the classification model.

6. Results and Analysis

The machine learning models were evaluated using accuracy, precision, recall, and F1-score to assess their effectiveness in detecting fake news. The results highlight the performance differences between traditional machine-learning models and transformer-based approaches. The key findings of this study emphasize the superior performance of transformer-based models in fake news detection. DistilBERT achieved the highest accuracy of 99.90%, significantly outperforming traditional machine learning models. This result underscores the effectiveness of transformer-based architectures in capturing contextual dependencies within textual data. Among traditional models, the Support Vector Machine (SVM) with TF-IDF features demonstrated strong performance, achieving high precision (98.65%) and recall (98.92%), making it a reliable choice for fake news classification. Logistic Regression, while competitive with an accuracy of 94.50%, was less effective than SVM and DistilBERT due to its linear decision boundary, which limits its ability to model complex textual patterns. Meanwhile, Naïve Bayes (NB), known for its computational efficiency, exhibited a slightly lower accuracy of 92.30%, primarily due to its assumption of word independence, which reduces its effectiveness in capturing nuanced language structures. Furthermore, the confusion matrix analysis (Table 2) revealed minimal false positives and false negatives, affirming the robustness and reliability of the implemented models in accurately distinguishing between real and fake news articles. The comparative evaluation of different models is summarized in Table 2, which presents their respective accuracy, precision, recall, and F1 scores.

Table 2 Model Performance Metrics

Model	Acc. (%)	Prec (%)	Rec (%)	F1 (%)
DistilBERT	99.90	99.86	99.95	99.91
SVM	98.80	98.65	98.92	98.78
LR	94.50	94.20	94.75	94.47
NB	92.30	91.80	92.50	92.14

The analysis of model performance reveals that transformer-based models, particularly DistilBERT, significantly enhance fake news detection by effectively capturing contextual relationships and semantic meaning within news articles. This deep learning approach outperforms traditional machine learning models by leveraging advanced language representations. Among conventional methods, the Support Vector Machine (SVM) with TF-IDF features emerges as a strong alternative, achieving high precision and recall, which underscores its effectiveness in text classification tasks. However, Logistic Regression and Naïve Bayes, while computationally efficient, exhibit inferior performance due to their inherent limitations in modelling complex textual structures.

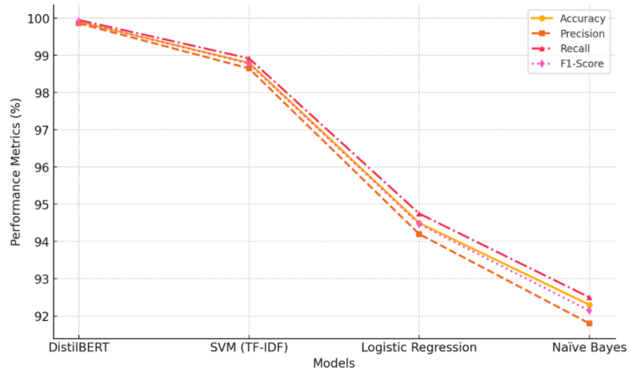

Figure 2. Performance Comparison of Fake News Detection Models Based on Accuracy, Precision, Recall, and F1-Score.

Overall, the findings indicate that deep learning-based models, particularly transformer architectures, offer superior accuracy and reliability for fake news detection. These results suggest that transformers are a promising direction for future research and real-world applications, highlighting their potential to combat misinformation effectively. The comparative performance of different models is visually represented in Figure 2, which illustrates the variations in accuracy, precision, recall, and F1-score among the evaluated approaches.

7. Conclusion and Future Work

This study demonstrated that transformer-based models, particularly DistilBERT, are highly effective in fake news detection, achieving an accuracy of 99.90%. The findings emphasize the superiority of deep learning architectures over traditional machine learning models, as they excel in capturing contextual and semantic information within textual data. Additionally, NLP-based feature extraction techniques play a crucial role in enhancing classification performance, making them essential for improving fake news detection accuracy. To further enhance the effectiveness and applicability of fake news detection systems, future research should explore several key directions. Dataset expansion is necessary to include multilingual fake news, enabling models to generalize across different languages and cultural contexts. Metadata integration, incorporating features such as author credibility and source reliability, can further strengthen classification accuracy.

Additionally, ensemble learning approaches should be investigated to combine transformer-based models with traditional machine learning techniques, enhancing robustness and interpretability. Finally, efforts should be directed toward developing a real-time fake news detection system for online platforms, ensuring timely identification and mitigation of misinformation. These advancements will contribute to more reliable and scalable solutions for combating fake news in diverse digital environments.

References

[1]. Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake News Detection on Social Media: A Data Mining Perspective. ACM SIGKDD Explorations Newsletter, 19(1), 22-36.
[2]. Ruchansky, N., Seo, S., & Liu, Y. (2017). CSI: A Hybrid Deep Model for Fake News Detection. Proceedings of the 2017 ACM on Conference on Information and Knowledge Management (CIKM).
[3]. Hanshal, O. A., Ucan, O. N., & Sanjalawe, Y. K. (2023). Hybrid Deep Learning Model for Automatic Fake News Detection. Applied Nanoscience, 13(4), 2957–2967.
[4]. Jing, J., et al. (2023). Multimodal Fake News Detection via Progressive Fusion Networks. Information Processing & Management, 60(1), 103120.
[5]. Rawat, G., et al. (2023). Fake News Detection Using Machine Learning. In Proceedings of the 2023 International Conference on Artificial Intelligence and Smart Communication (AISC). IEEE.
[6]. Aslam, N., et al. (2021). Fake Detect: A Deep Learning Ensemble Model for Fake News Detection. Complexity, 2021, 1–10.
[7]. Ashik, S. S., et al. (2021). M82B at CheckThat! 2021: Multiclass Fake News Detection Using BiLSTM. In Proceedings of the 2021 CLEF Conference.
[8]. Chauhan, T., & Palivela, H. (2021). Optimization and Improvement of Fake News Detection Using Deep Learning Approaches for Societal Benefit. International Journal of Information Management Data Insights, 1(2), 100051.
[9]. Mewada, Arvind, and Rupesh Kumar Dewang. “SUH-AIFRD: A self-training-based hybrid approach for individual fake reviewer detection.” Multimedia Tools and Applications 83.26 (2024): 67643-67671.
[10]. Choudhary, M., et al. (2021). BerConvoNet: A Deep Learning Framework for Fake News Classification. Applied Soft Computing, 110, 107614.
[11]. Sadr, M. M. (2021). The Use of LSTM Neural Network to Detect Fake News on Persian Twitter. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12(11), 6658–6668.
[12]. Seyam, A., et al. (2021). Deep Learning Models to Detect Online False Information: A Systematic Literature Review. In Proceedings of the 7th Annual

International Conference on Arab Women in Computing.

[13]. Umer, M., et al. (2020). Fake News Stance Detection Using Deep Learning Architecture (CNN-LSTM). IEEE Access, 8, 156695–156706.

[14]. Kaliyar, R. K., et al. (2020). FNDNet–A Deep Convolutional Neural Network for Fake News Detection. Cognitive Systems Research, 61, 32–44.

[15]. Zhang, J., et al. (2020). FakeDetector: Effective Fake News Detection with Deep Diffusive Neural Network. In Proceedings of the 2020 IEEE 36th International Conference on Data Engineering (ICDE), 1826–1829. IEEE.

[16]. Lee, D. H., et al. (2019). Fake News Detection Using Deep Learning. Journal of Information Processing Systems, 15(5), 1119–1130.

[17]. Drif, A., et al. (2019). Fake News Detection Method Based on Text-Features. In Proceedings of the Ninth International Conference on Advances in Information Mining and Management (IMMM).

[18]. Ramya, S. P., & Eswari, R. (2021). Attention-Based Deep Learning Models for Detection of Fake News in Social Networks. International Journal of Cognitive Informatics and Natural Intelligence (IJCINI), 15(4), 1–25.

[19]. Singh, Satya Prakash, Mohd Aquib Ansari, and Lalit Kumar. (2023). Analysis of Website in Web Data Mining Using Web Log Expert Tool. 2023 IEEE 12th International Conference on Communication Systems and Network Technologies (CSNT), IEEE.

[20]. Rai, Sandeep, et al. (2024). Optimizing Wireless Sensor Networks: Dynamic Load Balancing Strategy with AOMDV Routing. 2024 3rd International Conference on Power Electronics and IoT Applications in Renewable Energy and its Control (PARC), IEEE.

[21]. Mewada, Arvind, et al. “SentiBERT: A novel approach for fake review detection incorporating sentiment features with contextual features.” Proceedings of the 2023 Fifteenth International Conference on Contemporary Computing. 2023.

[22]. Ahmad, Shahnawaz, et al. (2024). Smart Market Insights: A Deep Learning Approach to Trend Analysis. 2024 13th International Conference on System Modeling & Advancement in Research Trends (SMART). IEEE.