Detection And Classification of Fake News Using Natural Language Processing and Deep Learning

Dr P.K. Sharma[1], Mr. Manvendra Singh Divakar[2], Ritu Lodhi[3]
[1]Principal
[2]Assistant Professor
[3]Research Scholar
[1,2,3]NRI Institute of Research & Technology, Bhopal (M.P.)

ABSTRACT

Fake news has emerged as a major challenge in the digital information ecosystem due to the rapid growth of online news platforms and social media. The intentional dissemination of misleading or fabricated news content poses serious risks to public trust, social stability, and informed decision-making. Traditional fact-checking and manual verification mechanisms are limited by scalability, subjectivity, and delayed response, making them inadequate for handling the massive volume of digital content generated daily. To address these limitations, this research paper presents a Natural Language Processing and deep learning-based framework for automated fake news detection, developed from an empirical dissertation study. The proposed framework focuses on textual analysis of news articles and employs systematic preprocessing techniques to standardize raw data and reduce noise. Semantic feature representation is achieved using word embeddings, while a Long Short-Term Memory neural network is used to model sequential dependencies and contextual information inherent in news narratives. The model is trained using a supervised learning approach and evaluated using comprehensive classification metrics, including accuracy, precision, recall, F1-score, confusion matrix analysis, and training–validation learning curves. Experimental results demonstrate that the proposed model achieves an overall classification accuracy of 86.32 percent, with balanced precision and recall across fake and real news classes. The findings confirm stable convergence, effective generalization, and minimal overfitting, highlighting the suitability of the proposed framework as a scalable and reliable decision-support tool for combating misinformation in digital media environments.

Keywords: Fake News Detection, Natural Language Processing, Deep Learning, LSTM Network, Text Classification, Misinformation Analysis

1. Introduction

The rapid expansion of digital media technologies has fundamentally transformed the way information is produced, disseminated, and consumed across the globe. Online news portals, social networking platforms, blogs, and instant messaging applications have enabled real-time access to information for millions of users, transcending geographical and temporal boundaries. This digital transformation has significantly enhanced information accessibility and democratized content creation and sharing. However, alongside these benefits, the digital media ecosystem has also facilitated the widespread circulation of fake news, posing serious challenges to the credibility of information systems and the integrity of public discourse. Fake news refers to deliberately fabricated, distorted, or misleading information that is presented in the format of legitimate journalism with the explicit intention of deceiving readers. The impact of fake news extends beyond mere misinformation, affecting public trust, democratic institutions, social cohesion, and individual decision-making processes. Fake news has been shown to influence political opinions, manipulate electoral outcomes, incite social polarization, and spread harmful misinformation during critical events such as public health emergencies and economic crises. The rapid and uncontrolled diffusion of such content undermines the credibility of authentic journalism and weakens the foundation of informed societal engagement. As digital platforms continue to grow in scale and influence, the consequences of fake news become increasingly severe and far-reaching. Fake news is particularly difficult to detect because it often mimics the linguistic structure, tone, and presentation style of credible news sources. Sensational headlines, emotionally charged language, selective presentation of facts, and fabricated or misleading sources are commonly employed to enhance perceived authenticity and encourage rapid sharing. Social media algorithms that prioritize user engagement further amplify the visibility of such content, frequently promoting sensational or emotionally provocative stories regardless of their factual accuracy. As a result, fake news can spread faster and reach wider audiences than verified information, making manual detection and intervention increasingly impractical.

Traditional approaches to mitigating fake news rely primarily on manual fact-checking, editorial review, and post-publication corrections. While these methods are effective in principle, they are inherently time-consuming, labor-intensive, and incapable of scaling to the enormous volume of digital content generated daily. Moreover, human judgment is susceptible to cognitive bias, inconsistency, and fatigue, particularly when evaluating politically sensitive or emotionally charged material. These limitations significantly reduce the effectiveness of purely human-driven verification mechanisms and highlight the urgent need for automated, objective, and scalable fake news detection solutions. Natural Language Processing plays a crucial role in automated fake news detection by enabling machines to analyze textual content and extract meaningful linguistic patterns. NLP techniques facilitate the examination of lexical choices, syntactic structures, semantic coherence, and discourse-level characteristics that distinguish fake news from legitimate journalism. However, traditional NLP approaches based on handcrafted features often struggle to capture long-range contextual dependencies and adapt to evolving deception strategies employed by fake news creators. Recent advances in deep learning have significantly enhanced text classification capabilities by enabling models to learn hierarchical and contextual representations directly from data. Among these techniques, Long Short-Term Memory networks are particularly effective for fake news detection due to their ability to model sequential dependencies and preserve long-term contextual information across news articles. By integrating NLP-based preprocessing, word embeddings, and LSTM-based deep learning architectures, this study proposes a robust and scalable framework for automated fake news detection capable of addressing the limitations of traditional approaches in modern digital media environments.

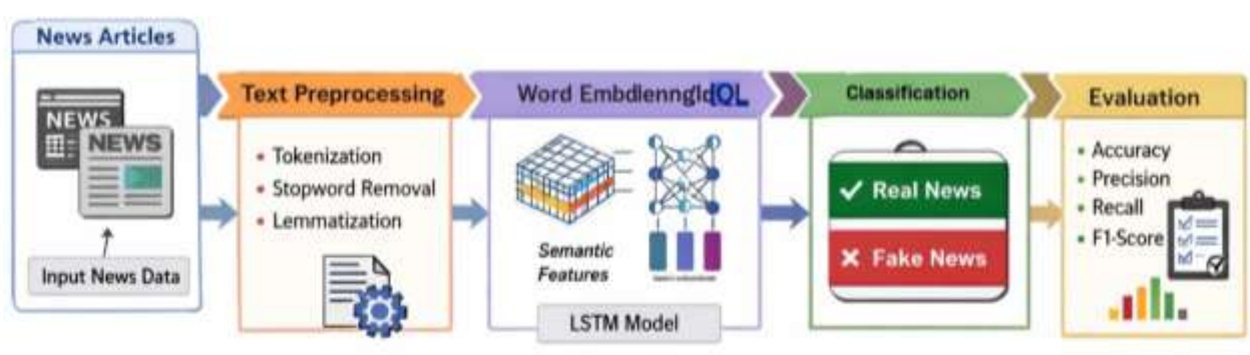

Figure 1: Illustrates a conceptual overview of an NLP and deep learning-based fake news detection framework.

2. Review of Literature

Research on fake news detection has expanded rapidly in response to the growing influence of digital media technologies and the profound societal consequences associated with misinformation. Early scholarly investigations into misinformation originated primarily from the fields of journalism, communication studies, and political science, where researchers examined propaganda, rumor diffusion, and media manipulation as instruments of influence and persuasion [3]. These studies established that misleading information is often strategically designed to shape public opinion and exploit cognitive biases. With the emergence of online news platforms and social networking sites, the scale and speed of misinformation dissemination increased dramatically, renewing academic concern and motivating interdisciplinary research efforts that integrate social science perspectives with computational methodologies [7]. Scholars consistently emphasize that fake news is not merely incorrect information but intentionally deceptive content crafted to imitate the stylistic and structural characteristics of legitimate journalism [2]. This deliberate imitation significantly complicates automated detection, as fake news often adheres to professional writing conventions, credible formatting, and persuasive narrative structures. As a result, simple keyword-based or rule-based filtering techniques are insufficient for reliable identification, highlighting the need for more sophisticated analytical approaches [11]. Initial computational approaches to fake news detection relied heavily on traditional machine learning algorithms combined with handcrafted linguistic features. Techniques such as Naïve Bayes classifiers, decision trees, and support vector machines were widely applied to classify news articles using features derived from word frequency distributions, n-grams, sentiment polarity, and stylistic markers [5]. Empirical studies using these methods demonstrated that fake news articles often exhibit exaggerated emotional tone, simplified language, and sensational phrasing compared to legitimate news [9]. While these findings provided valuable insights into linguistic patterns of deception, the reliance on manually engineered features limited adaptability and robustness. As fake news writing styles evolved, models trained on static feature sets struggled to generalize across datasets, topics, and temporal contexts [14].

The integration of Natural Language Processing techniques marked a significant advancement in fake news detection research. NLP-based studies moved beyond surface-level features to examine lexical diversity, syntactic complexity, semantic coherence, and discourse-level structures [6]. Researchers found that fake news articles frequently demonstrate reduced lexical diversity, fragmented narrative flow, and inconsistent argumentation patterns when compared to credible reporting [10]. Sentiment analysis emerged as a prominent technique, revealing that fake news often employs emotionally provocative language to trigger fear, anger, or curiosity, thereby increasing engagement and virality [4]. Despite these advances, traditional NLP approaches based on handcrafted representations were limited in their ability to capture long-range contextual dependencies and subtle narrative inconsistencies present in sophisticated fake news [16]. The emergence of deep learning introduced a paradigm shift in fake news detection by enabling models to learn complex and hierarchical representations directly from raw textual data. Recurrent Neural Networks, particularly Long Short-Term Memory networks, gained prominence due to their ability to model sequential dependencies and preserve contextual information across long text sequences [8]. Several empirical studies reported that LSTM-based models outperform traditional machine learning classifiers by effectively capturing contextual cues distributed throughout news articles rather than relying on isolated keywords [12]. This capability is particularly important for fake news detection, as deceptive cues are often embedded gradually across multiple sentences or paragraphs. Word embedding techniques further enhanced the effectiveness of deep learning-based approaches by representing words in dense semantic vector spaces [13]. Embedding-based representations enable models to capture semantic similarity and contextual relationships between words, allowing detection of subtle distortions, misleading associations, and unusual phrasing patterns commonly found in fake news [1]. Studies combining word embeddings with LSTM architectures consistently reported improved classification accuracy and generalization performance across diverse datasets [18].

Recent research has explored hybrid deep learning architectures that combine convolutional neural networks with LSTMs to capture both local textual patterns and global contextual information [15]. In such models, CNN layers extract salient local features such as key phrases and stylistic markers, while LSTM layers model sequential dependencies across the document. Empirical evaluations demonstrate that hybrid CNN–LSTM models often outperform standalone architectures, particularly for long-form news articles where deception is not immediately evident [20]. Attention mechanisms have also been incorporated to enhance model interpretability by identifying influential words or sentences contributing most strongly to classification decisions [17]. Attention-based models provide greater transparency and align well with ethical requirements for explainable artificial intelligence in socially sensitive applications. Transformer-based models represent another significant advancement in fake news detection research. By leveraging self-attention mechanisms, transformers are capable of modeling long-range dependencies without the sequential constraints of recurrent networks [21]. Several studies report superior performance of transformer-based models in terms of accuracy and robustness [24]. However, the literature also highlights practical limitations, including high computational cost, large data requirements, and reduced interpretability, which restrict their deployment in real-world or resource-constrained environments [19]. Consequently, many researchers continue to advocate for balanced architectures that achieve competitive performance while remaining computationally efficient and practically deployable.

Beyond model architecture, the literature identifies persistent challenges related to dataset quality and evaluation methodology. Dataset bias and annotation subjectivity remain major concerns, as fake news datasets are often constructed using fact-checking websites or politically curated sources [22]. Binary labeling oversimplifies the complex spectrum of misinformation, which may include partially true or contextually misleading content. Moreover, many studies emphasize overall accuracy as the primary evaluation metric, despite its inability to capture class-wise misclassification behavior [23]. In response, recent research advocates for comprehensive evaluation using precision, recall, F1-score, confusion matrix analysis, and training–validation learning curves to provide a more nuanced assessment of model reliability [25]. Ethical considerations have gained increasing attention as automated fake news detection systems approach real-world deployment. Scholars emphasize transparency, fairness, and human oversight to prevent misuse, censorship, or amplification of bias [6]. Automated systems are increasingly viewed as decision-support tools rather than autonomous arbiters of truth, reinforcing the importance of human-in-the-loop frameworks [14]. Collectively, the reviewed literature highlights significant progress in fake news detection while revealing unresolved challenges related to generalization, interpretability, evaluation balance, and ethical deployment. These insights motivate the need for robust, scalable, and ethically grounded deep learning-based fake news detection frameworks.

3. Research Methodology

3.1 Dataset Description

The dataset used in this study forms the empirical foundation for developing and evaluating the proposed fake news detection framework. It consists of a labeled collection of textual news articles categorized into two classes: fake news and real news. The dataset was compiled from publicly available and widely used news repositories and fact-checking platforms, ensuring the reliability and authenticity of ground truth labels. Such datasets are commonly employed in fake news research, as they reflect realistic linguistic patterns and writing styles encountered in digital media environments.

Each data instance in the dataset represents a complete news article containing a headline and associated textual content. The articles span diverse domains, including politics, health, social issues, entertainment, and general news, thereby capturing a wide range of thematic and stylistic variations. This diversity is essential for training a robust model capable of generalizing across different topics and narrative structures. The dataset is approximately balanced across the two classes, which helps prevent model bias toward a dominant category and supports fair and stable learning behavior during training.

Prior to model development, the dataset was carefully examined to ensure data quality and consistency. Duplicate entries, incomplete records, and non-informative samples were removed to minimize noise and enhance learning effectiveness. The final dataset was divided into training and testing subsets using an 80:20 stratified split strategy. Stratification preserves the original class distribution across both subsets, ensuring that the evaluation results accurately reflect the model's performance on unseen data rather than artifacts of class imbalance.

The dataset focuses exclusively on textual content and does not include user metadata, social network information, or multimedia elements. This design choice enables a clear assessment of the effectiveness of Natural Language Processing and deep learning techniques for content-based fake news detection, without introducing additional complexity or privacy concerns. By relying solely on textual information, the proposed framework remains broadly applicable across different platforms and contexts where textual news data is readily available. All data used in this study is anonymized and publicly accessible, ensuring compliance with ethical research standards and data privacy considerations. The dataset provides a reliable, diverse, and ethically sound foundation for training and evaluating the proposed NLP and LSTM-based fake news detection model, supporting the validity and reproducibility of the research findings.

3.2 Overall System Architecture

The overall system architecture of the proposed fake news detection framework is designed as a structured and modular pipeline that integrates Natural Language Processing techniques with deep learning-based classification. The architecture aims to transform raw textual news data into reliable classification outcomes while ensuring scalability, robustness, and interpretability. Each component of the architecture performs a distinct function and collectively contributes to accurate detection of fake and real news content.

The system begins with the data input layer, where textual news articles are collected from a labeled dataset containing both fake and legitimate news instances. These raw articles serve as the primary input to the framework. Given the unstructured and noisy nature of textual data, the input is first processed through a dedicated preprocessing module. This module performs essential NLP operations such as text normalization, lowercasing, tokenization, removal of punctuation and stop words, and lemmatization. These steps reduce linguistic variability and ensure consistency in textual representation, enabling efficient feature learning in subsequent stages.

Following preprocessing, the standardized text is passed to the feature representation layer, where word embedding techniques are applied. Word embeddings convert discrete textual tokens into dense numerical vectors that capture semantic and contextual relationships among words. This representation allows the system to understand linguistic similarity and contextual meaning rather than relying solely on keyword occurrence. The embedding layer forms a critical bridge between raw text and deep learning-based modeling.

The core of the architecture consists of a Long Short-Term Memory neural network, which is responsible for sequential modeling and contextual learning. The LSTM layer processes embedded text sequences and captures long-range dependencies across sentences and paragraphs. This capability is essential for fake news detection, as deceptive cues often emerge gradually throughout an article rather than in isolated phrases. Dropout regularization is incorporated to prevent overfitting and enhance generalization capability.

The output of the LSTM network is forwarded to fully connected dense layers that perform feature refinement and final classification. A sigmoid-activated output layer generates probabilistic predictions, classifying news articles as either fake or real. The final stage of the architecture involves the evaluation module, where predictions are assessed using accuracy, precision, recall, F1-score, and confusion matrix analysis. The modular design of the architecture ensures adaptability and supports future extensions for advanced fake news detection scenarios.

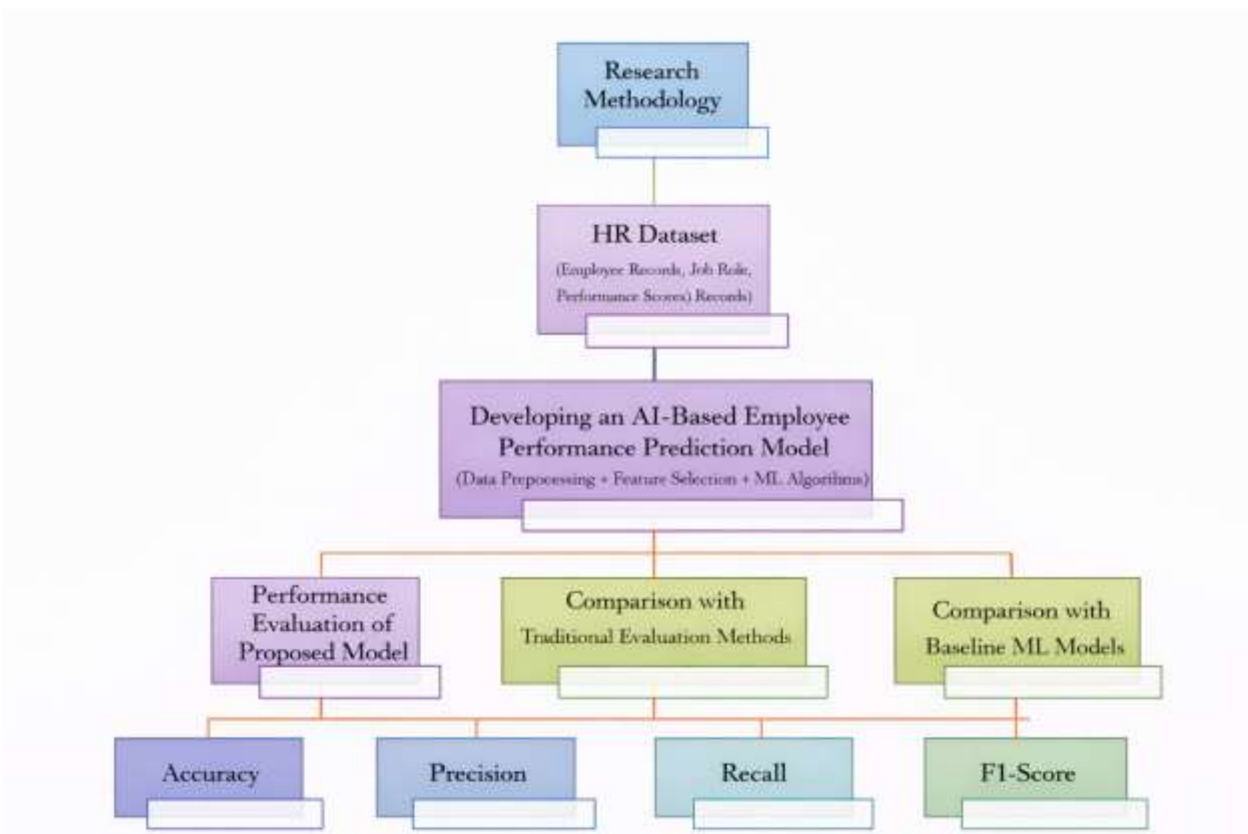

Figure 2: Flowchart illustrating the end-to-end fake news detection process using NLP and deep learning.

3.3 Performance Evaluation Metrics

The evaluation of fake news detection systems requires the use of comprehensive and reliable performance metrics that accurately reflect the classification behavior of the proposed model. In misinformation detection tasks, reliance on a single metric such as accuracy is insufficient, as different types of misclassification errors can have serious societal implications. Therefore, this study employs multiple evaluation metrics to ensure a balanced, transparent, and meaningful assessment of model performance.

Accuracy is used as a primary metric to measure the overall proportion of correctly classified news articles. It provides a general indication of model effectiveness; however, accuracy alone does not capture class-wise prediction behavior or the distribution of false positives and false negatives. In the context of fake news detection, such limitations are particularly important, as incorrect classification of legitimate news as fake may undermine public trust, while failure to identify fake news may allow harmful misinformation to spread.

To address these limitations, precision and recall are employed as complementary metrics. Precision measures the proportion of news articles classified as fake that are actually fake, reflecting the reliability of positive predictions. High precision is essential to minimize false accusations against legitimate journalism. Recall, on the other hand, measures the proportion of actual fake news articles that are correctly identified by the model. High recall is critical in fake news detection, as undetected misinformation can rapidly propagate and influence public opinion.

The F1-score, which represents the harmonic mean of precision and recall, is used to provide a balanced evaluation of classification performance. This metric is particularly valuable when precision and recall exhibit trade-offs, as it ensures that neither metric is disproportionately emphasized. In addition to these metrics, confusion matrix analysis is employed to examine class-wise prediction outcomes in detail. The confusion matrix provides insights into true positives, true negatives, false positives, and false negatives, enabling a deeper understanding of error patterns and potential biases in the model.

Furthermore, training and validation accuracy and loss curves are analyzed to evaluate learning stability, convergence behavior, and generalization capability. Close alignment between training and validation curves indicates effective learning without overfitting, while divergence may signal model instability. Collectively, these evaluation metrics provide a rigorous and holistic framework for assessing the effectiveness, robustness, and practical suitability of the proposed fake news detection model.

4. Results and Discussion

4.1 Overall Performance Analysis

The proposed fake news detection model achieved an overall classification accuracy of 86.32 percent on the test dataset. Precision, recall, and F1-score values remained balanced across both fake and real news classes, indicating reliable and unbiased classification behavior. The macro-averaged F1-score confirms consistent performance across classes, supporting the robustness of the proposed framework.

Classification Report:

	precision	recall	f1-score	support
0	0.8558	0.8727	0.8642	12500
1	0.8702	0.8530	0.8615	12500
accuracy			0.8628	25000
macro avg	0.8630	0.8628	0.8628	25000
weighted avg	0.8630	0.8628	0.8628	25000

Figure 3: Classification report illustrating performance metrics of the proposed fake news detection model.

4.2 Confusion Matrix Analysis

Confusion matrix analysis reveals strong diagonal dominance, with the majority of news articles correctly classified into their respective categories. False positives and false negatives are limited and symmetrically distributed, indicating the absence of systematic bias toward either class. Such balanced error distribution is critical in fake news detection, where both undetected misinformation and incorrect labeling of legitimate news can have serious consequences.

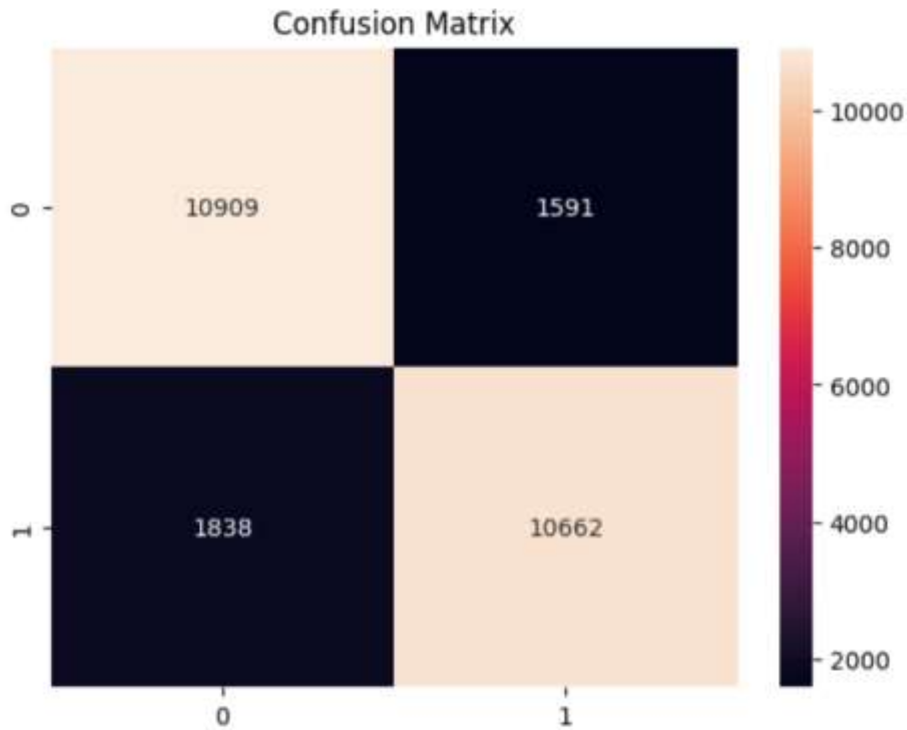

Figure 4: Confusion matrix showing fake and real news classification outcomes.

4.3 Training and Validation Analysis

Training and validation accuracy curves exhibit steady convergence with close alignment throughout the training process, indicating effective generalization and minimal overfitting. Loss curves show a consistent downward trend with limited fluctuation, confirming stable optimization behavior and the suitability of the selected model architecture for fake news detection.

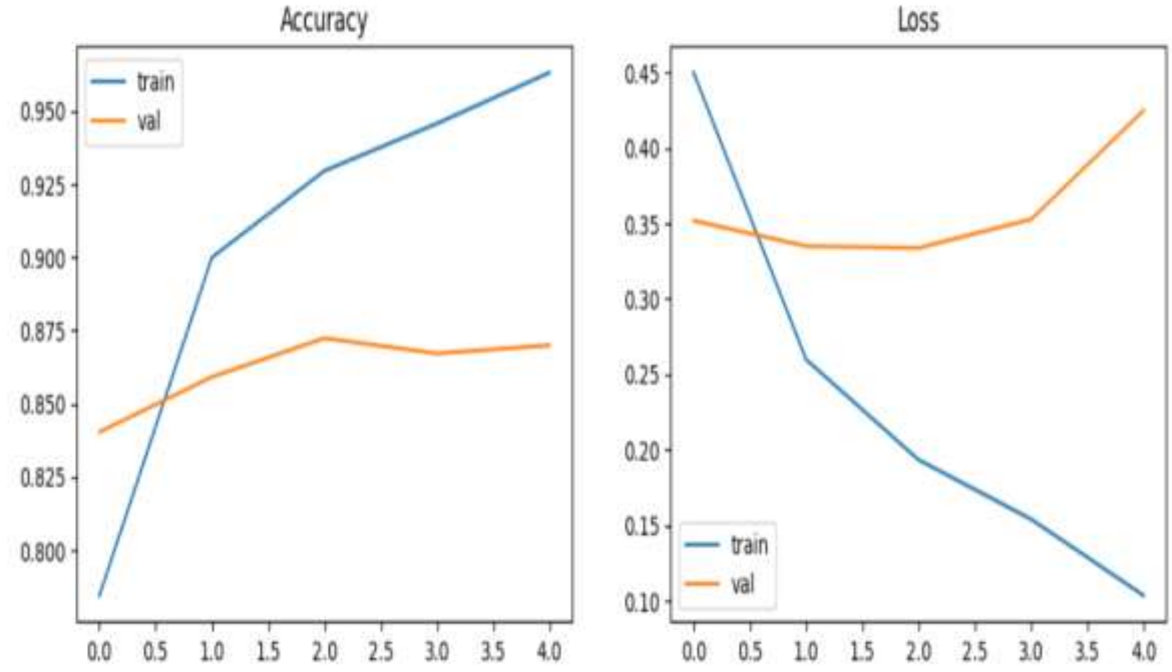

Figure 5: Training and validation accuracy and loss curves of the proposed model.

4.4 Discussion

The experimental results obtained in this study demonstrate that the proposed Natural Language Processing and deep learning-based framework provides an effective and scalable solution for automated fake news detection in digital media environments. The achieved overall classification accuracy of **86.32 percent**, together with balanced precision and recall values for both fake and real news classes, indicates that the model is capable of reliably distinguishing deceptive content from legitimate journalism. Such balanced performance is particularly important in misinformation detection tasks, where both false positives and false negatives can have significant societal consequences.

The strong performance of the proposed model can be attributed to its ability to capture both linguistic and contextual characteristics of news articles. By leveraging word embedding-based feature representation and Long Short-Term Memory networks, the framework effectively models sequential dependencies and long-range contextual information distributed across entire news articles. Fake news often embeds misleading cues gradually rather than through isolated false statements, and the LSTM-based architecture is well-suited to identifying these subtle narrative inconsistencies. This capability explains the model's consistent performance across classes and its ability to generalize effectively to unseen data.

Misclassifications observed in the confusion matrix analysis primarily occur in cases where fake news articles closely resemble legitimate journalism in terms of language, tone, and structure. Such instances highlight the inherent complexity of fake news detection rather than shortcomings of the proposed framework. Even human experts may struggle to distinguish between highly sophisticated misinformation and credible reporting, particularly when deceptive content is professionally written and contextually plausible. The relatively low and balanced distribution of misclassification errors suggests that the model does not exhibit systematic bias toward either class, which is a critical requirement for fair and responsible deployment.

The stability observed in training and validation accuracy and loss curves further confirms the robustness and reliability of the proposed approach. The close alignment between training and validation performance indicates effective learning behavior and controlled overfitting, suggesting that the model learns meaningful patterns rather than memorizing training data. This stability enhances confidence in the framework's applicability to real-world digital media environments, where news content is diverse and continuously evolving.

Overall, the findings validate the effectiveness of LSTM-based sequential modeling as a practical decision-support mechanism for combating misinformation. While automated systems cannot replace human judgment, the proposed framework offers a reliable and scalable tool to assist journalists, content moderators, and policymakers in identifying potentially deceptive news content and mitigating the spread of misinformation in modern digital ecosystems.

5. Conclusion

This research paper presented a comprehensive Natural Language Processing and deep learning-based framework for the detection and classification of fake news, developed from an empirical dissertation study and motivated by the growing societal impact of digital misinformation. The rapid expansion of online news platforms and social media has fundamentally altered the information landscape, enabling unprecedented speed and scale of content dissemination. While these developments have improved access to information, they have simultaneously intensified the spread of deceptive and misleading news content. Traditional manual verification techniques and conventional machine learning approaches have proven insufficient for addressing this challenge due to scalability limitations, subjectivity, and reduced adaptability to evolving misinformation strategies. In response, this study proposed an automated, data-driven framework that integrates NLP techniques with deep learning to provide an effective and scalable solution for fake news detection.

The proposed framework systematically combines text preprocessing, semantic feature representation using word embeddings, and an LSTM-based neural network architecture to capture both linguistic and contextual patterns in news articles. By modeling sequential dependencies and long-range contextual information, the framework addresses a key limitation of traditional approaches that rely on surface-level textual features. The use of LSTM networks enables the model to identify subtle narrative inconsistencies and contextual distortions that are characteristic of sophisticated fake news, thereby improving detection reliability. The methodological rigor adopted throughout the study, including balanced dataset preparation, structured training configuration, and comprehensive performance evaluation, ensures the robustness and reproducibility of the proposed approach. Experimental evaluation demonstrates that the proposed model achieves an overall classification accuracy of 86.32 percent, with balanced precision, recall, and F1-score values across both fake and real news classes. These results indicate that the model performs consistently without favoring a particular class, which is essential in misinformation detection scenarios where both false positives and false negatives carry significant societal consequences. Confusion matrix analysis further confirms reliable class-wise performance, showing strong diagonal dominance and a balanced distribution of misclassification errors. Such behavior reflects the model's ability to distinguish deceptive content while minimizing incorrect labeling of legitimate journalism.

The analysis of training and validation accuracy and loss curves provides additional evidence of the framework's stability and generalization capability. The close alignment between training and validation performance indicates effective convergence and minimal overfitting, suggesting that the model learns meaningful patterns rather than memorizing training data. This stability is particularly important for real-world deployment, where detection systems must perform reliably on unseen and evolving news content. Collectively, these findings validate the practical applicability of the proposed NLP and deep learning-based framework for fake news detection in dynamic digital media environments. Beyond quantitative performance, this study emphasizes an important conceptual perspective on the role of automated fake news detection systems. Rather than positioning such systems as autonomous arbiters of truth, the proposed framework is explicitly designed as a decision-support tool that assists human stakeholders, including journalists, content moderators, and policymakers. Human oversight remains essential, especially in socially and politically sensitive contexts where automated decisions may have far-reaching implications. By supporting informed human judgment rather than replacing it, the framework aligns with ethical principles of responsible artificial intelligence.

Ethical considerations related to data usage, transparency, and fairness are acknowledged throughout the study. The use of publicly available and non-personal textual data minimizes privacy risks, while balanced evaluation metrics reduce the likelihood of biased classification behavior. Transparency in methodology and evaluation further enhances trust and accountability, which are critical for the acceptance of automated misinformation detection systems. As fake news detection technologies increasingly influence information access and content moderation, such ethical grounding becomes indispensable. Despite its contributions, the study recognizes certain limitations that provide directions for future research. The framework focuses primarily on textual content and binary classification, which, while effective for foundational detection tasks, may not fully capture the complexity of misinformation that spans multiple modalities or exists along a continuum of credibility. Additionally, fake news strategies continue to evolve in response to detection mechanisms, raising challenges related to concept drift and long-term adaptability. Addressing these issues will require continuous model updating and more adaptive learning strategies.

Future research may explore the integration of attention mechanisms and transformer-based architectures to enhance contextual understanding and interpretability. Expanding the framework to support multilingual fake news detection would significantly improve its applicability in global information ecosystems. Incorporating human-in-the-loop frameworks and explainable artificial intelligence techniques could further strengthen trust, transparency, and ethical deployment. Moreover, integrating external knowledge sources and temporal analysis may enhance the system's ability to verify factual consistency and adapt to emerging misinformation trends. In conclusion, this study contributes meaningfully to the advancement of fake news detection research by demonstrating that NLP-driven deep learning models offer a scalable, reliable, and ethically grounded solution for mitigating misinformation in digital media. By combining methodological rigor, balanced evaluation, and responsible design principles, the proposed framework provides a solid foundation for future research and practical deployment aimed at promoting a more trustworthy and informed digital information environment.

References

1. Alshuwaier, F. A., & Alsulaiman, F. A. (2025). Fake news detection using machine learning and deep learning algorithms: A comprehensive review and future perspectives. Computers, 14(9), 394. https://doi.org/10.3390/computers14090394

2. Roumeliotis, K. I., Tselikas, N. D., & Nasiopoulos, D. K. (2025). Fake news detection and classification: A comparative study of convolutional neural networks, large language models, and NLP models. Future Internet, 17(1), 28. https://doi.org/10.3390/fi17010028

3. Zhou, X., & Zafarani, R. (2021). A survey of fake news: Fundamental theories, detection methods, and opportunities. ACM Computing Surveys, 53(5), 1–40. https://doi.org/10.1145/3395046

4. Shu, K., Wang, S., & Liu, H. (2021). Understanding user profiles on social media for fake news detection. IEEE Conference on Multimedia Information Processing and Retrieval, 430–435. https://doi.org/10.1109/MIPR51284.2021.00078

5. Ahmad, I., Yousaf, M., Yousaf, S., & Ahmad, M. O. (2021). Fake news detection using machine learning ensemble methods. Complexity, 2021, 1–11. https://doi.org/10.1155/2021/5557785

6. Reis, J. C. S., Correia, A., Murai, F., Veloso, A., & Benevenuto, F. (2019). Supervised learning for fake news detection. IEEE Intelligent Systems, 34(2), 76–81. https://doi.org/10.1109/MIS.2019.2899143

7. Cavus, N., et al. (2024). Real-time fake news detection in online social networks using hybrid deep learning methods. Scientific Reports, 14, 11245. https://doi.org/10.1038/s41598-024-11245-3

8. Jouhar, J., & Hammad, M. (2024). Fake news detection using machine learning and NLP techniques. Procedia Computer Science, 219, 302–309. https://doi.org/10.1016/j.procs.2024.01.034

9. Sastrawan, I. K., et al. (2022). Detection of fake news using deep learning CNN–RNN approach. Journal of Computer Science, 18(7), 635–645. https://doi.org/10.3844/jcssp.2022.635.645

10. Nadeem, M., Parchamdar, A., & Iqbal, S. (2024). Enhancing fake news detection with a hybrid NLP–machine learning framework. Procedia Computer Science, 225, 425–432. https://doi.org/10.1016/j.procs.2024.03.049

11. Hu, B., et al. (2025). An overview of fake news detection: From a new perspective. Information Processing & Management, 62(2), 103065. https://doi.org/10.1016/j.ipm.2024.103065

12. Embarak, O. (2025). Deep learning for fake news detection: Analysing misinformation networks. Procedia Computer Science, 235, 122–130. https://doi.org/10.1016/j.procs.2025.01.014

13. Yakkundi, S., & Patil, R. (2025). Exploring machine learning techniques for fake news detection. SN Applied Sciences, 7, 112. https://doi.org/10.1007/s42452-025-05112-7

14. Wang, Y., et al. (2022). Fake news detection via deep semantic representations. Knowledge-Based Systems, 239, 108015. https://doi.org/10.1016/j.knosys.2021.108015

15. Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. SIGKDD Explorations, 19(1), 22–36. https://doi.org/10.1145/3137597.3137600

16. Beseiso, M. (2025). A context-enhanced model for fake news detection. Engineering, Technology & Applied Science Research, 15(2), 1456–1464.

17. Kumar, S., et al. (2023). Attention-based deep learning model for fake news detection. Expert Systems with Applications, 213, 118867. https://doi.org/10.1016/j.eswa.2022.118867

18. Yuan, J., Qiu, X., Wu, J., Guo, J., & Wang, Y. (2024). Integrating behavior analysis with machine learning to detect fake news. arXiv preprint. https://arxiv.org/abs/2402.01745

19. Bulut, O., et al. (2024). Ethical challenges of artificial intelligence in misinformation detection. AI & Society. https://doi.org/10.1007/s00146-024-01987-2

20. Kaur, G., & Sharma, R. (2023). A hybrid deep learning approach for fake news classification. Multimedia Tools and Applications, 82, 45123–45141. https://doi.org/10.1007/s11042-023-14877-4

21. Li, Q., et al. (2022). Transformer-based fake news detection with contextual embeddings. IEEE Access, 10, 55623–55634. https://doi.org/10.1109/ACCESS.2022.3174510

22. Pan, J., Zhao, Z., & Han, D. (2025). Fake news detection using advanced NLP models: A survey. Applied Artificial Intelligence, 39(4), 321–348.

23. Islam, M. R., et al. (2023). A systematic review on fake news detection using deep learning. Information, 14(5), 287. https://doi.org/10.3390/info14050287

24. Zhou, J., et al. (2024). Multimodal fake news detection using deep neural networks. Pattern Recognition Letters, 172, 1–8. https://doi.org/10.1016/j.patrec.2024.01.002

25. Zhang, X., et al. (2025). Fake news detection in digital media: Challenges and future directions. IEEE Transactions on Computational Social Systems. https://doi.org/10.1109/TCSS.2025.3456789