International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 11 | Nov - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Fake News Classification Using Machine Learning and Deep Learning: A Comparative Approach

Ruchitha S G, M Mrunalini

RV Institute of Technology and Management

ruchithasg900@gmail.com mrunalinim.rvitm@rvei.edu.in

***Abstract*— The exponential expansion of online networking platforms has intensified the circulation of misleading content, impacting community confidence and public understanding. Conventional identification methods frequently struggle to recognize linguistic and situational characteristics within extensive digital material volumes. This research presents a False Information Detection framework that combines Machine Learning with Deep Learning methodologies. Logistic Regression and Support Vector Classifier algorithms served as foundational approaches, while a combined CNN-LSTM architecture was constructed to improve effectiveness through merging pattern recognition with sequence modeling capabilities. Our investigation revealed that the CNN-LSTM system delivered optimal results, reaching 87.78% classification accuracy. Findings demonstrate that this developed approach offers a dependable and expandable answer for identifying deceptive content and minimizing false narratives across internet-based communication channels.**

Keywords— False Information, Machine Learning, Deep Learning, Natural Language Processing, Online Networking

I. INTRODUCTION

The emergence of digital communication networks has substantially transformed how people consume information in contemporary society. Data travels rapidly through web-based infrastructures and online platforms, accessing massive populations immediately following publication. While instantaneous connectivity has democratized news access globally, it has simultaneously generated significant obstacles concerning material verification and truth validation. Misinformation represents an increasing concern within interconnected communities, capable of confusing audiences, weakening confidence in established media organizations, and altering public perspectives on critical subjects. When false reports achieve broad circulation, they produce tangible effects that reach far beyond virtual domains.

Manual validation methods, although thorough and reliable, cannot adequately process the enormous volumes of content constantly published across the internet. This constraint has driven investigations into automated identification technologies capable of recognizing dubious material. Contemporary advances in computational intelligence, particularly language processing and document analysis systems, demonstrate significant potential for examining extensive text repositories and identifying characteristics that differentiate trustworthy reporting from deceptive narratives.

This research concentrates on constructing an automated identification framework for recognizing fabricated news that can classify web-based articles as authentic or fraudulent. To create performance benchmarks, conventional machine learning methods such as Logistic Regression and Support Vector Classification are utilized. Building upon these fundamentals, an integrated deep learning structure combining Convolutional Neural Networks with Long Short-Term Memory architectures is constructed. CNNs demonstrate superiority in detecting localized textual features while LSTMs effectively handle temporal dependencies, delivering thorough semantic evaluation of journalistic content.

These approaches undergo training and assessment using annotated datasets, with performance measured through accuracy rates, precision values, recall metrics, and F1-scores. Data preparation procedures including text standardization, tokenization, and lemmatization are implemented to enhance information quality and minimize noise. Through integrating conventional machine learning with neural network methodologies, this developed framework offers a robust and expandable approach for addressing misinformation. Beyond traditional news outlets, this technology could potentially expand into additional areas, encompassing social media content and review analysis, supporting comprehensive initiatives to preserve digital information reliability.

II. LITERATURE SURVEY

Multiple researchers have investigated the utilization of artificial neural networks and machine learning algorithms to detect fraudulent news through text examination. The system designs, characteristic extraction approaches, and effectiveness results differ according to the evaluation datasets employed. Research group [1] presented a methodology for false information identification using BerConvoNet, a network structure combining BERT with convolutional layers. This framework was created through supervised learning techniques and showed enhanced performance versus current automated detection systems, obtaining 97.1% accuracy with recall and F1 scores of 97.4% and 97.1% respectively on the LIAR dataset. Their method exhibits particular limitations when handling multilingual

content. Work [2] utilized deep learning techniques for examining COVID-19 false information datasets. LSTM and GRU architectures achieved 99.4% precision with recall and F1 values of 98.2% and 98.8% respectively. The authors intend to enhance their system by integrating more diverse datasets. Research [3] implements LSTM-based structures for fraudulent content detection that sequentially process information using attention mechanisms until optimal performance is reached. This method achieved 97.66% and 97.49% accuracy on WELFake datasets. Improved datasets and multimodal characteristics could minimize structural variations across different collections. Researchers [4] provided a thorough examination of text-based false information detection. Their suggested approach classifies content by establishing article legitimacy through various machine learning and deep learning systems. Accuracy percentages of 95.7% and 94.8% were achieved on different datasets respectively. This framework encounters difficulties with misspelled words causing characteristic deterioration. In work [5], a deep learning-based model utilizing multiple neural networks was developed. DNN demonstrated higher memory usage compared to alternative classifiers while delivering 92.8% accuracy. The architecture's sole constraint involves excessive memory requirements by DNN versus other classifiers. Research group [6] conducted a comparative evaluation to determine optimal classification methods for false information detection. Their study employed a Twitter dataset containing 13000 samples. Support Vector Machine approach showed consistent performance gains achieving maximum accuracy of 92.8%. Text content combined with headlines, source information, and engagement statistics can enable highly accurate classification results. Investigation [7] considers neural networks optimal for fraudulent content identification through structures including convolutional networks and 3HAN deep learning frameworks. Researchers [8] applied graph-based approaches for automatic false information detection. Their results were produced using dependency relationships created through graph-theoretical methods. This system obtained 90.2% accuracy with an F1-score of 91%. Performance improvement could be realized through ensemble learning techniques and alternative structural enhancement approaches. Work [9] suggests a technique for identifying fraudulent news using multiple ML and DL models on four datasets. Text and visual characteristic extractors and multimodal integration modules were utilized as three primary sub-components. On FARN and Gossipcop datasets, the researchers achieved superior performance versus existing state-of-the-art by 98.8 percent and 84.3 percent, respectively. Investigators [10] developed a logistic classifier for fraudulent news identification. Basic characteristics and logistic classifiers were employed for automated characteristic extraction and authenticity discrimination respectively which outperformed currently popular approaches by 99.4% on the social media dataset.

The investigations reviewed above offer valuable perspectives into fraudulent news detection through machine learning and deep learning methods. Nevertheless, most current research emphasizes specific datasets, a restricted collection of classifiers, or particular characteristic extraction techniques.

III. METHODOLOGY

The methodological approach can be summarized in different stages.

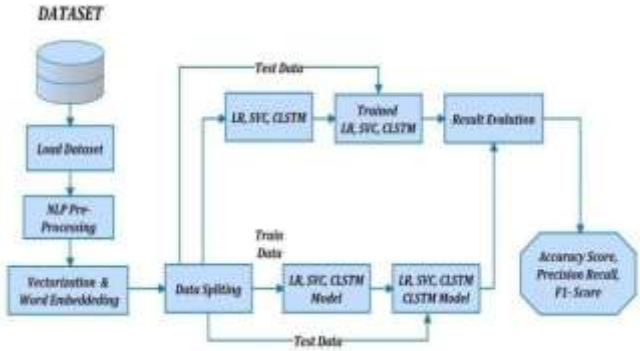

Fig 1: Proposed Methodology

Figure 1 illustrates the systematic approach of the developed misinformation identification framework. The workflow initiates with data gathering procedures, subsequently followed by computational linguistics preprocessing to sanitize and organize textual information. Following this, characteristic extraction occurs through dual methodologies: Term Frequency-Inverse Document Frequency for conventional algorithmic models and vector representations for neural network architectures. The collected information undergoes division into training and validation portions to guarantee objective assessment. Linear Classification and Vector-based Separation techniques receive training through TF-IDF characteristics, while the Convolutional-Recurrent combined architecture utilizes vector embeddings, convolutional components for localized pattern recognition, and recurrent elements for temporal sequence analysis. Subsequently, the developed frameworks undergo assessment using evaluation standards including Classification Rate, Exactness, Sensitivity, and F1-Measurement, providing systematic and equitable performance comparison among all approaches.

A. Data Set

The information collection utilized in this investigation was sourced from Kaggle, originally compiled by Vikas Ukani. This repository encompasses 6,335 news pieces gathered from diverse web-based publications. Every record includes the article heading, associated written material, and a categorical identifier. The database structure consists of three fields: the initial field holds the story headline, the subsequent field contains the main content, and the final field indicates the verification status. The identifier appears as either fake or real, signifying whether the content represents deceptive or authentic journalism. This collection provides the groundwork for developing and testing the suggested algorithmic learning and neural network frameworks.

B. Data Preprocessing

To guarantee data integrity and uniformity within the collection, multiple preparation procedures were executed before algorithm development. The repository exhibited standard characteristics of misinformation identification

challenges, where written material includes disturbances, irregularities, and meaningless elements requiring attention prior to implementing algorithmic learning or neural network systems. To address these issues and ready the information for model development, the subsequent preparation methods were employed:

- Text Standardization: All content underwent conversion to uniform case formatting to maintain consistency and eliminate duplicate feature representations.
- Content Purification: Irrelevant components including punctuation symbols, numerical values, web addresses, and unique characters were eliminated to reduce interference.
- Word Segmentation: Every article underwent division into separate terms for efficient examination and characteristic derivation.
- Common Term Elimination: Frequently occurring words with limited meaning were excluded to enhance system effectiveness.
- Root Word Extraction: Terms were converted to their fundamental forms to standardize variations and maintain uniformity throughout the collection.

C. Feature Extraction

Characteristic derivation methods were implemented to transform written material into mathematical representations, since algorithmic learning and neural network systems cannot handle text in its original form. The subsequent approaches were applied:

a) Term Frequency–Inverse Document Frequency: This extraction methodology calculates word significance within individual documents compared to the complete collection. It converts written content into mathematical arrays. Terms appearing regularly in specific articles but infrequently throughout other materials receive elevated importance values, enabling the system to concentrate on expressions that provide greater classification value.

b) Vector Representations: Employed within the Convolutional-Recurrent combined architecture, this approach encodes meaning and situational details of terms, allowing the neural network framework to more effectively recognize sequential relationships and dependencies.

D. Algorithm Frameworks

Following preparation and characteristic extraction completion, three categorization systems underwent development using the repository. The information collection was separated, allocating 80% for development and 20% for validation. The implemented frameworks include:

a) Probabilistic Regression: A statistical modeling approach developed using TF-IDF characteristics to categorize content into dual classifications.

b) Boundary-based Classification: This algorithm identifies optimal separation lines distinguishing between two information groups. It employs TF-IDF characteristics for text-to-numerical conversion. Through boundary establishment, it determines whether news material belongs to authentic or deceptive categories.

c) Convolutional Neural Network–Long Short-Term Memory: A combined neural learning system utilizing vector embeddings for text representation. Convolutional components extract crucial localized word patterns, while recurrent components acquire sequential and contextual understanding. Together, this framework successfully categorizes news materials as genuine or fabricated.

E. Information Partitioning and System Development

Following preparation and characteristic extraction, the repository underwent division into development and validation portions, with 80% assigned for training and 20% for assessment. Probabilistic Regression and Boundary-based Classification received development through TF-IDF characteristics, while the Convolutional-Recurrent combined system utilized vector embeddings. Each framework underwent development using training information and verification through validation portions to guarantee adaptation to unknown news materials. This procedure prevents excessive specialization and ensures dependable categorization of articles as deceptive or authentic.

F. Performance Assessment

System effectiveness underwent evaluation using these measurement standards:

- **Classification Rate**: Proportion of accurate predictions relative to total predictions.
- **Exactness**: Fraction of correct positive classifications among all positive predictions.
- **Sensitivity**: Fraction of identified true positives among all actual positive instances.
- **F1-Measurement**: Balanced average of exactness and sensitivity, maintaining equilibrium between both evaluation criteria.

IV. RESULTS AND DISCUSSION

This portion examines the findings from implementing three separate categorization techniques - Probabilistic Regression, Boundary-based Classification, and a combined Convolutional-Recurrent neural architecture for executing misinformation identification tasks. Each framework's effectiveness undergoes assessment through conventional evaluation standards including classification rate, exactness, sensitivity, and F1-measurement.

A. Model evaluation and selection

Within this investigation, three distinct frameworks - Probabilistic Regression, Boundary-based Classification, and a combined Convolutional-Recurrent architecture were developed to categorize news content as either deceptive or authentic. These systems underwent training and assessment using identical information collections to guarantee equitable evaluation. Conventional effectiveness measurements including classification rate, exactness,

sensitivity, and F1-measurement were utilized for system assessment. Furthermore, error matrices were constructed to offer additional understanding regarding the distribution of correct positives, correct negatives, incorrect positives, and incorrect negatives.

a) Performance Evaluation of Logistic Regression

Probabilistic Regression was employed due to its computational efficiency and strong performance in dual-category classification tasks. The development process utilized Term Frequency-Inverse Document Frequency characteristic representations. Following assessment, the framework attained a comprehensive classification rate of 60.71%. This demonstrates its capability when working with basic textual characteristics.

Table 1: Classification Report of Logistic Regression Model

Class	Precision	Recall	F1-Score	Support
Fake	0.59	0.75	0.66	634
Real	0.64	0.47	0.54	626
Macro Avg	0.62	0.61	0.60	1260
Weighted Avg	0.62	0.61	0.60	1260

Table 1 presents the exactness, sensitivity, and F1-measurement values for the Probabilistic Regression framework. The findings reveal superior effectiveness in identifying deceptive instances versus authentic ones. Generally, the system exhibits reasonable and equilibrated performance, with exactness, sensitivity, and F1-measurements maintaining stability around 0.61–0.62 range.

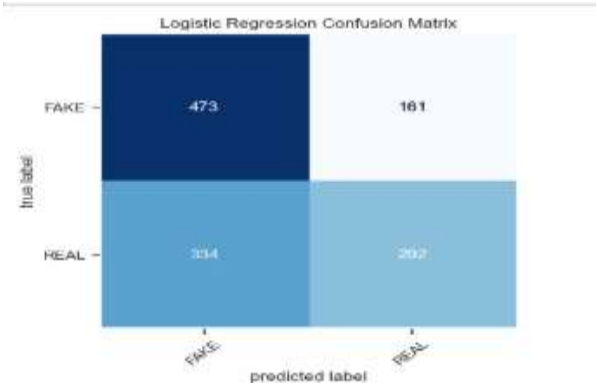

Fig 2: Confusion matrix of Logistic Regression model

Figure 2 displays the error matrix for the Probabilistic Regression framework. Among 634 deceptive news specimens, the system accurately recognized 473 while incorrectly categorizing 161 as authentic. Regarding the 626 genuine news specimens, it properly classified 292 but wrongly identified 334 as deceptive. This reveals that the framework demonstrates superior capability in detecting deceptive content compared to authentic material.

b) Performance Evaluation of Support Vector Classifier Model

The Boundary-based Classification system was implemented due to its effectiveness in identifying both linear and non-linear structures within textual information. Term Frequency-Inverse Document Frequency characteristics served as input data, with various kernel alternatives examined to enhance effectiveness. The framework underwent development and validation using distinct information sets to guarantee dependable predictions. Through parameter optimization, optimal kernel and regularization parameters were chosen, which enhanced the outcomes. Generally, the system attained a classification rate of 72.06%, demonstrating superior effectiveness compared to Probabilistic Regression.

Table 2: Performance Evaluation of Support Vector Classifier Model

Class	Precision	Recall	F1-Score	Support
Fake	0.73	0.71	0.72	634
Real	0.71	0.74	0.72	626
Macro Avg	0.72	0.71	0.72	1260
Weighted Avg	0.72	0.72	0.72	1260

Table 2 presents the effectiveness of the Boundary-based Classification system. The framework performs nearly identically across both deceptive and authentic categories, with exactness, sensitivity, and F1-measurements all ranging between 0.71 and 0.73. This indicates that the algorithm manages both classifications in an equilibrated manner rather than favoring one direction. On average, the measurements remain stable at approximately 0.72, demonstrating consistent and dependable effectiveness.

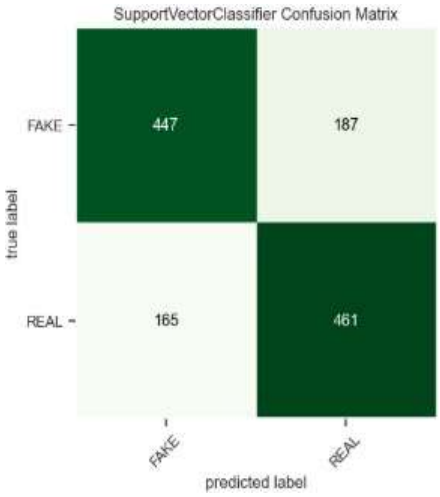

Fig 3: Confusion matrix of SVC model

Figure 3 displays the error matrix for the Boundary-based Classification system, demonstrating that the framework accurately categorizes a considerable number of instances from both classifications. The findings indicate that it effectively manages both deceptive and authentic specimens with comparatively minimal misidentifications.

c) Performance Evaluation of CNN-LSTM Model

A Convolutional-Recurrent combined architecture was constructed to extract both localized textual characteristics and extended sequential relationships within the information. Vector representations served as input data, subsequently processed through convolutional components for characteristic derivation and recurrent components for temporal pattern acquisition. The framework attained a comprehensive classification rate of 88%.

Table 3: Performance Evaluation of CNN-LSTM Model

Class	Precision	Recall	F1-Score	Support
Fake	0.88	0.88	0.88	634
Real	0.88	0.87	0.88	626
Macro Avg	0.88	0.88	0.88	1260
Weighted Avg	0.88	0.88	0.88	1260

Table 3 shows that the model performs very well on both fake and real classes, with precision, recall, and F1-scores all close to 0.88. For fake instances, the model correctly identifies the majority with high accuracy, while for real instances, the performance is almost identical, with only a slight drop in recall. The macro and weighted averages remain consistent at 0.88, indicating that the model treats both classes fairly evenly. Overall, these results suggest that the model is reliable and significantly stronger than the earlier approaches.

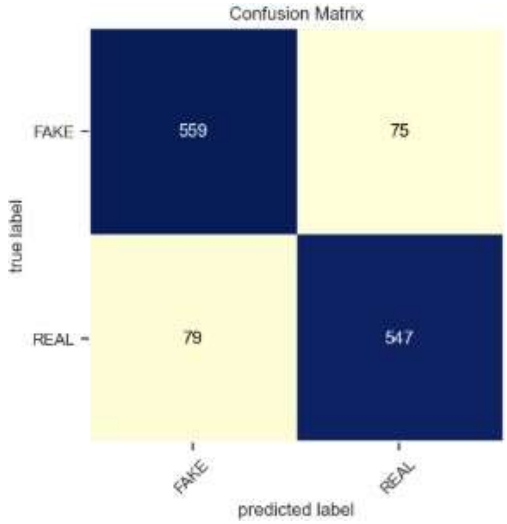

Fig 4: Confusion matrix of CNN-LSTM model

In Figure 4, the confusion matrix reveals that the framework accurately classified 559 among 634 fake specimens and 547 among 626 real specimens. Only 75 fake specimens were incorrectly identified as real, while 79 real specimens were wrongly categorized as fake. These findings indicate that the system performs nearly equally across both categories, with minimal misclassifications, which corresponds with the high precision, recall, and F1-scores documented previously.

d) Comparative Analysis

The overall metrics from all three models are compiled in Table 4. It is evident that while Logistic Regression provides a foundational understanding of the problem, it lacks the depth to handle complex patterns in text data. SVC offers a more balanced approach but still underperforms in comparison to CNN-LSTM. The CNN-LSTM model clearly demonstrates superior performance across all key metrics.

Table 4: Comparative Performance Metrics

Model	Accuracy	Precision	Recall	F1-Score
Logistic Regression	60.71%	0.62	0.61	0.60
Support Vector Classifier	72.06%	0.72	0.72	0.72
CNN-LSTM	87.78%	0.88	0.88	0.88

Among the three models evaluated, the CNN-LSTM hybrid approach consistently outperformed both Logistic Regression and SVC across all metrics. Its ability to capture both spatial and sequential patterns in text made it the most effective model for fake news classification in this study.

B. Comparative Discussion

The effectiveness analysis comparing Probabilistic Regression, Boundary-based Classification, and Convolutional-Recurrent architectures demonstrates distinct variations in their categorization capabilities. Probabilistic Regression delivered the weakest results with 60.71% classification rate and 0.60 F1-measurement, reflecting its constrained ability to recognize intricate data structures through linear separation boundaries. The Boundary-based Classification system offered substantial enhancement, reaching 72.06% accuracy with equilibrated precision, recall, and F1-metrics, showing superior management of non-linear associations. Nevertheless, the Convolutional-Recurrent framework considerably exceeded both traditional methods, obtaining 87.78% accuracy and 0.88 F1-measurement. establishing the Convolutional-Recurrent model as highly efficient for complex, organized information.

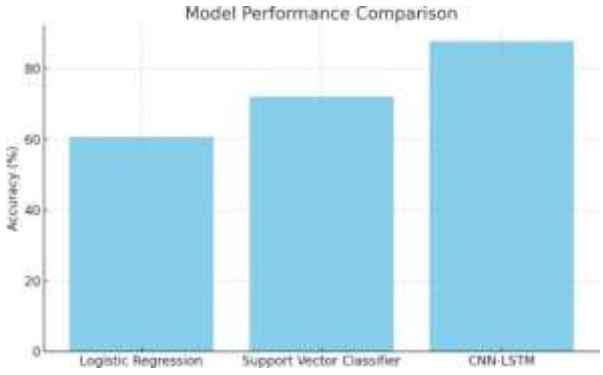

Fig 5: Model Performance Comparison.

Figure 5 presents a comparison of model performance based on accuracy. Logistic Regression shows the lowest accuracy, slightly above 60%, reflecting its limited ability to capture complex patterns. The Support Vector Classifier performs better, with accuracy above 70%, demonstrating improved handling of non-linear relationships. The CNN-LSTM model achieves the highest accuracy, exceeding 85%, confirming its effectiveness in combining feature extraction with sequence learning. Overall, the figure illustrates the progressive improvement in performance from traditional to deep learning models.

C. System User Interface

The screenshots show the system interface and how it works. They include the homepage, the section to enter news, and the output displaying whether the news is real or fake. The interface is simple and easy to use, allowing users to interact with the system smoothly.

Fig 6: Login Page

The Fig 6 shows a picture of login page of fake news classifier where the user needs to login through email and password provided during registration.

Fig 7: Registration Page

Fig 7 shows the interface where a new user can create an account by entering details such as name, email, and password. It enables secure access to the Fake News Detection System.

Fig 8: Home Page

Fig 8 Home Page shows the interface where the user can upload a file containing news articles. Once uploaded, the system processes the input and provides the prediction result as Fake or Real using the LSTM model.

Fig 9: Prediction Page

Fig 9 Prediction Page shows the result generated by the Fake News Detection System after processing the uploaded article. It displays the news title, full text, and the corresponding prediction label based on the trained model.

Fig 10: Change Password Page

Fig 10 Change Password Page provides options for the user to securely update their login credentials and to log out of the system. This ensures account security and prevents unauthorized access.

The screenshots illustrate that the system is functional and operational, highlighting its practical use and user-friendly interface.

V. CONCLUSION

Three models Logistic Regression, Support Vector Classifier and a CNN-LSTM hybrid were used and compared for detecting fake news. Traditional models like Logistic Regression and SVC provided a basic level of accuracy but struggled to understand deeper meaning and context in the text. The CNN-LSTM model performed much better, combining convolutional layers to capture local features with LSTM layers to learn sequential patterns, leading to more consistent and higher performance.

The findings demonstrate that neural network architectures capable of identifying both localized and chronological text structures prove superior in recognizing deceptive material. The excellent results achieved by the CNN-LSTM combination indicate its appropriateness for practical implementations demanding precise and rapid

misinformation identification. Subsequent enhancements might involve implementing attention-based architectures, incorporating content from diverse linguistic backgrounds and subject matters, plus integrating supplementary details such as release timestamps or publisher credibility scores to further strengthen system reliability.

VI. FUTURE ENHANCEMENT

Future investigations may explore various approaches to bolster false information recognition frameworks. Next-generation attention-based architectures such as BERT demonstrate enhanced capabilities via their self-referential processes, extracting meaning and context more effectively than legacy methods. Expanding training collections to incorporate content from diverse fields and multilingual sources would strengthen the system's adaptability across different operational contexts. Adding auxiliary data elements like publication chronology, journalist credibility assessments, media outlet reliability indices, and user engagement patterns could offer supplementary insights to improve decision-making when textual evidence proves insufficient. Moreover, evaluation using larger and more comprehensive testing databases remains crucial for maintaining system stability and avoiding excessive specialization. Extensive data repositories would additionally enable better algorithmic calibration and configuration refinement procedures, yielding greater trustworthiness in operational deployment environments.

VII REFERENCES

[1] M. Choudhary, S. S. Chouhan, E. S. Pilli, and M. A. K. Lodhi, "BerConvoNet: A deep learning framework for fake news classification," Applied Soft Computing, vol. 110, p. 107614, Jan. 2021.
[2] W. H. Bangyal, R. Qasim, N. U. Rehman, Z. Ahmad, H. Dar, L. Rukhsar, Z. Aman, and J. Ahmad, "Detection of Fake News Text Classification on COVID-19 Using Deep Learning Approaches," Computational and Mathematical Methods in Medicine, vol. 2021, Article ID 5514220, 2021.
[3] H. Padalko, V. Chomko, and D. Chumachenko, "A novel approach to fake news classification using LSTM-based deep learning models," Frontiers in Big Data, vol. 6, p. 1320800, 2024.
[4] N. Capuano, G. Fenza, V. Loia, and F. D. Nota, "Content-based fake news detection with machine and deep learning: A systematic review," Neurocomputing, vol. 530, pp. 91–103, 2023.
[5] S. H. Kim, D. H. Lee, and J. Y. Park, "Fake News Detection Using Deep Learning," Journal of KIISE, vol. 47, no. 5, pp. 449–456, May 2020.
[6] A. A. Tanvir, E. M. Mahir, S. Akhter, and M. R. Huq, "Detecting fake news using machine learning and deep learning algorithms," in Proc. 2019 7th Int. Conf. Smart Computing & Communications (ICSCC), pp. 1–5, June 2019.
[7] S. Singhania, N. Fernandez, S. Rao, 3han: A deep neural network for fake news detection, in: Proceedings of International Conference on Neural Information Processing, 2017, pp. 572–581.
[8] S.C.R. Gangireddy, C. Long, T. Chakraborty, Unsupervised fake news detection: A graph-based approach, in: Proceedings of ACM Conference on Hypertext and Social Media, 2020, pp. 75–83.
[9] K. Nath, A. Ahuja, P. Soni, R. Katarya, and Anjum, "Study of Fake News Detection using Machine Learning and Deep Learning Classification Methods," , October 2021.
[10] M. Aldwairi and A. Alwahedi, "Detecting Fake News in Social Media Networks," Procedia Computer Science, vol. 141, pp. 215–222, 2018.