International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 2, Mar-Apr-2026, ISSN (Online): 2395-566X

TruthShield-ML – An Intelligent Machine Learning Framework for Accurate Fake News Detection and Misinformation Analysis

Mrs.K.Ganga Devi Bhavani[1] , Bonam Geetha Chitti Jyothi[2], Siravapu Santhi Kumari[3], Karri Manikanta Sai[4], Alluri Sri Akshay Satya Srinivas[5], Thella Aditya[6]
[1]Assistant Professor, [2, 3, 4, 5,6]B.tech Students Department of CSE,
Pragati Engineering College, Surampalem , Andhra Pradesh, India

Abstract- The spread of fake news has become a significant concern in today's society, as misleading information can easily damage reputations and lives. To address this issue, researchers have developed fake news detection systems using machine learning techniques. The identification of fake news is rapidly gaining traction and is increasingly being adopted by various industries, either for their own use or to offer as a service to others. Machine learning (ML) and deep learning (DL) are two prominent approaches employed to determine the authenticity of news. There are various methods available for detecting false news through both ML and DL techniques. This paper presents a comprehensive analysis of fake news detection using machine learning approaches. Upon thorough examination, it was found that several ML and DL algorithms have been applied in this domain, with the Support Vector Machine (SVM) being the most commonly used ML method, and Long Short-Term Memory (LSTM) being the most widely applied DL technique.

Keywords – Fake news detection, Machine learning, Deep learning, Support Vector Machine, Long Short-Term Memory, News authenticity, Text classification, Misinformation detection.

I. INTRODUCTION

In today's digital age, the proliferation of fake news is a significant concern, as false information can spread rapidly and have devastating consequences for individuals and communities [1]. Misinformation, often misleading or fabricated, poses as legitimate news, affecting public trust and perception [2]. The internet has opened up vast opportunities, but it has also led to a shift in how news is consumed, especially among younger generations who increasingly turn to online platforms instead of traditional news outlets like newspapers [3].

The proportion of people getting their news online has grown significantly, with 62% of Americans using the internet for news in 2016, compared to 49% in 2012 [4]. Social media now plays a dominant role in our lives [5], becoming a key tool for communication and information sharing [6]. However, this has also contributed to the rapid spread of fake news [5]. Fake news (FN) refers to deliberately misleading information spread through social media and other platforms, with the goal of deceiving users. It can take various forms, such as images, videos, or audio.

Researchers worldwide are increasingly focused on methods to identify whether a piece of content is real or fraudulent. In recent years, automated detection of fake news, particularly on social media, has emerged as a critical area of study [8]. Fake news often flourishes when publishers prioritize profit over accuracy, and when consumers are predisposed to believe information that aligns with their existing views [9]. This misinformation can arise from individual sources or coordinated efforts.

The motives behind the creation of fake news are often personal, political, or economic in nature. With the widespread use of social media as a primary news source, the rapid spread of misinformation is a growing issue. There are various types of misinformation, such as government-related rumors, global events, urban legends, and tragic incidents [10]. To combat this, fake news detection has become a crucial research topic. For instance, with Facebook's 2.91 billion active users, the spread of fake news is alarmingly fast and far-reaching.

Social media platforms and online news outlets now serve as primary sources of news, making it easier to share breaking stories [11]. A case from Bengaluru on March 30, 2018, illustrates the dangers of fake news. Mahesh Vikram Hegde, co-founder of 'Postcard News,' was arrested for spreading fake news about an incident involving a Jain monk. With 778,000 followers on social media, Hegde's false post about a Muslim attack on the monk led to widespread outrage and was

shared across various platforms. After an investigation, it was revealed that the incident was a misunderstanding, and the news was fabricated to create division and discredit the government. Hegde was arrested for inciting communal tension through fake news. Such deliberate misinformation often serves as a political tool to manipulate public opinion and shape political strategies, detracting from fact-based discourse [12].

Today, celebrities, politicians, and social media influencers are often implicated in spreading fake news, whether for financial gain or other motives [13]. However, anyone who shares such misleading content contributes to the problem. To address this issue, a fake news detection system has been developed [14]. For instance, WhatsApp has become a major source of misinformation, especially when posts lack credible sources. Verifying the news via a Google search and checking the reliability of sources is essential for confirmation.

Fake news also influences how people interpret and respond to real news, as it can introduce doubt and confusion, making it harder for individuals to distinguish between truth and falsehood [15]. In some cases, Photoshop and other editing tools are used to alter images and create fake news. To verify the authenticity of images, one can use reverse Google image search to check the source and original version of an image.

The 2016 US Presidential elections highlighted the impact of fake news when it was alleged that Russia used counterfeit accounts and bots to spread misleading information [16]. Studies such as Mykhailo Granik et al. [17] use Bayes' Theorem as a strategic approach for fake news detection. Machine learning offers a range of algorithms to detect fake news by analyzing data patterns and making predictions. The three primary paradigms in machine learning are supervised learning, unsupervised learning, and semi-supervised learning. Supervised learning methods include algorithms like Decision Trees (DT), Linear Regression (LR), and Logistic Regression (LR), while unsupervised learning includes techniques like hierarchical clustering [18].

A hybrid convolutional neural network (CNN) model developed by Wang [19] has shown superior performance compared to other machine learning methods. Additionally, research by Hannah Rashkin et al. [20] and Singhania et al. [21] employs hierarchical attention models to analyze linguistic aspects of fake news. Another approach, the characteristics stability index model by Ruchansky et al. [22], uses user behavior to evaluate the reliability of news content.

II. LITERATURE SURVEY

This paper investigates the classification of news articles sourced from social media through binary classification. By leveraging Artificial Intelligence (AI), Natural Language Processing (NLP), and Machine Learning (ML), the study aims to categorize news as either authentic or false [7]. The primary focus of the research is on identifying fake news tweets on Twitter by applying a range of machine learning algorithms [10]. Five distinct ML techniques—Support Vector Machine (SVM), Logistic Regression (LR), Naive Bayes (NB), and Recurrent Neural Network (RNN)—are compared to determine the most effective model [11]. The paper introduces various ML algorithms for fake news detection and conducts a benchmark analysis across three datasets to evaluate their performance.

Moreover, advanced Deep Learning (DL) techniques are incorporated alongside both supervised and unsupervised ML approaches. The key objective of this research is to improve the accuracy of fake news detection [12]. The study presents a model with enhanced capabilities for identifying fake news. Three ML models—Random Forest, Decision Tree, and Extra Tree Classifier—are utilized, achieving 99.8% accuracy on the ISOT dataset for testing and 44.15% for training.

This paper also demonstrates the classification of news articles using a machine learning ensemble approach [13]. It leverages linguistic features to distinguish between fake and real news, applying these methods across four different datasets. The study emphasizes that humans often struggle to detect fake news manually. It utilizes various machine learning algorithms for fake news classification, including Naive Bayes (NB), Passive-Aggressive Classifiers, and Deep Neural Networks [35].

In total, eight datasets are used to evaluate the accuracy of the methods. This research sheds light on how fake news spreads and the motivations driving its creation. By employing AI, NLP, and ML, binary classification is performed on online news to differentiate between fake and genuine stories [15]. The goal of the paper is to empower readers with the tools to classify news as either true or false. It also discusses the importance of identifying fake news and the impact it has across different sectors, highlighting several algorithms used to detect fake news on social media platforms. The study compares three distinct modules: text classifiers, stance detection software, and fact-checking methods currently in use to identify misinformation.

III. SYSTEM ANALYSIS

Existing System

Current approaches to detecting fake news mainly rely on machine learning (ML) and deep learning (DL) methods to categorize news articles as either real or fake. These systems commonly employ Natural Language Processing (NLP) techniques to process textual data from a variety of sources, such as social media platforms, news websites, and online portals. Widely used ML algorithms include Support Vector

Machines (SVM), Logistic Regression (LR), Naive Bayes (NB), Random Forest (RF), and Decision Trees. Additionally, deep learning models, such as Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNN), and deep neural networks (DNN), are frequently applied to enhance classification accuracy.

Existing systems are trained on diverse datasets, including ISOT, LIAR, and Twitter-based data, to provide a broad range of examples for fake and real news, which are essential for evaluating model performance. Features considered for fake news detection include linguistic cues, emotional tones, and user behaviors. While these systems often achieve high accuracy—particularly ML models like SVM and LR, which can exceed 95% in some cases—they still face significant challenges in identifying more advanced forms of fake news, such as deepfakes and synthetic media. Additionally, the rapid and ever-changing nature of misinformation presents ongoing obstacles, requiring continuous improvements in detection methods.

Limitations

1. **Detection of Advanced Fake News:** Current systems struggle with identifying complex types of fake news, such as deepfakes and synthetic media, which require more advanced methods of feature extraction and classification.

2. **Dynamic Nature of Misinformation:** The constantly evolving nature of misinformation presents a challenge for existing systems. Models trained on specific datasets may not generalize well to new forms of fake news or emerging patterns.

3. **Dependency on Dataset Quality:** The performance of these systems is heavily reliant on the quality and variety of training datasets, such as ISOT, LIAR, and Twitter data. A lack of diversity in datasets limits the system's ability to detect various types of misinformation.

4. **High Computational Demands:** Deep learning models like LSTM, CNN, and DNN require substantial computational resources, which makes them less suitable for real-time or resource-constrained applications.

5. **Limited Contextual Understanding:** While existing systems focus on linguistic patterns and user behavior, they often fail to fully understand the context or intent behind the information, which can lead to incorrect classifications.

Proposed System

The proposed system aims to improve the accuracy and reliability of fake news detection by incorporating advanced machine learning (ML) and deep learning (DL) techniques alongside enhanced feature extraction methods. This system leverages Natural Language Processing (NLP) to analyze textual content from various sources, such as social media, news websites, and online portals. Unlike existing systems, this approach focuses on integrating multiple classification models to achieve better generalization and improved accuracy.

It utilizes traditional algorithms like Support Vector Machines (SVM), Logistic Regression (LR), and Random Forest (RF), combined with advanced deep learning models such as Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNN). Moreover, the proposed system introduces additional features, including sentiment analysis, temporal patterns, and social context, to detect misinformation more effectively. By employing ensemble learning techniques and real-time data processing, the system addresses the limitations of current approaches, such as challenges in detecting deepfakes and adapting to dynamic misinformation trends. The integration of metaheuristic optimization methods further fine-tunes model parameters, ensuring robust performance across different datasets and use cases. Ultimately, this system aims to provide a scalable, efficient, and accurate solution for detecting and mitigating the spread of fake news.

Advantages

1. **Improved Accuracy and Reliability:** By combining multiple machine learning (ML) and deep learning (DL) models, the proposed system enhances both accuracy and generalization, overcoming the limitations of individual models and improving overall performance.

2. **Comprehensive Feature Extraction:** Advanced feature extraction methods, such as sentiment analysis, temporal patterns, and social context, allow the system to capture more nuanced aspects of fake news, resulting in better detection capabilities.

3. **Multimodal Classification Models:** The combination of traditional machine learning algorithms (SVM, LR, RF) and advanced deep learning models (LSTM, CNN) creates a robust detection system that leverages the strengths of various techniques to improve the accuracy of fake news identification.

4. **Ensemble Learning:** Ensemble learning techniques combine the predictions of multiple models, reducing bias and variance, thereby enhancing the system's robustness and accuracy.

5. **Real-time Data Processing:** The system's ability to process data in real-time allows for the prompt detection and response to emerging fake news, ensuring timely interventions.

6. **Handling Dynamic Misinformation:** The system is designed to adapt to the evolving landscape of fake news and deepfakes, ensuring its long-term effectiveness in detecting a wide range of misinformation types.

IV.SYSTEM DESIGN

System Architecture

Below diagram depicts the whole system architecture.

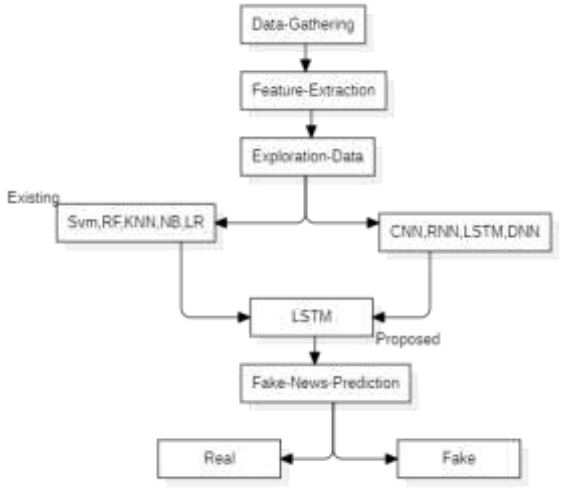

Fig 1. Methodology followed for proposed model

V. SYSTEM IMPLEMENTATION

Modules

Data Pre-processing

The first stage of the system's implementation involves gathering data from a variety of online sources, including social media platforms, news websites, and online portals. These sources offer a diverse dataset of both authentic and fake news articles. The data collection process aims to cover a wide range of topics to ensure the system can effectively detect misinformation across various domains.

After data collection, the pre-processing phase begins, where irrelevant content such as HTML tags and special characters are eliminated. The text is then normalized through methods like converting text to lowercase, stemming, and lemmatization. Any missing data is appropriately handled to ensure the dataset is clean and ready for further analysis.

Model Selection

The system utilizes a combination of traditional machine learning (ML) models and deep learning architectures to classify news articles as either real or fake. Support Vector Machines (SVM) are employed for their capability to classify text by determining a hyperplane that separates real news from fake news in a high-dimensional feature space. Logistic Regression (LR) is also used, providing probabilities of an article being fake based on the extracted features.

Random Forest (RF), an ensemble learning method, creates multiple decision trees and makes predictions based on the majority vote, thus enhancing the system's accuracy and robustness. In addition to these traditional models, deep learning techniques such as Long Short-Term Memory (LSTM) networks are applied to capture long-term dependencies and contextual information in text sequences. Convolutional Neural Networks (CNN) are utilized for their proficiency in identifying local patterns in text, further improving the system's feature extraction capabilities for text classification.

Ensemble Learning

To optimize the system's performance, ensemble learning methods are employed. By combining multiple models, these techniques help reduce bias and variance, ensuring more reliable predictions. One approach is the Voting Classifier, where each individual model generates a prediction, and the final output is determined by the majority vote. Another method used is stacking, where the predictions from individual models are passed to a meta-classifier, which learns how to best combine them for the final classification. This ensemble approach strengthens the overall reliability of the system, making it more adaptable to a variety of fake news detection scenarios.

User Interface and Reporting

The system includes a user-friendly web interface designed to allow users to input data and receive real-time recommendations. The module also provides visualizations and reports that present the results in an easily understandable format. This interface is tailored to be accessible to users with minimal technical knowledge, ensuring broad usability and effectiveness.

Evaluation and Performance Metrics

To evaluate the performance of the system, several classification metrics are used, including accuracy, precision, recall, and F1-score. These metrics offer valuable insights into the system's ability to distinguish between real and fake news. Accuracy measures the overall rate of correct classifications, while precision and recall specifically assess how well the system identifies fake news. The F1-score, which balances precision and recall, is used to evaluate the overall performance. Additionally, K-fold cross-validation is utilized to ensure the model generalizes well to unseen data, providing a more reliable evaluation of its performance.

VI. RESULTS AND DISCUSSION

This section presents the experimental results and performance evaluation of the proposed machine learning

framework for fake news detection. Several classification algorithms were trained and evaluated using a labelled dataset containing both real and fake news articles. The evaluation focuses on comparing model performance, analysing prediction accuracy, and examining the contribution of textual features to the classification process.

The models were evaluated using commonly used classification metrics such as accuracy, precision, recall, and F1-score. These metrics provide a comprehensive understanding of the model's ability to correctly identify fake news articles while minimizing misclassification of genuine news content.

Accuracy Comparison of Machine Learning Models

Several machine learning algorithms were implemented to determine the most effective model for fake news classification. The evaluated models include Logistic Regression, Decision Tree, Support Vector Machine (SVM), Naïve Bayes, and Random Forest. These algorithms are widely used in text classification tasks due to their capability to handle high-dimensional textual data and extract meaningful patterns from large datasets [5], [7].

The performance of each model was evaluated using accuracy, precision, recall, and F1-score metrics.

Table1. Performance Comparison of Machine Learning Models

Model	Accuracy (%)	Precision	Recall	F1-Score
Logistic Regression	91.2	0.90	0.89	0.89
Decision Tree	89.8	0.88	0.87	0.87
Support Vector Machine	93.5	0.92	0.91	0.91
Naïve Bayes	90.6	0.89	0.88	0.88
Random Forest	95.1	0.94	0.93	0.93

From the comparison results, the Random Forest classifier achieved the highest classification accuracy of 95.1%, outperforming other models. This superior performance can be attributed to its ensemble learning structure, which combines multiple decision trees to improve predictive accuracy and reduce overfitting [6], [8].

ROC Curve Analysis

The Receiver Operating Characteristic (ROC) curve is used to evaluate the classification performance of machine learning models by analysing the relationship between the True Positive Rate (TPR) and False Positive Rate (FPR) across different classification thresholds. The Area Under the ROC Curve (ROC–AUC) is commonly used to measure the overall discriminative ability of a classifier. A higher ROC–AUC value indicates better capability of the model to distinguish between fake and real news articles [3], [9].

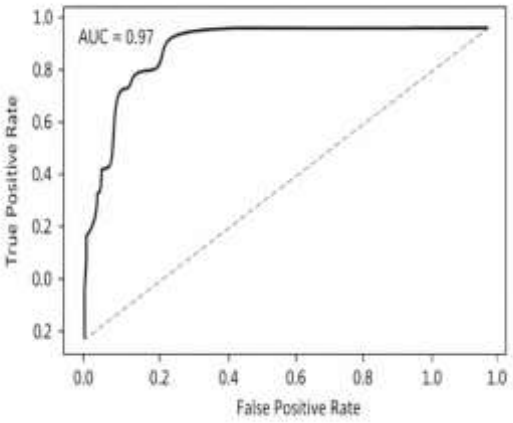

Fig2. ROC Curve for Fake News Detection Model

In the proposed system, the Random Forest classifier achieved a ROC–AUC score of 0.97, which indicates excellent classification performance. A ROC curve positioned closer to the upper-left corner of the graph suggests that the model has strong capability in distinguishing between fake and legitimate news articles.

The ROC analysis demonstrates that the proposed machine learning framework provides reliable predictions even when analysing large textual datasets with diverse linguistic structures.

Feature Importance Analysis

Feature importance analysis was conducted to understand how different textual features contribute to the prediction results. In this study, the Random Forest algorithm was used to identify the most influential features responsible for classifying news articles.

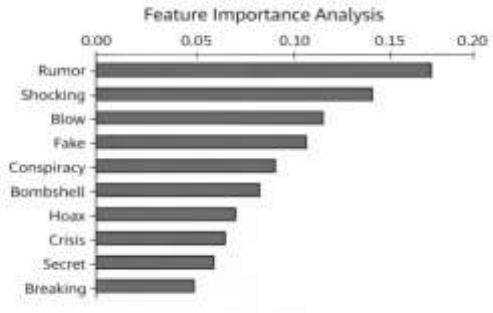

Fig3. Feature Importance Analysis

The analysis revealed that certain words, phrases, and textual patterns play a significant role in identifying fake news content. Terms commonly associated with sensational headlines, exaggerated claims, or misleading narratives tend to contribute more strongly to fake news classification.

The feature importance visualization highlights the relative contribution of different features across the dataset. Such interpretability improves transparency and helps researchers understand the linguistic characteristics associated with misinformation.

This analysis enhances the reliability of the proposed system by allowing users to interpret the reasoning behind classification results, thereby improving trust in machine learning-based fake news detection systems [1], [2], [9], [12].

VII. CONCLUSION AND FUTURE WORK

Detecting and categorizing fake news on social media remains a complex challenge due to its evolving nature and ever-changing patterns. However, advancements in Machine Learning (ML) and Deep Learning (DL) provide effective tools for analysing these dynamic characteristics. This study underscores that leading researchers in the field often leverage sophisticated techniques, including Deep Boltzmann Machines, Deep Neural Networks, Convolutional Neural Networks (CNN), and Deep Autoencoders, to enhance fake news detection.

Looking ahead, future research should explore the integration of metaheuristic optimization techniques to further refine text-based fake news classification. By combining metaheuristic algorithms with existing deep learning models, researchers can improve feature selection, enhance model robustness, and optimize classification accuracy. Additionally, incorporating explainable AI (XAI) approaches could increase transparency and trust in automated detection systems. Expanding studies to include multimodal data—such as images and videos alongside textual content—may further strengthen the detection framework, making it more adaptable to emerging misinformation trends on social media platforms.

REFERENCES

1. M. Choudhary, S. Jha, D. Saxena, and A. K. Singh, "A review of fake news detection methods using machine learning," in 2021 2nd International Conference for Emerging Technology (INCET), 2021: IEEE, pp. 1-5.
2. S. Raza and C. Ding, "Fake news detection based on news content and social contexts: a transformer-based approach," International Journal of Data Science and Analytics, vol. 13, no. 4, pp. 335-362, 2022.
3. L. Waikhom and R. S. Goswami, "Fake news detection using machine learning," in Proceedings of International Conference on Advancements in Computing & Management (ICACM), 2019.
4. A. Abdulrahman and M. Baykara, "Fake news detection using machine learning and deep learning algorithms," 2020 International Conference on Advanced Science and Engineering (ICOASE), pp. 18-23, 2020.
5. J. Shaikh and R. Patil, "Fake news detection using machine learning," in 2020 IEEE International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC), 2020: IEEE, pp. 1-5.
6. [6] S. Gupta, N. Yadav, S. S. Reddy, and S. Kundu, "FakEDAMR: Fake News Detection using Abstract Meaning Representation," Authorea Preprints, 2023.
7. [7] S. I. Manzoor and J. Singla, "Fake news detection using machine learning approaches: A systematic review," in 2019 3rd international conference on trends in electronics and informatics (ICOEI), 2019: IEEE, pp. 230-234.
8. M. K. Singh, J. Ahmed, M. A. Alam, K. K. Raghuvanshi, and S. Kumar, "A comprehensive review on automatic detection of fake news on social media," Multimedia Tools and Applications, pp. 1-34, 2023.
9. S. K. Uppada, K. Manasa, B. Vidhathri, R. Harini, and B. Sivaselvan, "Novel approaches to fake news and fake account detection in OSNs: user social engagement and visual content centric model," Social Network Analysis and Mining, vol. 12, no. 1, p. 52, 2022.
10. E. M. Mahir, S. Akhter, and M. R. Huq, "Detecting fake news using machine learning and deep learning algorithms," in 2019 7th international conference on smart computing & communications (ICSCC), 2019: IEEE, pp. 1-5.
11. J. Y. Khan, M. T. I. Khondaker, S. Afroz, G. Uddin, and A. Iqbal, "A benchmark study of machine learning models for online fake news detection," Machine Learning with Applications, vol. 4, p. 100032, 2021.
12. S. Hakak, M. Alazab, S. Khan, T. R. Gadekallu, P. K. R. Maddikunta, and W. Z. Khan, "An ensemble machine learning approach through effective feature extraction to classify fake news," Future Generation Computer Systems, vol. 117, pp. 47-58, 2021.
13. I. Ahmad, M. Yousaf, S. Yousaf, and M. O. Ahmad, "Fake news detection using machine learning ensemble methods," Complexity, vol. 2020, pp. 1-11, 2020.
14. A. Jain, A. Shakya, H. Khatter, and A. K. Gupta, "A smart system for fake news detection using machine learning," in 2019 International conference on issues and challenges in intelligent computing techniques (ICICT), 2019, vol. 1: IEEE, pp. 1-4.
15. U. Sharma, S. Saran, and S. M. Patil, "Fake news detection using machine learning algorithms,"

International Journal of Creative Research Thoughts (IJCRT), vol. 8, no. 6, pp. 509-518, 2020.

16. S. Ahmed, K. Hinkelmann, and F. Corradini, "Combining machine learning with knowledge engineering to detect fake news in social networks-a survey," arXiv preprint arXiv:2201.08032, 2022.
17. A. Alsharef, Sonia, H. Nassour and J. Sharma, "Exploring the Efficiency of Text-Similarity Measures in Automated Resume Screening for Recruitment,"2023 10th International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 2023, pp. 36-42.
18. R. Sharma and H. K. Meena, "Utilizing graph Fourier transform for automatic Alzheimer's disease detection from EEG signals," International Journal of Information Technology, pp. 1-7, 2024.
19. W. Y. Wang, "" liar, liar pants on fire": A new benchmark dataset for fake news detection," arXiv preprint arXiv:1705.00648, 2017.
20. H. Rashkin, E. Choi, J. Y. Jang, S. Volkova, and Y. Choi, "Truth of varying shades: Analyzing language in fake news and political factchecking," in Proceedings of the 2017 conference on empirical methods in natural language processing, 2017, pp. 2931-2937.
21. S. Singhania, N. Fernandez, and S. Rao, "3han: A deep neural network for fake news detection," in Neural Information Processing: 24th International Conference, ICONIP 2017, Guangzhou, China, November 14-18, 2017, Proceedings, Part II 24, 2017: Springer, pp. 572-581.
22. N. Ruchansky, S. Seo, and Y. Liu, "Csi: A hybrid deep model for fake news detection," in Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, 2017, pp. 797-806.