Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 14, Issue 18s (2025)
https://www.jneonatalsurg.com

OPEN ACCESS

Hindi Fake News Detection Using Machine Learning Models

Varun Rathore[1], Dr. Latika Jinda[l2]

[1]Research Scholar, Department of Computer Science and Engineering, Medi-Caps University, Indore, Madhya Pradesh, India, en23cs501034@medicaps.ac.in
[2]Associate Professor, Department of Computer Science and Engineering, Medi-Caps University, Indore, Madhya Pradesh, India, latika.mehrotra@medicaps.ac.in

Cite this paper as: Varun Rathore, Dr. Latika Jindal, et.al (2025) Hindi Fake News Detection Using Machine Learning Models. *Journal of Neonatal Surgery*, 14 (18s), 299-305.

ABSTRACT

Social media and online platforms amplify the spread of misinformation, directly impacting politics, economics, and society. The results for this study will discuss the application of machine learning and deep learning in detecting fake news in Hindi. The final results for this study were derived from a dataset consisting of over 2,100 Hindi news articles labeled as genuine or fabricated. This paper will systematize the results of three robust algorithms: XGBoost, Convolutional Neural Networks (CNN), and Random Forest (RF). These performances were evaluated against performance by doing preprocessing steps such as tokenization, stop word removal, and stemming, with feature extraction using TF-IDF. After all the comparisons, the best performance was given by XGBoost with 94% accuracy, beating Random Forest with an accuracy of 85% and CNN with an accuracy of 84%. Moreover, XGBoost beat on other metrics too, such as RMSE and MAE. The above findings have underlined the strong potential of both ensemble and deep learning models in the detection of fake news in Hindi.

Keywords: Fake, Machine learning algorithms, Random Forest, CNN, XG Boost.

1. INTRODUCTION

Traditionally, people got information from reputable institutions that had set standards for their work. With the invention of the internet came an explosion of new platforms of information sharing and dissemination, most of which do not follow any rules or standards. This situation has further worsened the situation of credibility determination of news items, since more and more people depend on online sources and social media. Factors such as information overload and a lack of understanding about digital ecosystems further accelerate the spread of fake news and misinformation. Social media, in particular, has become a key driver in amplifying such narratives. "False information" refers to deliberate misrepresentation of facts intended to deceive, while disinformation often emanates from political or economic motives. These stories are spread using bots or fraudsters to accelerate dissemination and employ various tactics to evade detection. Therefore, with the expansion of social media, the spread of misinformation has grown by leaps and bounds. Social media sites like Facebook, Instagram, and WhatsApp are undeniably important channels for the spread of misinformation. This document consists of a survey report on several research publications within this subject, along with a suggested model that forecasts the authenticity of an article using NLP and Machine Learning techniques. This model will receive an article title as input and ascertain if the article is bogus or legitimate. In recent years, the identification of false information has attracted the focus of various researchers globally. This section focusses on contemporary methods used to predict or detect misinformation. Many current disinformation detection techniques mostly depend on feature extraction. The several classifiers included in these solutions have been suitably designated. The advent of the internet and social media has precipitated considerable challenges. Although social media is an efficient medium for the quick and cost-free dissemination of information, it concurrently enables the generation and propagation of misinformation. With the increasing daily number of internet users, who are subjected to a constant stream of information and narratives, the incidence of misleading false news may escalate swiftly. Misinformation often has lasting repercussions and is difficult to rectify. An person will hone their ideas based on their experiences. The exposure may be intentional or unconscious. If the information they encounter is erroneous, they may base their thinking on falsehoods. Furthermore, the fast spread of inaccurate and detrimental information may negatively impact individual cognition and perhaps impair large organisations and the stock market. Individuals may succumb to Truth Bias, Naive Realism, and Confirmation Bias, facilitating the manipulation of customers via misinformation. Truth-bias denotes an individual's inclination to assume the accuracy of information, resulting in the evaluation of news as credible until a specific contextual factor introduces scepticism. Humans inherently struggle to detect deception, exacerbated by their scepticism over the likelihood of being misled. Social media users are unaware that many posts, tweets, and articles are specifically crafted to alter customer perceptions and influence their choices. The manipulation of information is a multifaceted subject that is often difficult to understand, particularly when conveyed by a credible source. Individuals disregard the verification of news authenticity and accept every disinformation as truth. The situation becomes further precarious when adolescents rely heavily on these platforms for knowledge on politics, major events, and breaking news [3-5]. In 2021, 67% of U.S. people accessed news via social media, up from less than 49% in 2012, demonstrating the rapid rise in news consumption via social media platforms. Moreover, people start to see their viewpoints on such news as the only legitimate views, and when others express disagreement, those persons are categorised as "biassed, irrational, or uninformed," a phenomenon also known as a philosophical system [6, 7]. This relates to the notion of confirmation bias. Confirmation bias is the tendency of people to prioritise information that only corroborates their own convictions. Individuals often adhere to material that corroborates their opinions and do not seek data that challenges their notions. An individual may fervently advocate for a political stance, utilising any available online information to substantiate their beliefs, regardless of the sources—be they reputable articles, shared posts from acquaintances, or retweeted content; if it aligns with their convictions, they consider it accurate. Individuals often evade exposure to knowledge that challenges their views, resulting in the acceptance of inaccurate information. They will not let people to prioritise their chosen information, instead promoting disinformation, which leads to confirmation bias. Only persons who attempt to check news to an academic level may dodge prejudice and disinformation; conversely, the normal person, without knowledge of lies, cannot resist these inadvertent impulses. Moreover, ignorance negatively affects people and is harmful to society overall. The spread of disinformation among the public may destabilise the news ecosystem. During the 2016 Presidential election, disinformation was distributed more widely on Facebook than via the most reliable news sites. This example demonstrates how a person may disproportionately concentrate on distorted news rather than factual facts. This problem emerges not only because such disinformation persuades customers to embrace deceptive or distorted narratives to achieve the manipulator's aim and strengthen their influence, but it also effects how consumers react to true news. The principal objective of those spreading false information is to induce confusion, so impairing individuals' ability to distinguish between truth and falsehood. Alongside its effects on people, political motives and manipulation are contributing factors to the proliferation of fake news.

2. PURVEYORS OF MISINFORMATION

While news consumers are mostly authentic persons, the origins of misinformation are not just human. Individuals responsible for disseminating false information generally belong to three classifications: humans, bots, and cyborgs. The little expense of setting up social media accounts enables the construction of deceptive profiles, which may be used to propagate false and detrimental information.

2.1. Social Bots

An account that autonomously participates in conversation on social media is characterised as a bot. Not all bots are intrinsically detrimental; their influence is contingent upon their programming. These are primarily designed to confuse human cognition, allowing manipulators to propagate deceptive material on social media [9, 10]. For instance, “Studies indicate that over 10 million people were in queue to exchange currency notes in India during 2016 demonstrating the significance of these bots to influence the data available.

2.2. Trolls

Although social bots play a substantial role, they are not the exclusive means of propagating false news. Humans significantly contribute to the creation and propagation of erroneous myths. Trolls are real persons who often aim to provoke others worldwide by spreading false information. Evidence suggests that "1000 Russian individuals were engaged to create and disseminate disinformation against Hillary Clinton during the 2016 United States presidential elections," illustrating the intent of specific individuals to manipulate information to sway public opinion [12]. The basic aim of trolls is to cultivate unfavourable opinions among customers, so fostering distrust and scepticism. Upon establishing a viewpoint, consumers serve as channels for the propagation of deceptive and harmful information.

2.3. Cyborgs

A cyborg is a fictional person whose powers are enhanced by the integration of technology. The advantage of using cyborgs is that fraudsters create accounts masquerading as human identities while employing computer programs to operate on social media. The ability to switch functionalities between human and bot amplifies its efficacy, offering a unique edge for doing such tasks [12]. The previous sections demonstrate the impacts and factors associated with false news. The innovation of the present work is in the accurate categorisation and detection of counterfeit material in news, analysed by several artificial intelligence-enhanced machine learning algorithms. The article later used Random Forest, XGBoost, and Convolutional Neural Networks to assess the veracity of the news stream.

3. METHODOLOGY

3.1. Data Set

A significant impediment to the identification of false information is the availability of a high-quality dataset. While several datasets are available for the English language, there are far less datasets accessible for Hindi. A dataset named "Hindi Fake and True Dataset" was created for this investigation. News articles from several Hindi news station websites were analysed to examine both fraudulent and authentic material. This dataset contains more than 2,100 news pieces sourced from several leading news organisations. The dataset consists of two CSV files: fake.csv and true.csv. The dataset statistics are shown in Table 1, along with instances of assertions disseminated online.

Table 1 Counts per data set

Ruling	**Count**	**Sample**
True	95	जम्मू और कश्मीर में टारगेट किलिंग की घटनाओं में वृद्धि हुई है, जिसके कारण पर्यटन में कमी आई है।
False	256	न्यूज़ीलैंड ने भारत के खिलाफ टेस्ट सीरीज़ में उत्कृष्ट प्रदर्शन करते हुए 3-0 से क्लीन स्वीप किया। यह पहली बार है जब न्यूज़ीलैंड ने भारतीय सरज़मीं पर टेस्ट सीरीज़ जीती है। इससे भारत की घरेलू टेस्ट सीरीज़ में 12 वर्षों से चली आ रही अजेयता का सिलसिला टूट गया।
True	63	दिल्ली में प्रदूषण का स्तर खतरनाक, स्कूलों में छुट्टी की घोषणा।
True	67	भारतीय सॉफ्टवेयर कंपनियों ने अमेरिका में रिकॉर्ड निर्यात दर्ज किया।

Table 2 Distribution of news data

Subject	**Percentage**
News	29
Politics	36
Business	12
Sports	5
Others	18

Each shall have four columns: title, description, date of publication, and topic. Table 2 outlines the many categories of news reports. Identifying misinformation is a complex task, since it requires interaction with languages; hence, natural language processing is used for data pretreatment.

3.2. Preprocessing

Raw text must be translated into numerical format prior to training, since machines cannot comprehend language. Preprocessing denotes the alteration or cleansing of data before to its use to improve performance. Numerous pre-processing procedures include data cleansing. Elimination of Stop Words Tokenization Stemming in TF-IDF. The following sections address these steps.

3.3. Data Sanitization

The method involves eliminating null, superfluous, improperly formatted, noisy, or incomplete data from the dataset. Maintaining pristine data will enhance overall productivity.

3.4. Elimination of stop-words

रही, लेकिन, क्या, कि, हम, किया, खुद को, अपने आप को, हैं, पर, स्वयं, हमारे, इसको, आप, बीच, साथ, जब तक, के लिए, मैंने, वह, क्यों, था, आप, तुम, जिससे, किया जा रहा है, को, के बारे में, खुद की, किया है, में, है, जो, जिसके, जब तक, तुम्हारा, उनके, करने का, हम, इसका, लिए, के खिलाफ, खिलाफ, के, हो रहा है, है, के लिए

Fig. 1 Word stops

Stop words are often used terms that lack impact on the ultimate outcome and may be disregarded prior to training. Figure 1 presents Hindi stop words. We have gathered Hindi stop words from Mendeley.

3.5. Stemming

Stemming is the procedure of eliminating suffixes from tokens to transform them into their basic form. For instance, "Playing" is the token that will be transformed into "play" with "ing" being eliminated. The Hindi suffixes are sourced from KCDON4, as seen in Fig. 2. Table 3 presents the dataset subsequent to pre-processing.

"ती", "गई", "ते", "गए", "आओ", "ना", "ताई",
"आए", "उओं", "आई", "तू", "उआ", "उए", "थक",
"गया", "आई", "तेगा", "नी", "नै", "ताई", "आउं",
"देगी", "ता", "तेगी", "आए", "उआं", "आउं",
"उओ", "आऊं", "तेने", "नै", "आई", "आई",
"देगा", "ती", "ताई", "आऊं।"

Fig. 2 Suffixes in stemming

Table 3 Pre-processed data set

Ruling	Sample	Remark
0	दिल्ली में वायु प्रदूषण के स्तर में फिर से वृद्धि, हेल्थ इमरजेंसी की चेतावनी	True
1	संसद में महिला आरक्षण बिल पर बहस, जल्द हो सकता है पारित।	True
2	गुजरात में चुनावी माहौल गर्म, सभी पार्टियां मैदान में।	False
3	अंतरिक्ष में भारतीय वैज्ञानिकों ने एक और सफलता हासिल की।	False
4	नई दिल्ली में अंतरराष्ट्रीय फिल्म महोत्सव का आयोजन।	True

Table 4 Sample response based on Random Forest technique

S. No.	Sample	Prediction	Actual
1	संसद में महिला आरक्षण बिल पर बहस, जल्द हो सकता है पारित।	Not Fake	Fake
2	नई दिल्ली में अंतरराष्ट्रीय फिल्म महोत्सव का आयोजन।	Not Fake	Not Fake
3	गुजरात में चुनावी माहौल गर्म, सभी पार्टियां मैदान में।	Fake	Not Fake

Table 5 Sample response based on XG Boost Technique

S. No.	Sample	Prediction	Actual
1	संसद में महिला आरक्षण बिल पर बहस, जल्द हो सकता है पारित।	Not Fake	Fake
2	नई दिल्ली में अंतरराष्ट्रीय फिल्म महोत्सव का आयोजन।	Fake	Fake
	गुजरात में चुनावी माहौल गर्म, सभी पार्टियां मैदान में।	Not Fake	Fake

Table 6 Sample response based on CNN technique

S. No.	Sample		Prediction	Actual
1	संसद में महिला आरक्षण बिल पर बहस, जल्द हो सकता है पारित।		Fake	Not Fake
2	नई दिल्ली में अंतरराष्ट्रीय फिल्म महोत्सव का आयोजन।		Not Fake	Fake
3	गुजरात में चुनावी माहौल गर्म, सभी पार्टियां मैदान में।		Not Fake	Fake

3.6. Machine learning algorithms

The present work used machine learning classifiers such as Random Forest, XGBoost, and CNN to identify false news material across various news platforms and websites. These classifiers assume that all predictions are independent of one another. It used fundamental probability concepts, assuming that all inputs are independent, to predict the result. RF is the fundamental technique used for constructing classification models. The dataset was acquired from several news sources. It consists of around 2000 articles for training and 927 items for testing [13]. The dataset is divided into three segments: training set, testing set, and validation set. The training dataset is used to instruct the model. The validation dataset is used to adjust the classifier's global parameters. Testing is done on the test dataset to check the accuracy of the model. Through the process of stemming, one is able to get the general meaning of a word. This approach has achieved an accuracy of 74%. This means that even the simplest models can achieve sufficient accuracy when solving such complex classification problems as the problem of fake news detection. On this dataset, XGBoost Classifiers, Random Forests, and Convolution Neural Networks will be used. A variety of news items within a different period-from the year 2023 up until the first week of 2024-and many news inlets helped create the dataset. After going into detail to provide appropriate training after normalization cleaning using these datasets, it had an apex outcome from testing each of these algorithms to about 99.90 percent. The tools used for this purpose are Count Vectorizer and TF-IDF, which are the most important methods of Natural Language Processing in fake news detection. TF-IDF calculates the importance of words in a statement, which helps in effective detection. In the preparation of the dataset, frequently occurring English stop words like "the," "a," and "is" are removed.

It also applied three powerful algorithms: CNNs, Random Forest, and XGBoost, which are each state-of-the-art for specific tasks in regression and classification.

Random Forest: The models developed by this algorithm, by generating many decision trees and averaging their predictions, improve accuracy and reduce overfitting. It does particularly well in cases of missing or imbalanced data in datasets. Its ensemble learning procedure also selects feature importance for better interpretability, handling both continuous and categorical variables effectively.

XGBoost: Extreme Gradient Boosting is a powerful ensemble method that builds trees one after another, each correcting the errors of its predecessor. Because of its efficiency and performance due to features like parallel computing and tree pruning, it also uses regularization techniques-L1 and L2-in order to handle overfitting. This makes it adaptable to both structured and unstructured data.

Convolutional Neural Networks: While these have been designed for image data, they perform well in text classification, too. CNNs hierarchically and locally extract features from the text embeddings using convolutional layers. For handling inputs of variable lengths and having fewer parameters than a fully connected network, it uses pooling layers.

Together, these methodologies constitute a versatile and powerful toolkit to tackle complex classification challenges-like fake news.

4. RESULTS AND DISCUSSION

This work evaluated several deep learning and machine learning methods based on chosen datasets. While exploring, it was observed that the used algorithms were CNN, Random Forest, and XGBoost. We tried to classify between fake and true news stories spreading through communication networks. To elaborate on our results, Tables 4, 5, and 6 depict the predicted versus actual results of the machine learning models.

We have evaluated the performances of three algorithms, namely, Random Forest, XGBoost, and CNN, based on two important performance metrics, Mean Absolute Error and Root Mean Squared Error. Among these models studied, XGBoost was identified to be one of the best, while considering the least MAE obtained was 0.0083 and its RMSE as 0.0046; thus, being better in not giving higher prediction errors. Next, the model of Random Forest, which came next in ranking among these ensemble methods with an RMSE of 0.0052 and MAE of 0.0085.

On the other hand, CNN has its strong points, but since it is highly complex and has to be tuned for hyperparameters, the errors were a little higher: the MAE was 0.0091 and the RMSE 0.0061. However, it is nonetheless a potent tool in our armory. From Figure 3, it can clearly be seen that XGBoost outperforms both Random Forest and CNN for accuracy.

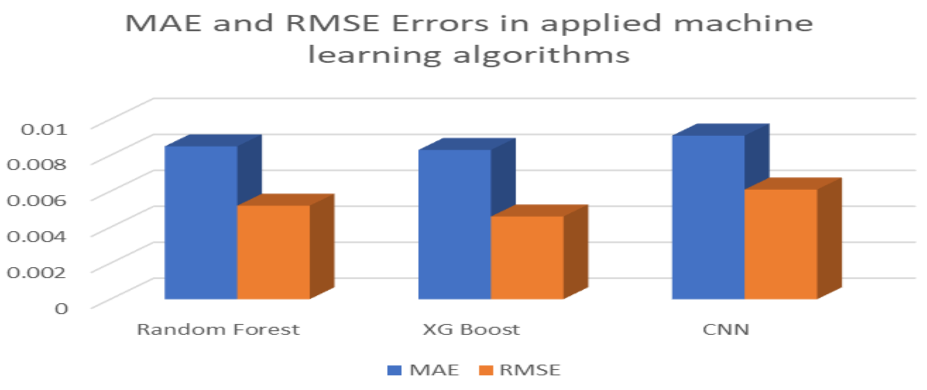

Fig. 3 Errors in machine learning algorithms

Figure 4 shows a bar graph comparing the % accuracy of Random Forest, XG Boost, and CNN, three machine learning algorithms. Thanks to its improved performance, which is probably attributable to its sophisticated boosting algorithms and efficient management of feature interactions, XG Boost obtains the greatest accuracy of the models at 94%. Following closely after with an accuracy of 85%, Random Forest demonstrates its resilience and aptitude as a dependable ensemble approach. Although CNN is successful, it only manages an 84% accuracy rate, suggesting that variables like dataset size or hyperparameter adjustment may have affected its effectiveness. When considering computational efficiency or simplicity as priorities, Random Forest and CNN are good alternatives to XG Boost, which stands out as the best-performing model overall.

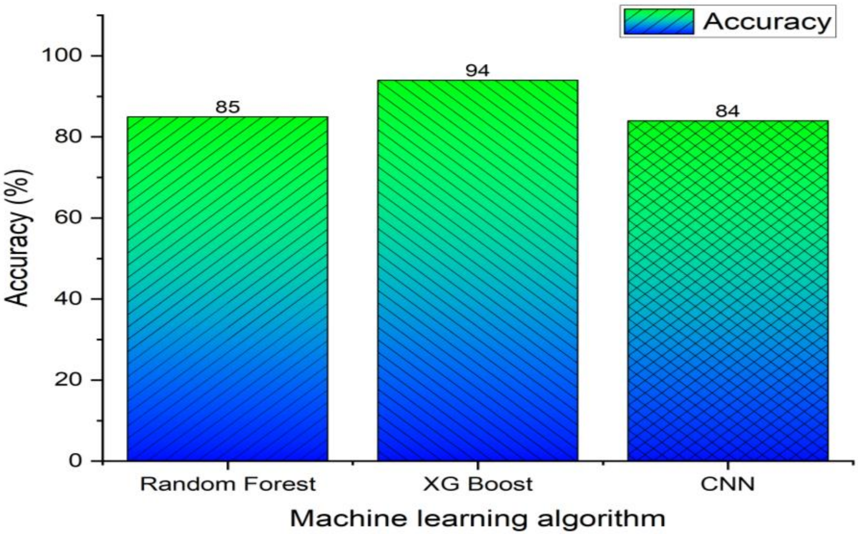

Fig. 4 Accuracy analysis of machine learned models

5. CONCLUSION

The current study has explored the application of deep learning and machine learning techniques, namely CNN, XGBoost, and Random Forest, for the detection of false news stories published in Hindi. Indeed, XGBoost performed really well, giving as high as 94% accuracy with very low error rates: Mean Absolute Error of 0.0083 and Root Mean Square Error of 0.0046, outperforming the other models. Following closely is Random Forest and CNN with an overall performance, which yielded 85% and 84% accuracy, respectively.

The findings in this regard strongly indicate the importance of feature extraction approaches like TF-IDF and, more importantly, rigorous preprocessing techniques in improving arrangement performance. This approach thereby allowed us to take up challenging tasks like distinguishing between false news in Hindi, based on which there is a promise that these models, when applied carefully, can go a long distance toward detecting such misinformation in other regional languages.

This calls for further improvement in the accuracy of detection in the future, considering hybrid models, larger datasets, and contextual factors. This study significantly contributes to efforts against disinformation in regional languages and aids in laying foundations for more trustworthy news ecosystems.

REFERENCE

1. Breiman, L. (2001). Random Forests. Machine Learning, 45(1), 5-32. https://doi.org/10.1023/A:1010933404324
2. Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785-794. https://doi.org/10.1145/2939672.2939785
3. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. Nature, 521(7553), 436-444. https://doi.org/10.1038/nature14539
4. Kim, Y. (2014). Convolutional Neural Networks for Sentence Classification. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 1746-1751. https://doi.org/10.3115/v1/D14-1181
5. Zhou, Z. H., & Feng, J. (2017). Deep Forest: Towards an Alternative to Deep Neural Networks. Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI), 3553-3559. https://doi.org/10.24963/ijcai.2017/497
6. Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake News Detection on Social Media: A Data Mining Perspective. ACM SIGKDD Explorations Newsletter, 19(1), 22-36. https://doi.org/10.1145/3137597.3137600
7. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT), 4171-4186. https://doi.org/10.18653/v1/N19-1423

8. Feng, S., Banerjee, R., & Choi, Y. (2012). Syntactic Stylometry for Deception Detection. Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics, 171-175. https://doi.org/10.18653/v1/P12-2034
9. Pang, B., & Lee, L. (2008). Opinion Mining and Sentiment Analysis. Foundations and Trends in Information Retrieval, 2(1-2), 1-135. https://doi.org/10.1561/1500000011
10. Pennington, J., Socher, R., & Manning, C. D. (2014). GloVe: Global Vectors for Word Representation. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 1532-1543. https://doi.org/10.3115/v1/D14-1162
11. Joachims, T. (1998). Text Categorization with Support Vector Machines: Learning with Many Relevant Features. Proceedings of the 10th European Conference on Machine Learning (ECML), 137-142. https://doi.org/10.1007/BFb0026683
12. Al-Saffar, A., & He, Y. (2018). Fake News Detection Using Machine Learning Algorithms. 2018 IEEE 9th International Conference on Dependable Systems, Services, and Technologies (DESSERT), 109-114. https://doi.org/10.1109/DESSERT.2018.8409109
13. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.
14. Ahmad, I., Yousaf, M., Yousaf, S., & Ahmad, M. O. (2020). Fake News Detection Using Machine Learning Ensemble Methods. Complexity, 2020, Article ID 8885861. https://doi.org/10.1155/2020/8885861
15. Nadeem, M., Horn, C., & Schütze, H. (2020). Fakeddit: A New Multimodal Benchmark Dataset for Fake News Detection. Proceedings of the 12th Language Resources and Evaluation Conference (LREC), 895-903.
16. Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed Representations of Words and Phrases and Their Compositionality. Proceedings of the 26th International Conference on Neural Information Processing Systems (NIPS), 3111-3119.
17. Zhang, Y., & Wallace, B. (2017). A Sensitivity Analysis of (and Practitioners' Guide to) Convolutional Neural Networks for Sentence Classification. Proceedings of the 8th International Joint Conference on Natural Language Processing (IJCNLP), 253-258.
18. Lan, Z., Chen, M., Goodman, S., Gimpel, K., Sharma, P., & Soricut, R. (2020). ALBERT: A Lite BERT for Self-Supervised Learning of Language Representations. Proceedings of the International Conference on Learning Representations (ICLR).
19. Rubin, V. L., Chen, Y., & Conroy, N. J. (2015). Deception Detection for News: Three Types of Fakes. Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community. https://doi.org/10.1002/pra2.2015.145052010083
20. Kaur, P., Kumar, H., & Kumar, M. (2019). A Study of Fake News Detection Using Machine Learning Algorithms. Proceedings of the 4th International Conference on Information Systems and Computer Networks (ISCON), 15-20. https://doi.org/10.1109/ISCON47778.2019.9036305