A Review on Fake News Detection using Machine Learning

Parul Saini
M.Tech Scholar
Department of Computer Science and Engineering,
Bansal Institute of Science and Technology,
Bhopal, MP, India
parulsaini8668@gmail.com

Virendra Khatarkar
Professor
Department of Computer Science and Engineering
Bansal Institute of Science and Technology
Bhopal, MP, India
virendra3012@gmail.com

Abstract: - Fake news, which is defined as material that has been shared with the intention of defrauding people, has been growing quickly and widely recently. This kind of misinformation is dangerous to social cohesion and wellbeing because it exacerbates political polarisation and public mistrust of authority figures. As a result, false news is an issue that has a big impact on our social lives, especially in politics. In order to address this issue, this study suggests brand-new methods based on machine learning (ML) and deep learning (DL) for the fake news identification system. This survey deals with a review of existing machine learning algorithms Naïve Bayes, Convolutional Neural Network, LSTM, Neural Network, Support Vector Machine proposed for detecting and reducing fake news from different social media platforms like Facebook, whatsapp, twitter, etc. This review provides a comprehensive detail including data mining perspective, evaluation metrics, and representative datasheets.

Keywords: Fake News, Machine Learning, Deep Learning, Neural Network, Convolutional Neural Network, Detection.

1. INTRODUCTION

Technology advancements have made knowledge widely available to all people. The internet offers a tonne of information, but its reliability depends on a variety of circumstances. Daily, a huge amount of information is released in print and online media, but it can be difficult to determine if the information is accurate or not. It necessitates a thorough examination and analysis of the narrative, which involves verifying the accuracy of the material by evaluating the sources used to support it, identifying the information's original source, determining the veracity of the authors, etc. The purpose of fabricating information is to intentionally harm or benefit a group, person, or entity's reputation. It can also be done for pure financial or political gain. The phrase "Fake News" was created to describe this type of manufactured information that deceives people. [1] We notice a lot of these falsified messages, news stories, and altered images spreading on social media throughout the Indian election campaigns.

These days, incorrect information is quickly distributed using social media. "A falsehood gets halfway across the world before the truth has a chance to get its pants on," is a famous Wiston Churchill quotation. Because there are so many people using social media, gossip and false information travel quickly. The reaction to this type of news may prove to be the deciding element in whether it is considered to be "fake" or "genuine." To support or refute the claim, the user offers evidence in the form of video or online links. [2] This approach-based classification would be a huge step in the right direction. I ran an experiment to see how often words connected to "fake" appeared in the responses in order to substantiate this claim. The term "deception," which is frequently used to refer to "fake news" in modern times, is defined as false or incorrect information that may be given with the intention of misleading the people who read it. Data or opinions that you can't help but disagree with might not actually be wrong.[3] Although the term "fake news" is frequently used as a derogatory in news reporting nowadays, this is an unreliable use of the phrase. To be sure, labelling reality-based reporting "deception" because it contradicts your political views might perhaps be considered lie in and of itself. Can we trust the majority of the news we read on social media and on ostensibly "legitimate" news websites? It is incredibly simple for anyone to post whatever they want, and while that may be acceptable, there is the idea of going too far. For example, posting false information online to incite panic, using lies to influence another person's decision, or essentially anything else that can have long-lasting effects, are all examples of going too far. [4]There is so much information available online that separating fact from fiction is becoming impossible. Consequently, this results in the issue of fake news.

2. LITERATURE REVIEW

In order to identify bogus news, Toumi et al. describe an ensemble learning model that makes use of CNN, LSTM, and C-LSTM. The ISOT and LIAR datasets were used in conjunction with one another to create a single corpus for this purpose. The proposed model produced good results when measured against the state-of-the-art. The proposed model produced competitive outcomes when measured against the state-of-the-art.

An ImageFake ensemble model, developed by Choudhary et al. VGG-16, VGG-19, Inception v3, SqueezeNet, and ResNet-101 are the pre-trained models utilised for false news detection and classification, and the bagging ensemble model is used to pick the best of the lot. The

ensemble model is able to get training and validation accuracy 97% and 66% respectively.

A fresh WELFake data set with over 72 000 articles was created by Verma et al.[10] that combines various data sets to produce an objective classification result. According to experimental findings, the WELFake model classifies news as real or false with a 96.73% accuracy rate, which is better than convolutional neural network (CNN) models and bidirectional encoder representations from transformer (BERT) models by 1.31% and 4.25%, respectively. Our frequency-based and targeted analysis of writing patterns model performs up to 1.73% better than comparable predictive-based works conducted using the Word2vec WE approach.

Xi and co. .[12] Create a COVID-19 fake news detector by combining three methods: CNN, LSTM models, and attention-based transformer. Evaluating the false news detector is one of the model performance criteria along with accuracy, F1 score, AUC score, and training time. The attention-based transformer model has demonstrated its benefits in light of the data thanks to its high accuracy of 75.9% and AUC score of 0.774.

The Nave Bayes Classifier, which was tested in a software system with a data set of Facebook news posts and had achieved an accuracy of 74%, the Convolutional Neural Network (CNN), for image visualisation, scored a mean accuracy of 92.85%, and the Recurrent Neural Network (RNN), for audio and text visualisation, had achieved an accuracy of 93% on the datasets collected from Brussels terrorist attacks. Patel et al. review various machine learning-based approaches. Some of the hybrid models are the Meta Optimization Semantic Evolutionary Search model (MOSES), which achieved a mean accuracy of 63%, and the early identification of false news on Twitter utilising propagation route classification using mixed CNN and RNN model, which had an accuracy of 85%.

Using a pre-train language model PhoBERT along with Term Frequency - Inverse Document Frequency (TF-IDF) for word embedding and Convolutional Neural Network (CNN) for features extraction, Pham et al. [14] propose a new approach for Vietnamese Fake News Detection on Social Network Sites. On the dataset of Reliable Intelligence Identification on Vietnamese SNSs (ReINTEL) shared task, our suggested model is trained and evaluated. To clarify the pre-processing hypothesis for data on social networks, we divide text data into two scenarios: raw data and processed data. Additionally, we apply the various added characteristics to raise the model's effectiveness. We contrast the baseline approaches with our suggested model. The proposed model produced remarkable outcomes with an AUC score of 0.9538 for the raw data.

Deep learning algorithms are used by Mansoui et al.[15] to identify bogus news. The approach uses a convolutional neural network to target both labelled and unlabeled input in a semi-supervised learning environment. In order to increase the impact of the estimated class in each step, the fitness function has also been altered. Results demonstrate that the proposed method, with a precision value of 95.5%, beats existing methods in terms of recall, specificity, and sensitivity.

A multi-CNN technique was suggested by Babu et al. [17] for identifying deepfakes. The CNN utilised in this case is a multipath CNN with three modules, each of which is stacked with a convolutional block attention mechanism. A Resnet module and a Densenet module are the first two modules in the dual-path paradigm. A one-dimensional feature reduction module with residual connections can be found in the parallel Inception Resnet module. With an accuracy and F1-score of 0.940 and 0.939, respectively, the proposed method produced the best results when its performance is compared to that of four deep learning-based approaches.

By taking into account the contextual information, Rani et al.[18] look at the rumour identification challenge The proposed model outperformed the baseline model, according to experimental findings. The accuracy of the proposed model is 90.93%.

Islam et al.'s[22] categorisation of Bengali false news takes into account the South Asian context. More over 200 million people worldwide speak Bengali, which is also their primary form of communicationWe have also developed a web interface based on our classifier to determine the validity of news articles in Bengali. When using a random forest classifier, the classification model achieves an accuracy of 85%.

To recognise bogus tweets, Kar et alteam[25] .'s suggests a BERT-based model enhanced with additional pertinent information taken from Twitter. We demonstrate through thorough trials that our approach outperforms state-of-the-art (SOTA) findings in false tweet identification, scoring around 89% F-Score. Additionally, we create the first benchmark for Bengali and Hindi, two Indic languages. Our model scores a 79% F-Score for Hindi and an 81% F-Score for Bengali Tweets using our annotated data. About 81% F-Score is achieved by our zero shot model in.

For the purpose of evaluating the reliability of web information contents, Meel et al.[26] developed a system based on five independent supervised machine-learned classifiers: Support Vector Machine, K-Nearest Neighbor, Logistic Regression, Naive Bayes, and Random Forest. The greatest accuracy reached by the projected technique on the Fake News Detection dataset using the SVM classifier is 96.61%, according to experiments and result analyses. The piece is also contrasted with other modern methods. Below table 1 presents the research contribution of researchers.

Table 1. Recent Research Contributions

Ref	Technique used	Dataset used	Results
[1]	CNN ,LSTM	LIAR	Accuracy 89.16 F1-score 95.03
[7]	CNN,VGG-16,VGG-19	MediaEval 2015	Accuracy training 96 Accuracy validation 97
[10]	CNN	WELFake	Accuracy 96.73
[12]	CNN	Kaggle	Accuracy 75.9 AUC score 0.774
[14]	CNN	ReINTEL	AUC 0.9538

[15]	LDA ,CNN	Kaggle	Accuracy 95.5
[17]	CNN,RESNET	Kaggle	Accuracy 0.940
[18]	CNN and BILSTM	Kaggle	F1 score 0.939 accuracy 90.93
[22]	Random Forest	South Asian context	Accuracy 85
[25]	BERT model	Kaggle	F1 score 81
[26]	SVM	Fake news detection	Accuracy 96.61

3. CONCLUSION AND FUTURE WORK

In this work, a framework is designed for news reality identification. For this user profile and their activities are observed. With their activities, and profile information correlation we are going to assign a reality score to each and every user. For this analysis statistical feature engineering and deep learning is integrated together in common framework to enhance the performance. The training datasets have been used to optimize and train the models, while the testing datasets were used to evaluate the models. It is important that we have some mechanism for detecting fake news, or at the very least, an awareness that not everything we read on social media may be true, so we always need to be thinking critically. This way we can help people make more informed decisions and they will not be fooled into thinking what others want to manipulate them into believing.

Conflict of Interest:The corresponding author, on behalf of all authors, confirms that there are no conflicts of interest to disclose.

REFERENCES

[1] C. Toumi and A. Bouramoul, "Ensemble learning-based model for fake news detection," *4th Int. Conf. Pattern Anal. Intell. Syst. PAIS 2022 - Proc.*, 2022, doi: 10.1109/PAIS56586.2022.9946895.

[2] J. Alghamdi, Y. Lin, and S. Luo, "Modeling Fake News Detection Using BERT-CNN-BiLSTM Architecture," *Proc. - 5th Int. Conf. Multimed. Inf. Process. Retrieval, MIPR 2022*, pp. 354–357, 2022, doi: 10.1109/MIPR54900.2022.00069.

[3] A. Gupta, V. Mishra, and B. J, "Fake news detection using Ensemble model," in *2022 Second International Conference on Interdisciplinary Cyber Physical Systems (ICPS)*, 2022, pp. 34–39. doi: 10.1109/ICPS55917.2022.00014.

[4] Q. Abbas, M. U. Zeshan, and M. Asif, "A CNN-RNN Based Fake News Detection Model Using Deep Learning," in *2022 International Seminar on Computer Science and Engineering Technology (SCSET)*, 2022, pp. 40–45. doi: 10.1109/SCSET55041.2022.00019.

[5] H. Ali *et al.*, "All Your Fake Detector are Belong to Us: Evaluating Adversarial Robustness of Fake-News Detectors Under Black-Box Settings," *IEEE Access*, vol. 9, pp. 81678–81692, 2021, doi: 10.1109/ACCESS.2021.3085875.

[6] X. Zhi *et al.*, "Financial Fake News Detection with Multi fact CNN-LSTM Model," in *2021 IEEE 4th International Conference on Electronics Technology (ICET)*, 2021, pp. 1338–1341. doi: 10.1109/ICET51757.2021.9450924.

[7] A. Choudhary and A. Arora, "ImageFake: An Ensemble Convolution Models Driven Approach for Image Based Fake News Detection," in *2021 7th International Conference on Signal Processing and Communication (ICSC)*, 2021, pp. 182–187. doi: 10.1109/ICSC53193.2021.9673192.

[8] X. Lyu, "News and Social Media Automatic Detection System based on Big Data Automatic Identification Technology," in *2022 IEEE 2nd International Conference on Electronic Technology, Communication and Information (ICETCI)*, 2022, pp. 1120–1123. doi: 10.1109/ICETCI55101.2022.9832122.

[9] K. Ivancová, M. Sarnovský, and V. Maslej-Krcšňáková, "Fake news detection in Slovak language using deep learning techniques," in *2021 IEEE 19th World Symposium on Applied Machine Intelligence and Informatics (SAMI)*, 2021, pp. 255–260. doi: 10.1109/SAMI50585.2021.9378650.

[10] P. K. Verma, P. Agrawal, I. Amorim, and R. Prodan, "WELFake: Word Embedding Over Linguistic Features for Fake News Detection," *IEEE Trans. Comput. Soc. Syst.*, vol. 8, no. 4, pp. 881–893, 2021, doi: 10.1109/TCSS.2021.3068519.

[11] F. Torgheh, M. R. Keyvanpour, B. Masoumi, and S. V. Shojaedini, "A Novel Method for Detecting Fake news: Deep Learning Based on Propagation Path Concept," in *2021 26th International Computer Conference, Computer Society of Iran (CSICC)*, 2021, pp. 1–5. doi: 10.1109/CSICC52343.2021.9420601.

[12] J. Xi and C. Zhang, "COVID-19 fake news detection using attention-based transformer," in *ICMLCA 2021; 2nd International Conference on Machine Learning and Computer Application*, 2021, pp. 1–7.

[13] K. Patel, Swati, and I. U. Hassan, "Identifying Fake News with Various Machine Learning Model," in *2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, 2021, pp. 1–5. doi: 10.1109/ICRITO51393.2021.9596330.

[14] N.-D. Pham, T.-H. Le, T.-D. Do, T.-T. Vuong, T.-H. Vuong, and Q.-T. Ha, "Vietnamese Fake News Detection Based on Hybrid Transfer Learning Model and TF-IDF," in *2021 13th International Conference on Knowledge and Systems Engineering (KSE)*, 2021, pp. 1–6. doi: 10.1109/KSE53942.2021.9648676.

[15] A. Kishwar and A. Zafar, "Predicting Fake News using GloVe and BERT Embeddings," in *2021 6th*

South-East Europe Design Automation, Computer Engineering, Computer Networks and Social Media Conference (SEEDA-CECNSM), 2021, pp. 1–6. doi: 10.1109/SEEDA-CECNSM53056.2021.9566243.

[16] R. Mansouri, M. Naderan-Tahan, and M. J. Rashti, "A Semi-supervised Learning Method for Fake News Detection in Social Media," in *2020 28th Iranian Conference on Electrical Engineering (ICEE)*, 2020, pp. 1–5. doi: 10.1109/ICEE50131.2020.9261053.

[17] R. B. P. and M. S. Nair, "Deepfake Detection using Multi-path CNN and Convolutional Attention Mechanism," in *2022 IEEE 2nd Mysore Sub Section International Conference (MysuruCon)*, 2022, pp. 1–6. doi: 10.1109/MysuruCon55714.2022.9972657.

[18] N. Rani, P. Das, and A. K. Bhardwaj, "A hybrid deep learning model based on CNN-BiLSTM for rumor detection," in *2021 6th International Conference on Communication and Electronics Systems (ICCES)*, 2021, pp. 1423–1427. doi: 10.1109/ICCES51350.2021.9489214.

[19] J. Gao, X. Li, and Z. Huang, "A Discriminative Deep Neural Network for Text Classification," in *2021 International Conference on Computer Information Science and Artificial Intelligence (CISAI)*, 2021, pp. 731–735. doi: 10.1109/CISAI54367.2021.00148.

[20] Berthet and J.-L. Dugelay, "A review of data preprocessing modules in digital image forensics methods using deep learning," in *2020 IEEE International Conference on Visual Communications and Image Processing (VCIP)*, 2020, pp. 281–284. doi: 10.1109/VCIP49819.2020.9301880.

[21] Ganesh and D. K. Anitha, "Implementation of Personality Detection and Accuracy Prediction for identification of fake and true news using Decision Tree and Random Forest Algorithms," in *2022 International Conference on Business Analytics for Technology and Security (ICBATS)*, 2022, pp. 1–5. doi: 10.1109/ICBATS54253.2022.9759039.

[22] F. Islam *et al.*, "Bengali Fake News Detection," in *2020 IEEE 10th International Conference on Intelligent Systems (IS)*, 2020, pp. 281–287. doi: 10.1109/IS48319.2020.9199931.

[23] Anjum, M. Keya, A. K. Mohammad Masum, and S. R. Haider Noori, "Fake and Authentic News Detection Using Social Data Strivings," in *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 2021, pp. 1–5. doi: 10.1109/ICCCNT51525.2021.9579777.

[24] M. Dong, L. Yao, X. Wang, B. Benatallah, X. Zhang, and Q. Z. Sheng, "Dual-stream Self-Attentive Random Forest for False Information Detection," in *2019 International Joint Conference on Neural Networks (IJCNN)*, 2019, pp. 1–8. doi: 10.1109/IJCNN.2019.8851765.

[25] D. Kar, M. Bhardwaj, S. Samanta, and A. P. Azad, "No Rumours Please! A Multi-Indic-Lingual Approach for COVID Fake-Tweet Detection," in *2021 Grace Hopper Celebration India (GHCI)*, 2021, pp. 1–5. doi: 10.1109/GHCI50508.2021.9514012.

[26] P. Meel and D. K. Vishwakarma, "Machine Learned Classifiers for Trustworthiness Assessment of Web Information Contents," in *2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)*, 2021, pp. 29–35. doi: 10.1109/ICCCIS51004.2021.9397228.