JOURNAL OF EMERGING TECHNOLOGY AND DIGITAL TRANSFORMATION
ONLINE ISSN
3006-9726
PRINT ISSN
3006-9718
VOLUME . 4 ISSUE . 2 (2025)

IDENTIFICATION OF FAKE NEWS ON SOCIAL MEDIA USING TEXT MINING

Areeba Razzaq
Department of Computer Science, University of Southern Punjab Multan
areebarazzaq97@gmail.com
Muhammad Sabir
Department of Computer Science, University of Southern Punjab Multan
muhammadsabir@usp.edu.pk
Mubasher H Malik
Department of Computer Science, University of Southern Punjab Multan
mubasher@usp.edu.pk
Kiran Shahzadi
Department of Computer Science, University of Southern Punjab Multan
kiranqasim859@gmail.com
Huma Asif
Department of Computer Science, University of Southern Punjab Multan
aries111.123@gmail.com

RECEIVED	ACCEPTED	PUBLISHED
2 july	23 july 2025	18 Aug 2025

ABSTRACT
The fake news detection is one of the most pressing problems in online social media and it has major social and political implications. Several automated techniques for the filtering out fake news have been suggested earlier. This research focuses on the issue of detecting fake news on social media employing text mining approach. We propose an ensemble approach based on hard voting, combining the strengths of three machine learning models. So there are Logistic Regression, Random Forest, Decision Tree. Social media text data is cleaned and normalized and turned into a structure that can be effectively analyzed. The ensemble method combines the decision that is produced by three models into one decision. By using the experimental results, it can be concluded that the proposed approach highly effective with an accuracy rate of 89% that helps in differencing the real and fake news. It is possible to identify fake news using machine learning through this approach, which can work as a perfect solution to the problem.
Index Terms Ensemble Learning, Hard Voting, Logistic Regression, Random Forest, Decision Tree.

JOURNAL OF EMERGING TECHNOLOGY AND DIGITAL TRANSFORMATION
ONLINE ISSN
3006-9726
PRINT ISSN
3006-9718
VOLUME . 4 ISSUE . 2 (2025)

Introduction

We have observed that with time the world is getting smaller and is increasingly becoming a global village. The technology that has surfaced in the last few decades has revolutionized and made communication among distant living people facile(Xu et al., 2025). Internet media is endless and has turned all the domains into the virtual world for exchanging information at a faster rate. According to the beginning of 2023, nearly 87.35 million users of internet were indicated leading to 36.7% penetration(Moin Khan et al., 2020).
The quantity of daily processing of data on the online media has been heightened to an infinite number(Matti & Yousif, 2025). Currently, many people express their thoughts and report events in social media interactions as posts, articles, tweets, retweets, blogs and stories; there is also a proliferation of geometries of abundance (Taher et al., 2022). Some of the problems that have been occasioned by circulation such information include: spread of fake news.
As a result of the changes that were being made in the technology(Steinebach et al., 2020), the usage of the internet was rising up to 4.95 billion which put the world social media active population at 61% of the overall (Saeed & Al Solami, 2023), world population and this ratio is expected to grow by 5%. As estimated, it will be 17 billion at the end of the year in the year 2024.
Popular social sites such as Face book, twitter and Instagram are the first ones in which users are interested in sharing their opinions, photos, audio and video clips in cross sectional communication (Essa et al., 2023). Such platforms where assists as positive media that boosts the users to increase their social activity and perspectives also have some significant demerits. Misinformation is as a result quantified as the 'core topic' of discussion on social media platforms(Ibrahim et al., 2023).
This means that social media users are autonomous because they can categorize any type of information, image, audio and even videos. Social media is like woods fire which has the ability to spread right and wrong information at anytime and anywhere(Farhoudinia et al., 2025).
This has continued to be a massive problem, especially in the current generation where even news is passed through the social media with the speed of the blink of an eye (Mookdarsanit & Mookdarsanit, 2021). Post-truth is a recent term to encompass rumors or 'lies' that have been turned into a convincing news story so that it is able to influence people's thinking and actions.
The detection of fake news has now become essential because fake news capable of altering the society's demographic distribution(Monti et al., 2019), disrupting political proceedings and diminishing the community's confidence in media outlets and other social institutions(Zellers et al., 2019).
Fake news (Benamira et al., 2019) are defined as news that are detected with the help of text mining approaches. Text mining techniques are also adopted for identifying correlation between the dataset to detecting fake news. Therefore, this research aims to establish and evaluate related models that enable us to compare state-of-art Machine Learning techniques for fake news detection(Lakzaei et al., 2025). Besides contributing a major effort to fight one of the largest present day issues for the benefit of the society, we expand text mining and machine learning research involved.

Related Works

Researchers (Aïmeur et al., 2023) uses a lot of models that use supervised learning

ONLINE ISSN
3006-9726
PRINT ISSN
3006-9718
VOLUME . 4 ISSUE . 2 (2025)

techniques include; regression, KNN, Random forest, decision trees and many more in which the labeling procedure is not very well understood during training. Below are the subsections of supervised classification based detection models such as CNN, SVM as well as RNN.

The paper under discussion (Reddy et al., 2020) provides the understanding of the CNN-based approach to distinguish real and fake information in videos with the emphasis on the usage of Deep Learning and Feature-Based methodologies. Authors(Sharma et al., 2021) rewarded for Fake News and Reviews, Deep Fake Detecting Tools. The best performing learning algorithms were the conventional and deep learning ones for the identification of fake news preferred SVM, KNN and decision trees.

For this purpose, 44 videos connected to the topic of health films available on the Chinese Bilibili site were obtained(Englmeier, 2021). Thus, the sample of the movies was composed of authentic movies and fakes. By using the linear methods, it is possible to identify 490 authentic and 210 fraudulent out of 700 transactions(Suryavardan et al., 2022). To train this feature set as seen in two dimensions, CNN was incorporated and included into the system.

In this investigation(Wan et al., 2021), the model that was developed used precision at a level of 1 percent and accuracy of 90 percent. Thus, the authors felt that it was necessary to set a stricter approach to develop this model by limiting the focus of the study to the articles related to You Tube videos.

Internet news is also projected to be more popular in the future due to issues like, Relative ease of access to spectators and Relative flexibility of options(Kaddoura et al., 2022). This is true because news can be spread over the internet; therefore, there has been cases of fake news. In this work (Rahman et al., 2023), we apply the ML and text mining techniques to the fake news detection problem.

In this work the chosen method for feature extraction was TF-IDF and given text has been preprocessed by removing words present in them. Kaggle provided for the data set used in this analysis while the political content is inherent to the authors(Forest, 2022). Feature validation involved SVM and Naïve Bayes and logistic regression in addition to the probability of occurrence of every class was weighed. Last algorithm used in the analysis was Decision Tree which achieved the highest accuracy of 99. 57% to the ML model that analyzed the fake and the genuine news.

The authors' studies of the ML models such as Random Forest and Decision tree approach pointed out that feature selection from large amounts of text were done using TF-IDF and Word2Vec to a reasonable extent(Mallick et al., 2023). Due to over-sampling of the dataset the data became balanced. The analysis of this method's performance we have used mean of accuracy, F1 Score, Precision and Recall as well as the AUC curve deems this approach to be the most accurate and best-performing model with of 94.24% of accuracy.

The author of (Taher et al., 2022) presents interesting ideas in increasing the odds of detecting simulated reports. Pre-processing was also done in targets and features were extracted through BERT and Glove in the course of the study. Similarly, with deep learning methods, other machine learning methods like SVM, Logistic Regression and Naïve Bayes were used likewise, where through the Gossip cop dataset (10670 headlines) Kaggle dataset (6328 articles) and McIntyre dataset (19390 articles); content-based data were recognized with 93.33% accuracy.

JOURNAL OF EMERGING TECHNOLOGY AND DIGITAL TRANSFORMATION
ONLINE ISSN
3006-9726
PRINT ISSN
3006-9718
VOLUME . 4 ISSUE . 2 (2025)

Table 1 Related papers along references and results

References	Dataset	Method	Results
(Xu et al., 2025)	WELFake	Large Language Model(LLM)	F1 score was 1.5% better than existing models
(Mookdarsanit & Mookdarsanit, 2021)	Global open Covid 19	BERT, ULMFit	72% with ULMFit
(Ahmad et al., 2020)	ISOT Fake News	SVM , XGBoost	92.5%
(Murti et al., 2025)	FakeNewsDetection Dataset	K-NN	Mean Squared Error (RMSE) of 0.077, Mean Absolute Error (MAE) of 0.011
(Zhi, 2023)	Self-built dataset from public opinion	MDFM-MF , BERT	90.44%
(Jwa et al., 2019)	Daily Mail dataset	BERT, exBake	74%
(Matti & Yousif, 2025)	Arabic Dataset	LSTM, Glove, BERT	BERT achieved 0.98 better accuracy
(Altheneyan & Alhadlaq, 2023)	FNC-1	TF-IDF , Logistic Regression	92.45%
(Biradar et al., 2023)	Twitter	XLNet , ELMo , BERT	97%
(Sahoo & Gupta, 2021)	Raw dataset generated from user posts	SVM, LSTM,KNN	85%
(Wang et al., 2022)	Weibo dataset	FMFN , Roberta	88.5%

Proposed Model

In this study, we have proposed an ensemble model comprising of three algorithms as logistic regression, random forest and decision tree for the detection of fake news from social media platforms that accurately identifies real and fake news. Fig 1 shows the overall working of the proposed model.

ONLINE ISSN
3006-9726
PRINT ISSN
3006-9718
VOLUME . 4 ISSUE . 2 (2025)

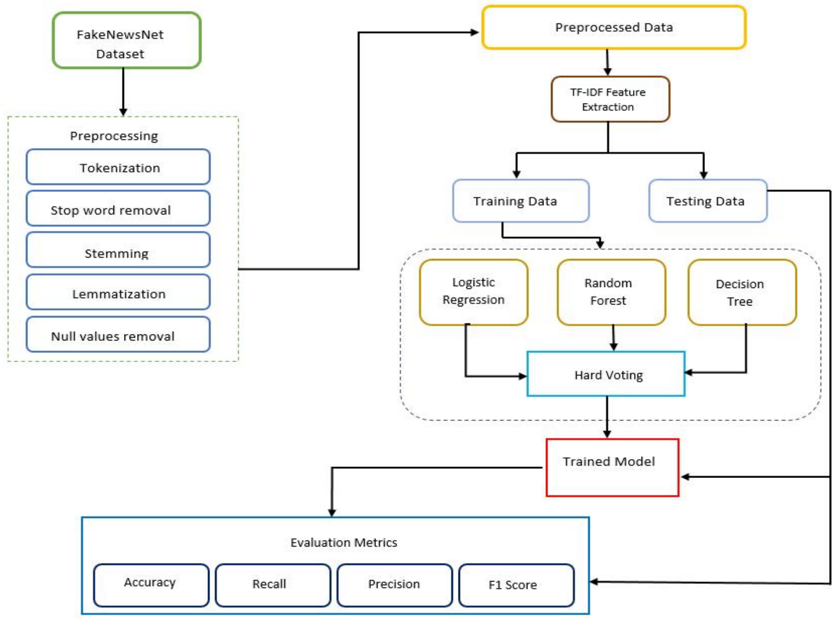

Figure 1 Proposed Methodology

Data Preprocessing

The proposed methods are train and tested on a dataset named "FakeNewsNet" that is taken from kaggle source. The dataset used in the research is not completely clean and includes some additional information which may impact on the performance of the proposed models. To preprocess the dataset, NLTK is used. To remove those entries, in the preprocessing section tokenization, lemmatization, stop words and null values were removed from the dataset.

Feature Extraction

Although, higher variability is given by the TF-IDF Vectors which is not advanced further based on the word counting. The term "term frequency" (TF) and the "inverse document frequency" (IDF) are the two components of TF-IDF Vectors, which are determined by the following formulas: When it comes to calculate the TF-IDF Vector two values which is defined is the (TF) and (IDF).

$$TF(t,d) = \frac{\textit{Number of times term t appears in document d}}{\textit{Tottal number of terms in document d}} \quad (1)$$

The specification of the matrix representation in different levels.

Words level – means that the scores are TF-IDF of terms.

N-gram level - it is the vectors of TF-IDF scores of the N-Grams.

Character level – We described it as the number of relevance scores of n-grams, which was calculated with the help of the TF-IDF method which takes into consideration the character level of the words.

Ensemble Learning Methods

In ensemble learning approaches(Villela et al., 2023) there are several machine learning algorithms used and their outputs combined to form the final output of a single learning algorithm. The rationale of ensemble learning is always to combine several individually distinct 'weak learners' into one

ONLINE ISSN
3006-9726
PRINT ISSN
3006-9718
VOLUME . 4 ISSUE . 2 (2025)

'strong learner'. It will also be important to note that; usually, the goals of ensemble methodologies are to decrease bias, add variance, or enhance the predictive capability.

For improvement of results, the voting method used in the ensembling involves combination of results from several classifiers' predictions. Using Decision Trees (DT), Random Forests (RF) and Logistic Regressions (LR) voting operates as follows:

LR: This linear model known as LR gives probabilities for binary classification and is considered best when two classes, one for fake and other for real was required after estimating the relationship of the features and the result. The member type of a given class is returned by it in case that a certain input is contained in the specified class.

$$P(y = 1|x) = \frac{1}{1+e^{-(\omega tx+b)}} \quad (2)$$

RF: The good thing is that if a random portion of the data and features is selected then each and every decision tree present in the RF is built. It reduces variance in an attempt to strongly predict by providing the over-riding majority class which each individual tree had considered.

$$y = \frac{1}{N}\sum_{i=1}^{N} Ti(x) \quad (3)$$

DT: In comparison with the other types of learning DTs classify data according to the feature based decision rule learning. This is a structure in which every decision has been made with regards to a feature, and the resultant final class label corresponds to the node.

$$IG(S,A) = Entropy(S) - \sum_{v\epsilon A}\frac{|Sv|}{|S|}.Entropy(Sv) \quad (4)$$

Hard Voting: The prediction made by each of the models (LR, RF, DT) produces a class with a hard voting. Here the class that receives most of the vote is taken as the predicted output class for the given data set. For example, while the classifiers namely LR, RF and DT classify an input text respectively as 'Fake' and 'Real', the final prediction given will be 'Fake' since majority of classifiers for this particular case predictable 'Fake'.

Result and Discussion

In this research work new ensemble model is presented, which includes LR, RF and DT model and it is tested on FakeNewsNet dataset. As observed and because of spread of fake news, social and security challenges are experienced which have to addressed.

The code splits data into testing and training data set of features *x* and the labels data set *y*. In relation to this, it divides the data into the training and testing subsets of which the testing is set at 20% (*test_size=0. 20*). The *random_state=10* increases the possibility of getting a specific split by setting the random number generator at a specific point and in this case is set to 10.

The parameter that is used in the stratify=y is helpful in the sense that it ensures that both in the training as well as in the testing stage, the proportion of the classes in the new data set is proportionate to the proportion of the initial data set which maintains the balance of the classes.

JOURNAL OF EMERGING TECHNOLOGY AND DIGITAL TRANSFORMATION
ONLINE ISSN
3006-9726
PRINT ISSN
3006-9718
VOLUME . 4 ISSUE . 2 (2025)

Figure 2 Bar graph representation of training and testing dataset

Then the given below heat map of vectorization conducted by TF-IDF in which terms indicates the rows and documents indicates the columns. The heat map helps in visualizing the values of higher TF-IDF scores with darker colors.

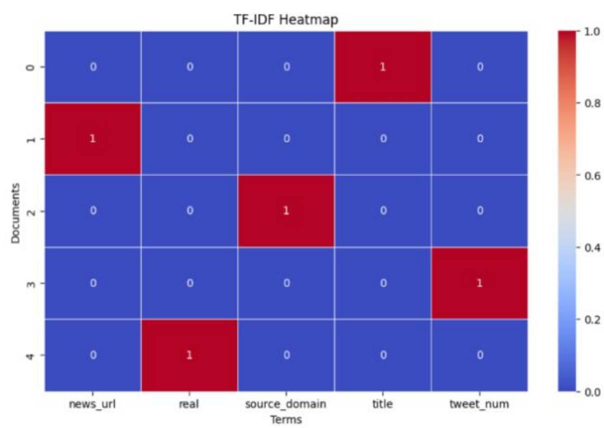

Figure 3 Heat map showing results of vectorization(TD-IDF) on dataset

Given below are the individual confusion matrix that shows the highest number of positive values in each class of all three applied models (LR, RF, DT) individually. The confusion matrix for hard voting of the model shows the best values diagonally.

JOURNAL OF EMERGING TECHNOLOGY AND DIGITAL TRANSFORMATION
ONLINE ISSN
3006-9726
PRINT ISSN
3006-9718
VOLUME . 4 ISSUE . 2 (2025)

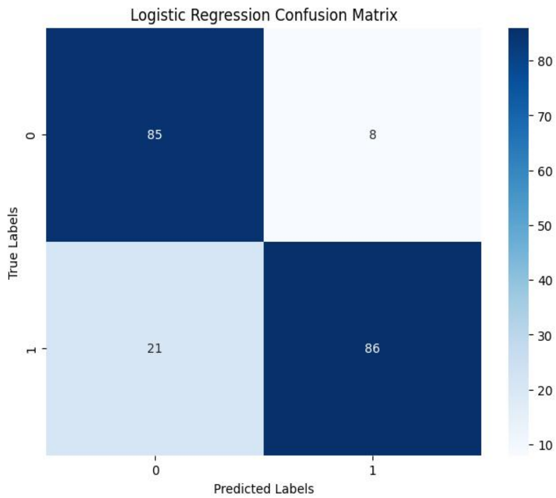

Figure 4 Confusion matrix of Logistic Regression

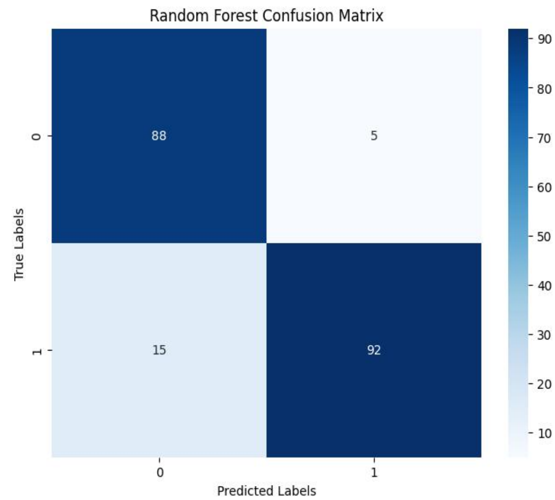

Figure 5 Confusion matrix of Random Forest

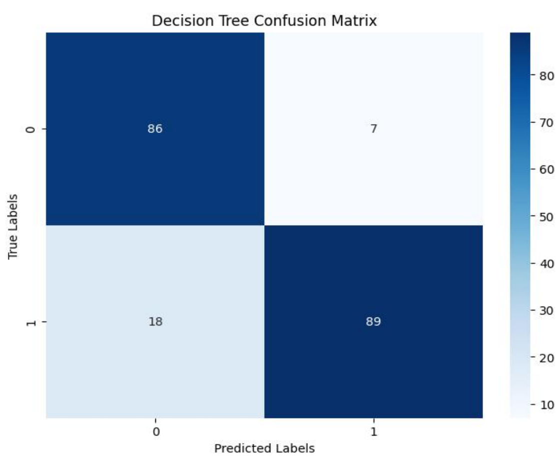

Figure 6 Confusion matrix of Decision Tree

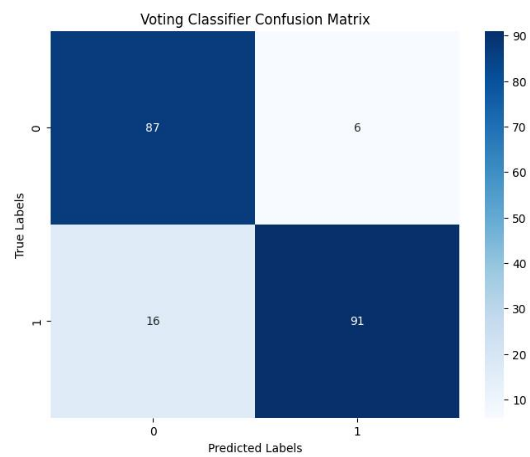

Figure 7 Confusion matrix of Voting on Three models (LR, RF & DT)

The number of true positives values obtained through voting classifier are greater and better in numbers. From these results, our suggested model performed admirably on the dataset of fake news. This study revealed that the selected ensemble model serves good for the FakeNewsNet dataset. The accuracy of all ensemble model used work was found to be highest at about 89%.

The bar graph shown here reveals the comparison of model accuracies of Logistic Regression, Random Forest, Decision Tree and our proposed Hard Voting Ensemble classifiers used in fake news detection from social media platforms.

JOURNAL OF EMERGING TECHNOLOGY AND DIGITAL TRANSFORMATION
ONLINE ISSN
3006-9726
PRINT ISSN
3006-9718
VOLUME . 4 ISSUE . 2 (2025)

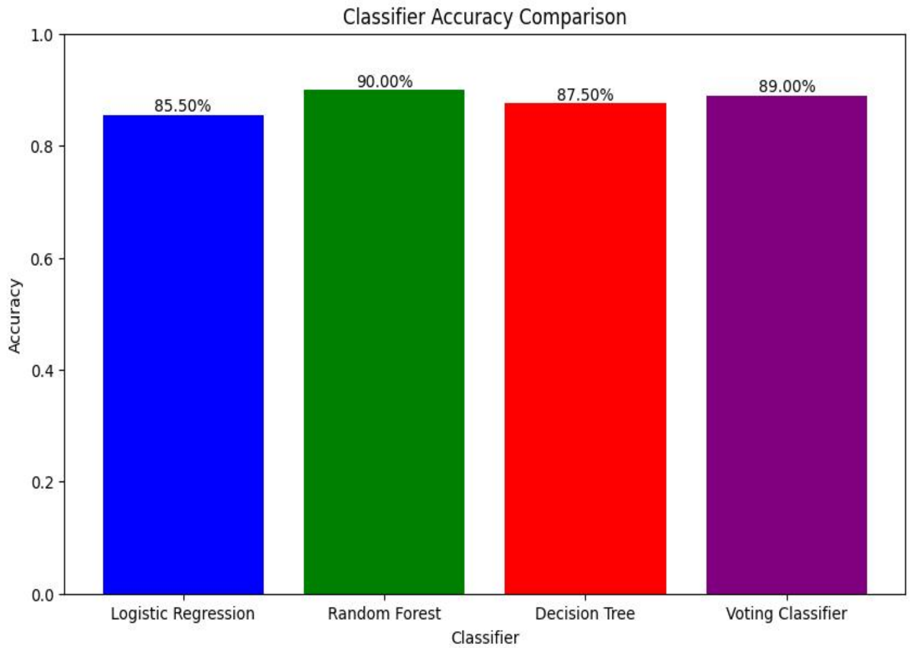

Figure 8 Bar graph of Voting on Proposed Model

Comparison of our proposed model with other models is shown in the table presented below, Outcome of all models is shown in the following table.

Table 2 Comparison Analysis of all Models (Accuracies)

Model	LR	RF	DT	SVM	Ensemble
Accuracy	83%	82%	78%	84%	89%
Precision(Class 0)	78%	72%	56%	81%	84%
Precision(Class 1)	84%	84%	85%	84%	94%
Recall(Class 0)	46%	47%	53%	45%	94%
Recall(Class 1)	96%	94%	86%	96%	85%
F1-Score(Class 0)	58%	57%	54%	58%	89%
F1-Score(Class 1)	90%	89%	86%	90%	89%
Macro Avg Precision	81%	78%	70%	82%	89%
Macro Avg Recall	71%	70%	70%	71%	89%
Macro Avg F1-	74%	73%	70%	74%	89%

ONLINE ISSN
3006-9726
PRINT ISSN
3006-9718
VOLUME . 4 ISSUE . 2 (2025)

Score					
Weighted Avg Precision	83%	81%	78%	83%	89%
Weighted Avg Recall	83%	82%	78%	84%	89%
Weighted Avg F1-Score	82%	81%	78%	82%	89%

The bars represent accuracy of four models. From the above it can be noted that as compared to other classifiers, hard voting ensemble has achieved the highest accuracy of 89% which demonstrates the capability of the model in predicting better than classifiers under discussion used in this study.

This is several individual classifiers combined into one system it can function better than individual ones because they harness collective intelligence resulting to better model. This has been marked with nearly 89% of accuracy, this indicates that classifier approach is highly accurate when practiced in this task itself to classify fake news.

All in all, it can be determined that although all the models perform fairly well, the ensemble models such as Hard Voting Ensemble are often the most accurate(Rastogi & Bansal, 2023). It points out the possible worthwhile benefit of using a number of learning algorithms to improve the evaluative results in highly differentiated areas such as fake news. Ensembling is useful because it mimics patterns and detail regarding data thus coming up with a model that is more superior and applicable.

Conclusion

The current study managed to come up with a model of the identification of fake news in digital platforms employing text mining and machine learning approach which yielded about 89% success rate with the Hard Voting ensemble model, the model comprises of three models (Logistic Regression, Random Forest and Decision Tree) combine working. Therefore, based on the results we can state that text mining methods are useful in this field therewith pointing out what can be potentially improved, such action of sarcasm and multimedia data.

Given the identified limitations and with the help of following the proposed directions for the future research, fake news detection systems can be improved, that can further help to provide a better solution to deal with the problem of fake news regarding the social media platforms.

1. **References**

2. Ahmad, I., Yousaf, M., Yousaf, S., & Ahmad, M. O. (2020). Fake News Detection Using Machine Learning Ensemble Methods. *Complexity*, *2020*. https://doi.org/10.1155/2020/8885861
3. Aïmeur, E., Amri, S., & Brassard, G. (2023). Fake news, disinformation and misinformation in social media: a review. In *Social Network Analysis and Mining* (Vol. 13, Issue 1). Springer Vienna. https://doi.org/10.1007/s13278-023-01028-5
4. Altheneyan, A., & Alhadlaq, A. (2023). Big Data ML-Based Fake News Detection Using Distributed Learning. *IEEE Access*, *11*(March), 29447–29463. https://doi.org/10.1109/ACCESS.2023.3260763
5. Benamira, A., Devillers, B., Lesot, E., Ray, A., & Malliaros, F. (2019). *Semi-Supervised Learning and Graph Neural Networks for Fake News Detection HAL Id :*

ONLINE ISSN
3006-9726
PRINT ISSN
3006-9718
VOLUME . 4 ISSUE . 2 (2025)

hal-02334445. https://doi.org/10.1145/3341161.3342958

6. Biradar, S., Saumya, S., & Chauhan, A. (2023). Combating the infodemic: COVID-19 induced fake news recognition in social media networks. *Complex & Intelligent Systems*, *9*(3), 2879–2891. https://doi.org/10.1007/s40747-022-00672-2
7. Englmeier, K. (2021). The role of text mining in mitigating the threats from fake news and misinformation in times of corona. *Procedia Computer Science*, *181*(2019), 149–156. https://doi.org/10.1016/j.procs.2021.01.115
8. Farhoudinia, B., Ozturkcan, S., & Kasap, N. (2025). Fake news in business and management literature: a systematic review of definitions, theories, methods and implications. *Aslib Journal of Information Management*, 77(2), 306–329. https://doi.org/10.1108/AJIM-09-2022-0418
9. Forest, R. (2022). *FNU-BiCNN: Fake News and Fake URL Detection using Bi-CNN*. *13*(2), 477–488.
10. Ibrahim, Z., Tulay, P., & Abdullahi, J. (2023). Multi - region machine learning - based novel ensemble approaches for predicting COVID - 19 pandemic in Africa. *Environmental Science and Pollution Research*, 3621–3643. https://doi.org/10.1007/s11356-022-22373-6
11. Jwa, H., Oh, D., Park, K., Kang, J. M., & Lim, H. (2019). exBAKE: Automatic fake news detection model based on Bidirectional Encoder Representations from Transformers (BERT). *Applied Sciences (Switzerland)*, *9*(19), 1–9. https://doi.org/10.3390/app9194062
12. Kaddoura, S., Chandrasekaran, G., Popescu, D. E., & Duraisamy, J. H. (2022). A systematic literature review on spam content detection and classification. *PeerJ Computer Science*, *8*. https://doi.org/10.7717/PEERJ-CS.830
13. Lakzaei, B., Chehreghani, M. H., & Bagheri, A. (2025). *A Decision-Based Heterogenous Graph Attention Network for Multi-Class Fake News Detection*. 1–26. http://arxiv.org/abs/2501.03290
14. Mallick, C., Mishra, S., & Ranjan, M. (2023). A cooperative deep learning model for fake news detection in online social networks. *Journal of Ambient Intelligence and Humanized Computing*, *14*(4), 4451–4460. https://doi.org/10.1007/s12652-023-04562-4
15. Matti, R. S., & Yousif, S. A. (2025). Leveraging Arabic BERT for High-Accuracy Fake News Detection. *Iraqi Journal of Science*, *66*(2), 751–764. https://doi.org/10.24996/ijs.2025.66.2.18
16. Moin Khan, Amisha Jain, Rishi Chouhan, & Sakeeb .H. Sheikh. (2020). Fake News Identification on Social Media. *International Journal of Engineering Research And*, *V9*(01), 365–370. https://doi.org/10.17577/ijertv9is010183
17. Monti, F., Frasca, F., Eynard, D., Mannion, D., & Bronstein, M. M. (2019). *Fake News Detection on Social Media using Geometric Deep Learning*. 1–15. http://arxiv.org/abs/1902.06673
18. Mookdarsanit, P., & Mookdarsanit, L. (2021). The covid-19 fake news detection in thai social texts. *Bulletin of Electrical Engineering and Informatics*, *10*(2), 988–998. https://doi.org/10.11591/eei.v10i2.2745
19. Murti, H., Sulastri, S., Santoso, D. B., Diartono, D. A., & Nugroho, K. (2025).

ONLINE ISSN
3006-9726
PRINT ISSN
3006-9718
VOLUME . 4 ISSUE . 2 (2025)

Design of Intelligent Model for Text-Based Fake News Detection Using K-Nearest Neighbor Method. *Sinkron*, *9*(1), 1–7. https://doi.org/10.33395/sinkron.v9i1.14306

20. Rahman, S., Jahan, N., Sadia, F., & Mahmud, I. (2023). Social crisis detection using Twitter based text mining-a machine learning approach. *Bulletin of Electrical Engineering and Informatics*, *12*(2), 1069–1077. https://doi.org/10.11591/eei.v12i2.3957
21. Rastogi, S., & Bansal, D. (2023). A review on fake news detection 3T's: typology, time of detection, taxonomies. *International Journal of Information Security*, *22*(1), 177–212. https://doi.org/10.1007/s10207-022-00625-3
22. Reddy, H., Raj, N., Gala, M., & Basava, A. (2020). Text-mining-based Fake News Detection Using Ensemble Methods. *International Journal of Automation and Computing*, *17*(2), 210–221. https://doi.org/10.1007/s11633-019-1216-5
23. Saeed, A., & Al Solami, E. (2023). Fake News Detection Using Machine Learning and Deep Learning Methods. *Computers, Materials and Continua*, *77*, 2079–2096. https://doi.org/10.32604/cmc.2023.030551
24. Sahoo, S. R., & Gupta, B. B. (2021). Multiple features based approach for automatic fake news detection on social networks using deep learning. *Applied Soft Computing*, *100*, 106983. https://doi.org/10.1016/j.asoc.2020.106983
25. Sharma, S., Saraswat, M., & Dubey, A. K. (2021). Fake News Detection Using Deep Learning. *Communications in Computer and Information Science*, *1459 CCIS*, 249–259. https://doi.org/10.1007/978-3-030-91305-2_19
26. Steinebach, M., Gotkowski, K., & Liu, H. (2020). Fake news detection by image montage recognition. *Journal of Cyber Security and Mobility*, *9*(2), 175–202. https://doi.org/10.13052/JCSM2245-1439.921
27. Suryavardan, S., Mishra, S., Patwa, P., Chakraborty, M., Rani, A., Reganti, A., Chadha, A., Das, A., Sheth, A., Chinnakotla, M., Ekbal, A., & Kumar, S. (2022). *Factify 2: A Multimodal Fake News and Satire News Dataset*. https://github.com/surya1701/Factify-2.0.
28. Taher, Y., Moussaoui, A., & Moussaoui, F. (2022). Automatic Fake News Detection based on Deep Learning, FastText and News Title. *International Journal of Advanced Computer Science and Applications*, *13*(1), 146–158. https://doi.org/10.14569/IJACSA.2022.0130118
29. Villela, H. F., Corrêa, F., Ribeiro, J. S. de A. N., Rabelo, A., & Carvalho, D. B. F. (2023). Fake news detection: a systematic literature review of machine learning algorithms and datasets. *Journal on Interactive Systems*, *14*(1), 47–58. https://doi.org/10.5753/jis.2023.3020
30. Wan, J., Cao, X., Yao, K., Yang, D., Peng, E., & Cao, Y. (2021). *Data Mining Technology Application in False Text Information Recognition. 2021.*
31. Wang, J., Mao, H., & Li, H. (2022). FMFN: Fine-Grained Multimodal Fusion Networks for Fake News Detection. *Applied Sciences (Switzerland)*, *12*(3). https://doi.org/10.3390/app12031093
32. Xu, X., Yu, P., Xu, Z., & Wang, J. (2025).

ONLINE ISSN
3006-9726
PRINT ISSN
3006-9718
VOLUME . 4 ISSUE . 2 (2025)

A Hybrid Attention Framework for Fake News Detection with Large Language Models. 1–4. http://arxiv.org/abs/2501.11967

33. Zellers, R., Holtzman, A., Rashkin, H., Bisk, Y., Farhadi, A., Roesner, F., & Choi, Y. (2019). Defending against neural fake news. *Advances in Neural Information Processing Systems*, *32*, 1–21.
34. Zhi, X. (2023). *A Novel Multilayer Framework for Fake News Detection Based on Multivariate Data Features Mining Read Full License A Novel Multilayer Framework for Fake News Detection Based on Multivariate Data Features Mining.*