SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

The Role of Machine Learning in Identifying Fake News and Detecting Anomalies According to a Survey

Ali Saad Flih[1] *, Areej Kadhim Abbas[2]

[1]Department of Computer Systems and Data Security, Al-Russafa Polytechnic College for Administrative specializations, Middle Technical University-Baghdad, Iraq. E-mail: alisaad.iraq@mtu.edu.iq
[2] Department of Computer Systems and Data Security, Al-Russafa Polytechnic College for Administrative specializations, Middle Technical University-Baghdad, Iraq. E-mail: areejalzuhairy@mtu.edu.iq

Abstract
Digital media has contributed to the spread of fake news. The large amount and speed of information spread is creating major societal, political, and security challenges, in which manual methods of fact-checking have become outdated and instead require automated and scalable solutions, that is why, this survey paper explores the role of machine learning (ML) in achieving the goals of dealing with this problem, with two interconnected areas, namely te classification of fake news and the detection of abnormal propagation patterns. We evaluated different models of machine learning as well, i.e., Naive Bayes, Random Forest, K-Neighbors, and Long Short-Term Memory (LSTM) networks on a selected collection of news articles. Our methodology included a three-step process of modeling, along with Natural Language Processing (NLP) steps that included Word Embeddings (Word2Vec) and Word Cloud Term Frequency analysis. Moreover, these findings also revealed that ensemble algorithms, such as the Random Forest, demonstrate a higher level of performance; i.e., their accuracy and F1-score reach 99 percent and 0.99, respectively, which is significantly higher than other algorithms. We also discuss the use of similar ML-based methods to detect anomalies in general and the section about critical issues, such as the bias of the datasets used, the adaptability of the model, and adversarial attacks, and suggest some ways of how such improvements can be reached in the future, including hybrid models and blockchain-inc Because of that, our analysis comes to a conclusion that ML provides a powerful means of improving the integrity of the information, but their effective implementation demands ongoing adjustment and the integration of the multiple disciplines.

Keywords: Fake and True news, Machine learning, Random forest model, and NB model.

Introduction

The digital age has made information both more democratic and, at the same time, contributed to the spread of misinformation and disinformation as never before [1]. The swift spread of fake news on social media and internet news media disintegrate the trust of the people, affect the political procedures, and jeopardize the health promotion efforts of the people [2, 3]. Traditional fact-checking with the help of manual methods simply cannot be scaled to address this issue because of the quantity and rate of content creation on the internet [4, 5]. Therefore, machine learning (ML) can be used to provide automated detection systems that are now indispensable. Textual and meta-information at a large scale can be examined by ML algorithms to identify trends showing falsified information, and they can be trained on the features of linguistic style, source credibility, and propagation processes through social networks [6, 7, 8]. In addition to fake news, ML is highly efficient in the general basis of anomaly detection- the detection of patterns that do not conform to normalcy [9]. The methods play a critical role in cybersecurity, fraud, and network surveillance [10]. Anomalies in the news context could be in the form of aberrant sharing or coordinated inauthenticity, as early warning signs of disinformation campaigns [11]. ML models are able to discover these signals proactively, and hence, preempt mitigation can be done [12]. The current paper introduces survey-based research assessing the power of ML in two interrelated fields, namely, the dichotomous classification of news pieces as authentic or fake, and the recognition of aberrant data patterns. We evaluate the performance of various ML models, define the implications, and stipulate the future research directions so that our digital information ecosystem will become more resilient.

Background and Related Work

1. The Fake News Landscape

Fake news also covers a continuum of content, both inadvertent misinformation and the intentional disinformation. This multifaceted character renders it an omnipresent menace to people, organizations, and democracy [13]. The problem of identification is further complicated in

social media, where the user networks are not centralized, and the dissemination of information occurs at a fast, peer-to-peer rate. This makes the curation and validation of things manually unfeasible [14]. Moreover, the asymmetry of available data, as the number of true news can greatly exceed the number of fake news that is labeled as fake, is a major challenge to the training of supervised ML models [15, 16].

2. Understanding the basics of anomaly Detection.

The term anomaly detection can be described as the identification of unusual items, events, or observations that arouse suspicions because they do not follow the majority trends of the data [17]. Examples of applications of this field include:

o Financial Security: Credit Card Fraud.
o Cybersecurity: Network intrusion detection.
o Health care: Fraudulent insurance claims or disease outbreaks detection.
o Critical Systems: Failure observation of industrial infrastructure [18, 19].

The fundamental concept of deviations of a norm could be directly applied to fake news identification in terms of its deviations in the linguistic features or distribution characteristics.

Methodology

In order to estimate the effectiveness of ML as a fake news detection tool, a survey was performed over a collection of news articles merged together around a variety of online sources, such as news agencies, social media sites, and search engines. The project pipeline for processing and analysis of data was divided into three consecutive projects aimed at comparing and contrasting modeling. Data Curation: There were 23,481 news records in the first dataset. Prior to modeling, there was a lot of pre-processing done, such as data cleaning, source verification, and labeling. The news was divided into types (e.g., Politics, World News, Government News) to get an idea of the spread of fake news by domains.

Project 1: Exploratory Data Analysis (EDA) & Sequential Modeling:-

o Task: Cleaning the dataset and initial feature analysis.
o Techniques: An elaborate cleaning policy was applied. Frequent terms and patterns that characterized fake and real news were revealed by means of Exploratory Data Analysis (EDA) and Word Cloud visualization (Figure 3, Figure 4). The later model utilized Word2Vec, which word-embeds model, and an LSTM network, which does classification to identify sequence dependencies in text.

Project 2: Machine Learning Traditional Classifiers:-

o Purpose: To categorize news with less complex interpretable models.
o Techniques: The data set was categorized with the help of the K-Neighbors Classifier and Gaussian Naive Bayes (GaussianNB) algorithms. A feature set that incorporated textual content, clarity in explanations, reputation of the source, and purported accuracy were used to classify them.

Project 3: Comparative Model Analysis:-

o Goal: To identify the best model to use in the detection of fake news.
o Methods: A direct comparison was made of a Multinomial Naïve Bayes (NB) model and a Random Forest (RF) Classifier. These models were trained and evaluated on the same dataset to test independently on their performance measures, like their accuracy, precision, recall, and the F1-score.
o Implementation: Python was used to implement and simulate all the algorithms in Google Colab notebooks. Figure 1 provides a summary of the overall methodology of fake news detection, and Figure 8 provides the same regarding anomaly detection.

The study's main contribution is that ML, specifically the Random Forest model, is an effective tool for automatically identifying fake news and anomalies, such as:-

1. A thorough review of machine learning models in the domain of fake news detection:

It systematically compares several machine learning (ML) models for the classification of news articles into "Fake" or "True" categories: Naive Bayes, Random Forest, K-Neighbors and LSTM networks. It offers a clear-cut comparison of their performance with near-perfect accuracy (up to 99%) and F1-scores (0.99), which fares much better than other models when it comes to ensemble techniques such as Random Forest.

2. Extension to Anomaly Detection:

In addition to classifying fake news, the study examines the general applicability of ML for anomaly detection, demonstrating its ability to detect anomalies in data (such as unusual sharing patterns or coordinated inauthenticity) that can be used as indicators for disinformation campaigns.

v. Results

The data clearly depict that Machine Learning (ML) models, especially the Random Forest, can be employed to accomplish the automatic classification of news articles as "Fake" or "True" with almost perfect accuracy, where these results have shown a potential application of ML to fake news detection.

Table 1 showed five examples labeled as "Fake News." The "News" column attributes sensational or negative headlines to political personalities (e.g., "An embarrassing new year is sent out by Donald Trump."). The "Form" column features a segment of the article body. The "Items" column confirms the label "News" (implied to be "Fake News").

Table 2 listed some five examples labeled "True News," specifically under politics. The headlines are more neutral and factual (e.g., "Transgender recruits will be accepted by the U.S. military."). In the "Form" column, there will most likely be what indicates a dateline (such as Washington (Reuters), a standard attribute of serious journalism, and likely a powerful element that the model can leverage to grab.

A. Fake and True News.

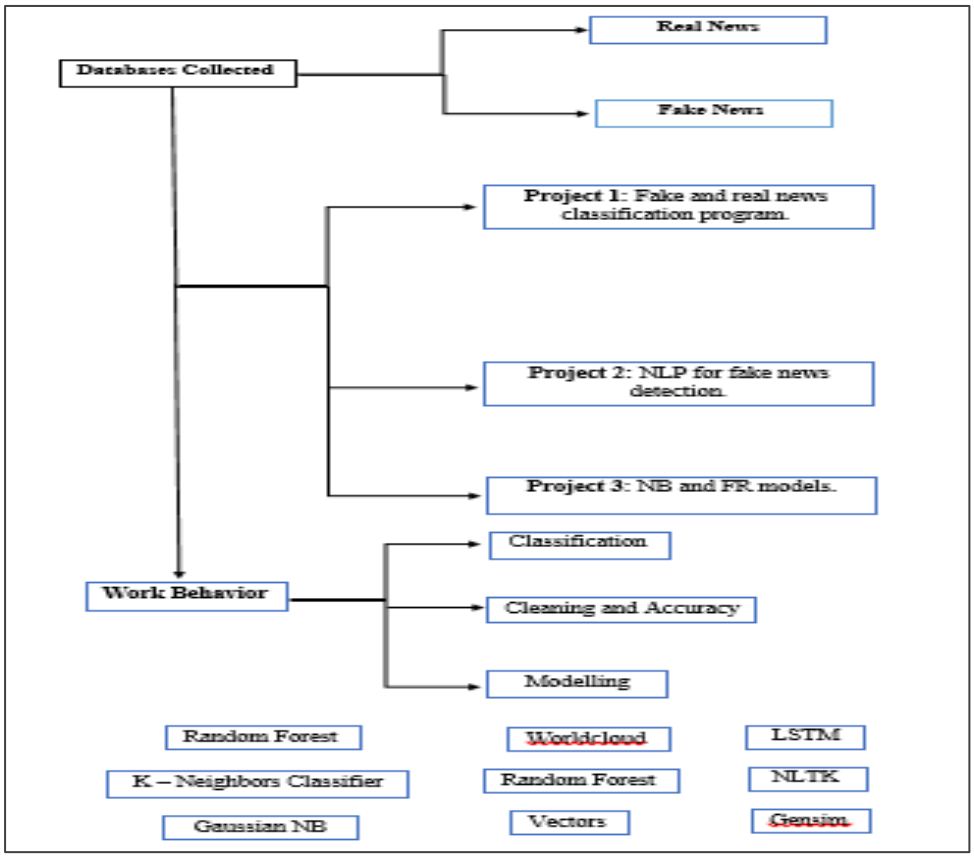


Figure 1. Flow chart of the databases' methodology of fake news detecting using machine learning.

Table 1. Classifications of databases into fake news.

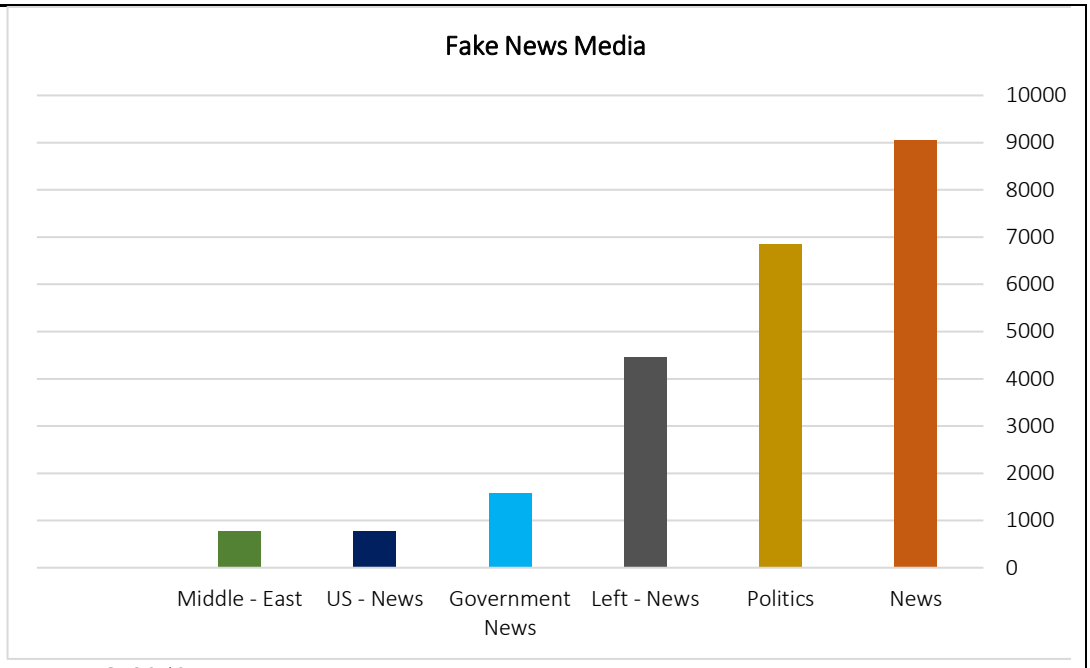
	News	Form	Items
0	An embarrassing new year is sent out by Donald Trump.	Donald Trump just was unable to hope for all Americans.	News
1	A drunken, boastful trump staffer began speaking Russian.	Devin Nu is the chairman of the House Intelligence Committee.	News
2	Sheriff David Clarke turns into a joke on the internet.	The former milwauk was made public on Friday.	News
3	Trump is so fixated with Obama that he even has his name.	Donald Trump made his announcement on Christmas Day.	News
4	Donald Trump dur was just called out by Pope Francis.	On Christmas Day, Pope Francis used his yearly mes.	News

Table 2. Classifications of databases into true news.

	News	Form	Items
0	The looming U.S. budget battle causes republicans to switch.	Washington (Reuters) - a conservationist's head.	Politics news
1	Transgender recruits will be accepted by the U.S. military.	Washington (Reuters) - individuals who identify as trans will.	Politics news
2	Senior republican senator from the United States: "Allow Mr. Muell."	Reuters/Washington: The special counsel is investigating.	Politics news
3	An Australian diplomat assisted the FBI's Russia investigation.	Washington (Reuters) - an aide to Trump's campaign.	Politics news
4	Trump wants to raise the postal service's fees significantly.	Reuters or Seattle or Washington of President Donald.	Politics news

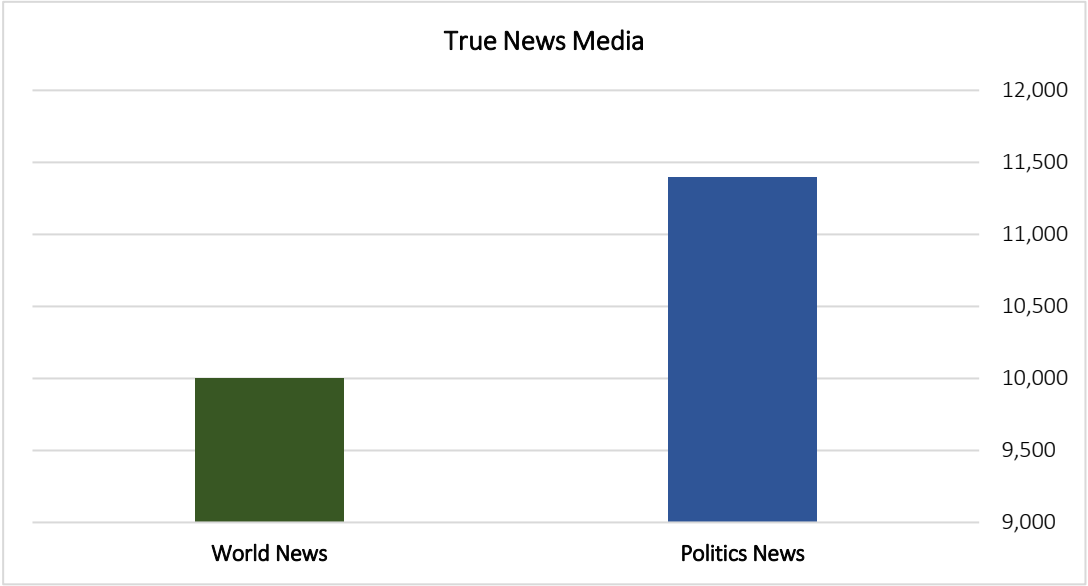
Table 3 summarized the performance of the final selected model on the test dataset (which the data was not trained on), which very high scores suggest good performance. 0 & 1: they represent the two classes, and we can infer that 0 = Fake News and 1 = True News (or else it can be opposed; practically, scores are close to each other). The accuracy near 0.98 suggests that, in all the articles the model considered as Fake, 98 percent were actually Fake, and hence a high degree of accuracy implies that the number of false positives (true news called Fake) is smaller. In the meantime, a Recall of 0.98 implies that the model identified all 98% of all Fake articles in the dataset; therefore, a high Recall value implies that there are fewer false negatives.

Moreover, Precision and Recall have a harmonic mean of the F1-score (~0.98), which is a solitary and strong statistic that is only obtained when both the Precision and Recall are adequately elevated. A value of 0.98 is rare. What are the number of examples of each of the classes in the test set? (Approximately 5898 in Class 0, approximately 5325 in Class 1), However, accuracy (0.98) indicates that the model correctly predicted the class of fake or true 98% of the time in an independent test set of 11,225 news articles. Macro Avg. & Weighted Avg. (~0.98) witness that the higher performance uniformly favors both classes and is not skewed by the datapoints of one class type slightly outweighed by the datapoints of another (i.e., a slightly imbalanced dataset of more Class 0 examples).



- NEWS: 9050
- POLITICS: 6841
- LEFT - NEWS: 4459
- GOVERNMENT NEWS: 1570
- US - NEWS: 783
- MIDDLE - EAST: 778
- TOTAL RECORDS: 23,481

A. Fake news



- POLITICS NEWS: 11,400
- WORLD NEWS: 10,000
- TOTAL RECORDS: 14,400

B. True news

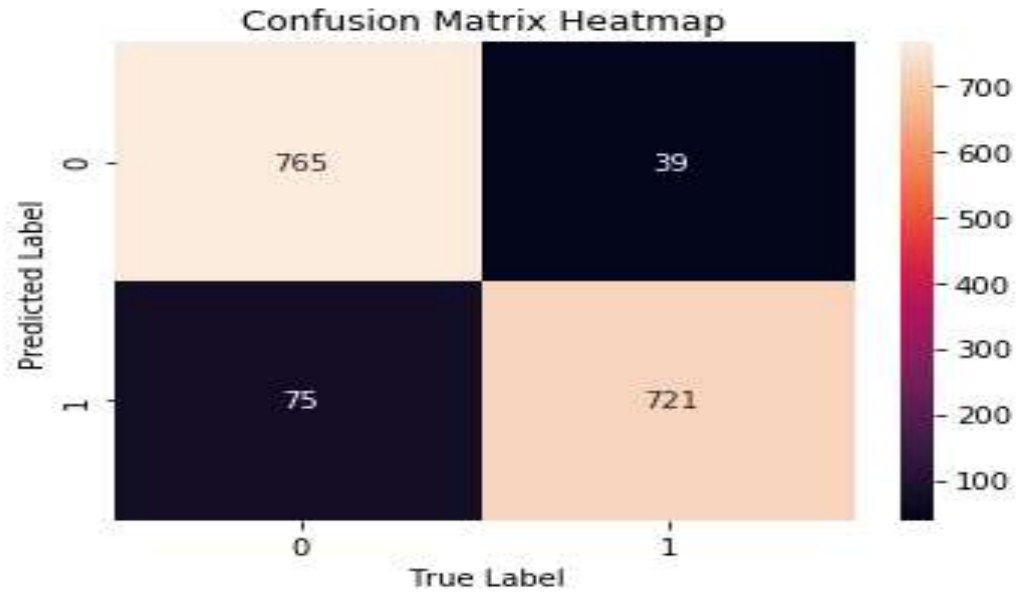
Figure 2. Identifying types of media that get fake news and true news.

The data that was described in **Figure 2** concerning the dataset referred to as Fake News. It has a total of 23,481 records, of which most of them fall under broad topics such as "news" (9,050) and "politics" (6,841). The large subgroup can further be classified as "left-news" (4,459), which suggests that the data on fake news has a possible political bias. Other prominent segments are also present, such as government news and geographical categories, such as US news and the Middle East.

Also, it described the dataset of True News, that is much smaller, comprising 14,400 records in total. The two main ones are the categories of politics news (11,400) and world news (10,000). These categorical groups constitute more than the total record, which means that the articles may fall into several categories.

Table 3. Determining classification variables in detecting news of output – test.

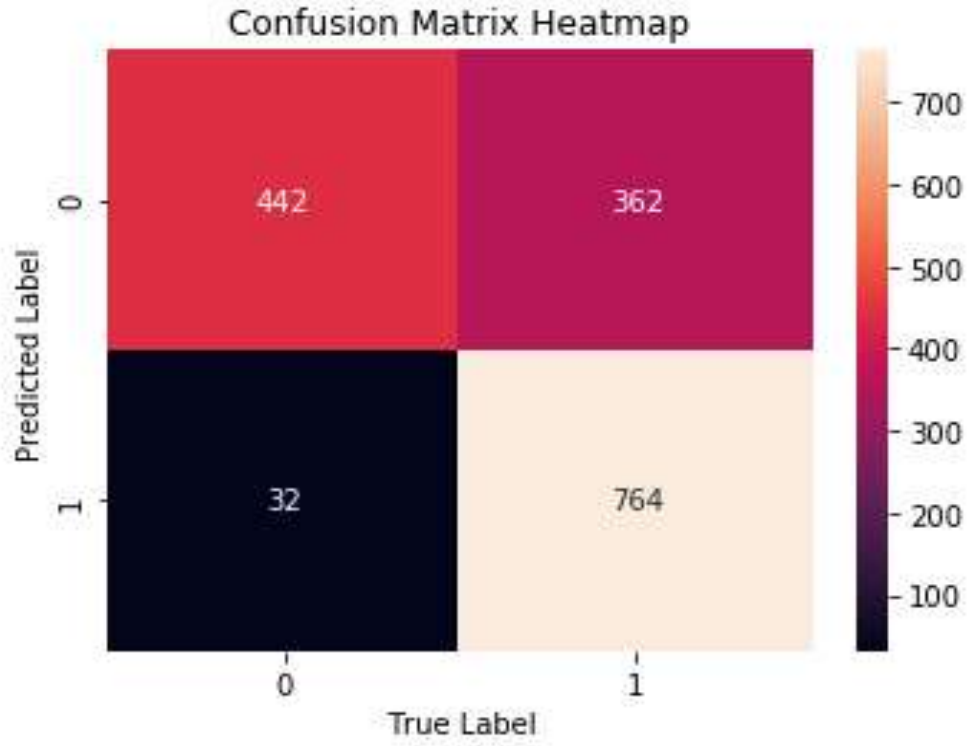
PARAMETERS	PRECISION	RECALL	F1 – SCORE	SUPPORT
0	0.98	0.98	0.98	5898
1	0.97	0.98	0.98	5325
ACCURACY	-	-	0.98	11225
MACRO, AVERAGE	0.98	0.97	0.98	11225
WEIGHTED, AVERAGE	0.97	0.97	0.98	11225



PRECISION	*RECALL*	*F1-SCORE*	*SUPPORT*
0 0.91	0.95	0.93	804
1 0.95	0.91	0.93	796
ACCURACY		0.93	1600
MACRO AVG	0.92	0.92	0.92 1600
WEIGHTED AVG	0.92	0.92	0.92 1600

Figure 3. Importing modelling of databases for news by the Gaussian NB model.

Figure 3 gives the performance statistics of the Gaussian Naive Bayes (NB) model. The model shows good and balanced performance; both the accuracy and recall of both classes (0 and 1, which are fake news and true news, respectively) are high. The F1-scores of 0.93 of each class suggest a balanced measure of accuracy and recall. This is indicated by the high overall accuracy of 93 percent and equivalent macro and weighted averages of 0.92.



PRECISION	RECALL	F1-SCORE	SUPPORT
0	0.93	0.55	0.69	804
1	0.68	0.96	0.80	796
ACCURACY			0.75	1600
MACRO AVG	0.81	0.75	0.74	1600
WEIGHTED AVG	0.81	0.75	0.74	1600

Figure 4. Importing modelling of databases for news using the K-Neighbors Classifier model.

The results of the K-Neighbors Classifier (KNN) were illustrated in **Figure 4**, and the performance in this test is quite the opposite. The model is highly imbalanced in terms of its measures between the two classes. In case 0 (fake news), the precision is good (0.93), but the recall is very low (0.55); that is, the case accurately recognizes a majority of the news that it labels as fake, but fails to detect most of the fake news in the dataset. The same thing happens with class 1 (true news), which have a recall as high as (0.96) but low precision (0.68), giving it a lot of false positives.
This **Table 4** is crucial as it compares two different machine learning algorithms, highlighting that the choice of model is a key "role" in effective fake news detection including Multinomial Naive Bayes (NB) Model of performance into Good, but significantly worse than Random Forest, where accuracy with 93.98% and F1-Score: 0.93, which found [5468 381] → 5468 correct Fake News predictions, 381 False Positives (True News mistaken for Fake).

Table 4. Predicting modelling fake and true news databases in terms of accuracy based on: a. NB model; b. RF model.

A. Multinomial Naive Bayes:

PRECISION	RECALL	F1-SCORE	SUPPORT
0	0.93	0.93	0.93	5849
1	0.93	0.92	0.93	5376
ACCURACY			0.93	11225
MACRO AVG	0.93	0.93	0.93	11225

WEIGHTED AVG	0.93	0.93	0.93	11225

- [5468 381]
- [407 4969]
- ACCURACY: 93.98%
- F1_SCORE : 0.93

B. RANDOM FOREST

PRECISION	RECALL	F1-SCORE	SUPPORT
0	0.99	1.00	0.99	5849
1	1.00	0.99	0.99	5376
ACCURACY			0.99	11225
MACRO AVG	0.99	0.99	0.99	11225
WEIGHTED AVG	0.99	0.99	0.99	11225

- [5835 14]
- [60 5316]
- Accuracy: 98.34%
- f1_score : 0.993

Also, 407 4969- True News predictions, 407 False Negatives (Fake News as True). The NB model is a more straightforward and probabilistic algorithm. It errs more (788 overall errors), allowing a fair deal of fake news through (407 FN) and wrongly classifying true news as fake (381 FP). Moreover, the RF Model has accuracy of 98.34% and an F1-score of 0.993, where [5835 14] would yield 5835 correct Fake News predictions, only 14 False Positives, and [60 5316] would yield 5316 correct True News predictions, only 60 False Negatives. Random Forest model is a stronger ensemble approach. It greatly minimizes mistakes (a total of 74 errors). It is also particularly good at not returning false positives (only 14); that is, it does not tend to block out true news. It is also far more effective at pouncing upon fake news (it only 60 make it through).

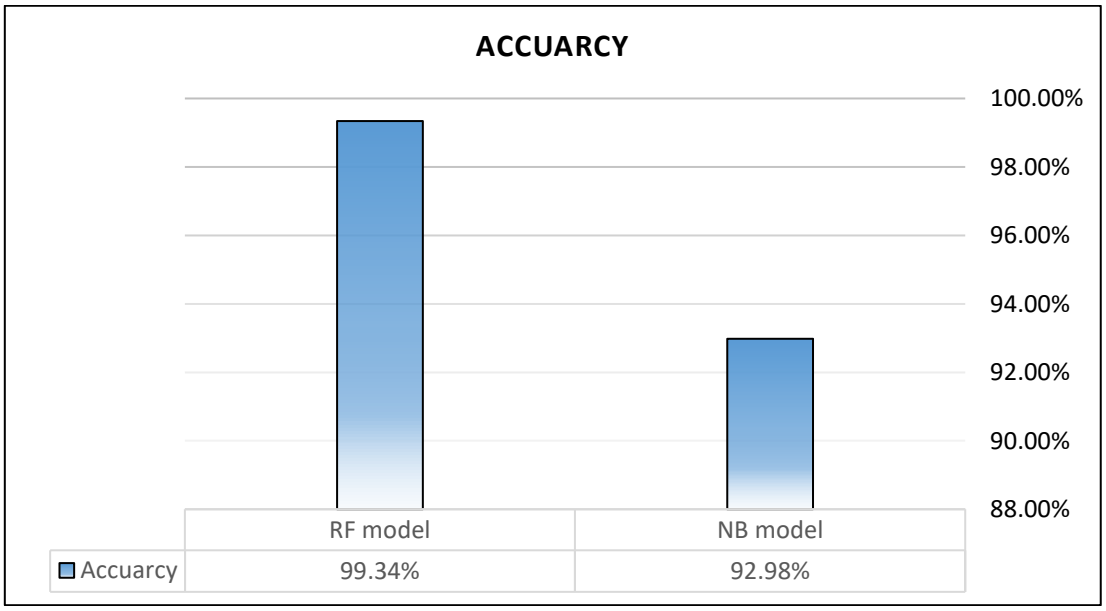
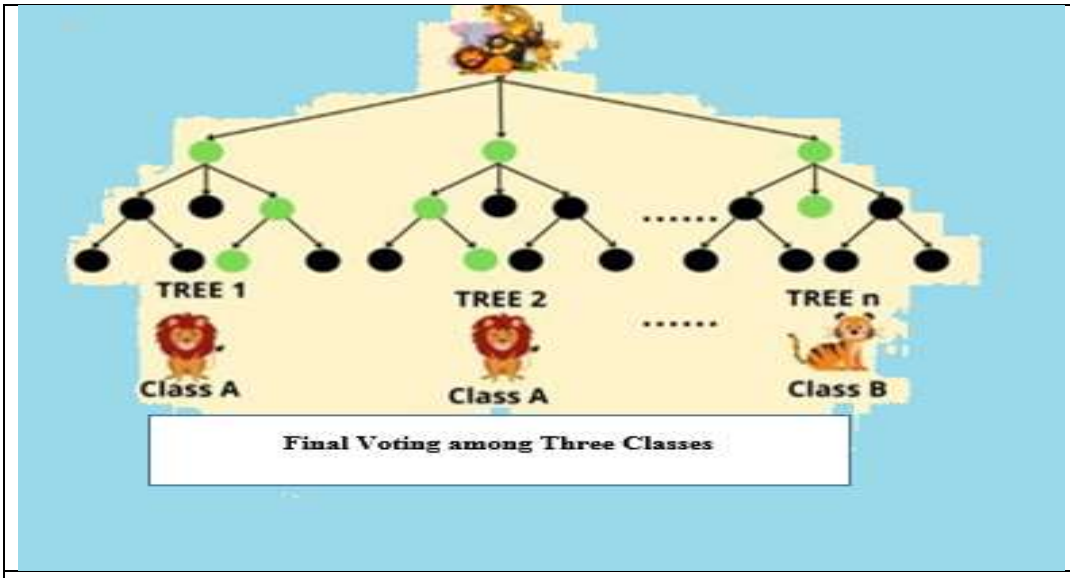


Figure 5. Determining the accuracy rate of the RF model and NB model in predicting of database modelling.

Figure 5 simply gave a direct point of comparison of the ultimate accuracy rates of the Random Forest (RF) model and the Naive Bayes (NB) model. The accuracy of the model is 99.34% in the Random Forest, which is significantly higher than the model of the Gaussian NB, as it is 92.98%. This implies that the ensemble learning technique of Random Forest, which is capable of capturing the complicated, non-linear relationships within the data, is much more efficient when it comes to this news classification problem.

B. Anomalies Detecting:



- Class A {Tree 1}: Lion
- Class A {Tree 2}: Lion
- Class B {tree n}: Cat

Figure 6. Detecting anomalies: Random forest modelling of a cat between two lions.

Figure 6 showed that it is our duty to classify the animals. In this case, we are applying a random forest. This results revealed various abnormalities that result in the discovery of just one fugitive among two lions. According to the findings of **Table 5**, the performance of the models in the specified classification task is highly dissimilar. The model of Random Forest proves to be the best in this analysis because it has an accuracy of 0.98, and it is the highest-performing algorithm in this analysis. Support Vector Machine (SVM) also demonstrates good performance with accuracy of 0.90, which is a sign of good classification, though not the best. Conversely, the Naive Bayes classifier produces a more average outcome with an accurate value of 0.74, indicating that it is not well-fit to the exact pattern or distribution of features that this dataset assumes. However, most notably, the Convolutional Neural Network (CNN) achieves a very low level of accuracy of 0.084 which is much lower than the chance level of a normal binary classification.

Table 5. Evaluating the accuracy of models in importing data.

MODELS	ACCURACY
RANDOM FOREST	0.98
CNN	0.084
SVM	0.90
NAÏVE BAYES	0.74

Discussion

The present digital age has made it extremely hard to tell the difference between misinformation and fake news due to the affordability of information. The growth of online news sources and social media platforms has expedited the virality of misleading news to pandemic levels. The pandemic degree of use of social media platforms and online news sources to transmit deceptive information into the audience has risen (at least two-thirds) [20]. The growth has led to the development of interest in using machine learning (ML) techniques to detect anomalies in the content of news and combat fake news. A more recent survey demonstrates that machine learning has been successful in these areas; hence, we need additional studies regarding its objectives, approaches, and implications. The impact of fake news on the election process, reaction to a nature crisis, and personal opinion is immense [21]. The need to have efficient ways of detecting and preventing such threats is making researchers and engineers resort to machine learning is a solution. Despite the development of machine learning, the deployment of such systems is not always been successful to date [23]. Bias in training data is one of them and can lead to wrong classifications and even strengthen preconceptions. [24, 25, 26].

A number of methods of identifying bogus news are the focus of a number of studies. The most important findings of the research review are presented below. Researchers have used labeled

datasets to test various machine learning classifiers, including Support Vector Machines (SVM), Decision Trees, and Naïve Bayes, among others [27]. These methods presuppose the selection of features, along with the quality of the features and data, to be effective, despite the high accuracy of these methods. User interaction tracking and news sharing have been examined through network monitoring in order to discover patterns of misinformation. [28] Word embedding, sentiment analysis, and metadata clues (like the trustworthiness of a source and author) can all play a critical role in the accuracy of machine intelligence, such as a Random Forest and K-Neighbor classifier. Machine learning models are superior to traditional machine learning methods, even though they need a substantial quantity of training data to stimulate the best results [29, 30]. Although paradoxically accurate, models are faced by hostile changes, biased data, and changing misinformation tactics. One of the developments in the future is the use of blockchain and machine training to screen news sources to create stronger models capable of overcoming new types of misinformation. In conclusion, despite the fact that machine learning may have potential tools in detecting abnormalities and helping to avoid fake information, there is a need to address the relevant issues to make sure that such technologies enhance the information ecosystem within society. [31, 32]

Conclusion

This survey has critically assessed the use of machine learning in detecting fake news and anomalous activities. We have shown by applying and comparing a set of models that high accuracy (>98%) can be obtained by high-quality algorithms (in particular, Random Forest) in misinformation classification. The paper also raises the suitability of their general applicability of similar ML paradigms to general anomaly detection, which supports their flexibility.

The issues of data bias, model vulnerability, and the dynamic character of threats are still a formidable challenge. To solve the problems, a combined approach to creating clear, flexible, and resilient systems is needed. Future studies are advised to focus on hybrid processes incorporating linguistic analysis with network science, and on the application of new technologies, such as blockchain, to verify them. Finally, the mitigation of fake news is not only a technical issue but a social one, as it requires the cooperation between computer scientists, policymakers, and social scientists to protect the discourse of the population.

References

1. A Multi-Aspect Classification Ensemble Approach for Profiling Fake News Spreaders on Twitter Notebook for PAN at CLEF 2020.
2. Fake News Spreaders Profiling Through Behavioural Analysis Notebook for PAN at CLEF 2020.
3. A Survey on Recent Advances in Machine Learning Techniques for Fake News Detection.
4. A survey on fake news and rumour detection techniques.
5. Fake News Detection on Social Media: A Data Mining Perspective.
6. A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities.
7. XINYI ZHOU, Syracuse University, USA.
8. REZA ZAFARANI, Syracuse University, USA. https://link.springer.com/chapter/10.1007/978-3-030-51859-2_31
9. Allcott, H., & Gentzkow, M. (2017). Social Media and Fake News in the 2016 Election. Journal of Economic Perspectives, 31(2), 211–236.
10. Alnabhan, M. Q., & Branco, P. (2024). Credibility Assessment of News Articles with Weak Supervision. 2024, 114435(114459), 114435–114459.
11. Andreadou, K., Papadopoulos, S., Apostolidis, L., Krithara, A., & Kompatsiaris, Y. (2015). Media Revealr: A Social Multimedia Monitoring and Intelligence System for Web Multimedia Verification. Intelligence and Security Informatics, 1–20. https://doi.org/10.1007/978-3-319-18455-5_1. Balmas, M. (2014). When Fake News Becomes Real. Communication Research, 41(3), 430–454.
12. Routledge. Bessi, A., & Ferrara, E. (2016). Social bots distort the 2016 U.S. Presidential election online discussion. First Monday, 21(11).
13. Bian, T., Xiao, X., Xu, T., Zhao, P., Huang, W., Rong, Y., & Huang, J. (2020). Rumour Detection on Social Media with Bi-directional Graph Convolutional Networks. Proceedings of the AAAI Conference on Artificial Intelligence, 549–556.
14. Birla, B., & Patel, S. (2014). An implementation on web log mining. International Journal of Advanced Research in Computer Science and Software Engineering, 4(2), 68–73.
15. Lai, S., Xu, L., Liu, K., & Zhao, J. (2015). Recurrent Convolutional Neural Networks for Text Classification. Proceedings of the AAAI Conference on Artificial Intelligence. Proceedings of the AAAI Conference on Artificial Intelligence.

16. Mridha, M. F., Keya, A. J., Hamid, M. A., Monowar, M. M., & Rahman, M. S. (2021). A comprehensive review on fake news detection with deep learning. IEEE Access, 9, 156151-156170.
17. Nguyen, D. Q., Vu, T., & Nguyen, A. T. (2020). BERTweet: A pre-trained language model for English Tweets. ArXiv:2005.10200.
18. Padiya, S. D., Pandit, R., & Patel, S. (2012). A System for MANET to detect selfish nodes using NS2. International Journal of Engineering of Science Innovation and Technology.
19. T. Salman, D. Bhamare, A. Erbad, R. Jain, and M. Samaka, Machine learning for anomaly detection and categorization in multi-cloud environments, in Proc. IEEE 4th Int. Conf. Cyber Secur. Cloud Comput. (CSCloud), 3rd IEEE Int. Conf. Scalable Smart Cloud (SSC), Jun. 2017, pp. 97–103.
20. Subutai Ahmad, Alexander Lavin, Scott Purdy, Zuha Agha —Unsupervised real-time anomaly detection for streaming data, Neurocomputing (2017).
21. Min Du, Feifei Li, GuinengZheng, Vivek Srikumar, DeepLog: Anomaly Detection and Diagnosis from System Logs, 2017.
22. D. A. Kumar and S. R. Venugopalan, A novel algorithm for network anomaly detection using adaptive machine learning, in Progress in Advanced Computing and Intelligent Engineering (Advances in Intelligent Systems and Computing), vol. 564. 2018, pp. 59–69.
23. S. Naseer, Y. Saleem, S. Khalid, M. K. Bashir, J. Han, M. M. Iqbal, and K. Han, Enhanced network anomaly detection based on deep neural networks, IEEE Access, vol. 6, pp. 48231– 48246, 2018.
24. T.-Y.KimandS.-B.Cho,__Web traffic anomaly detection using CLSTM neural networks,‘‘ Expert Syst. Appl., vol. 106, pp. 66–76, Sep. 2018.
25. B. G. Atli, Y. Miche, A. Kalliola, I. Oliver, S. Holtmanns, and A. Lendasse, Anomaly-based intrusion detection using extreme learning machine and aggregation of network traffic statistics in probability space, Cognit. Comput., vol. 10, no. 5, pp. 848–863, Oct. 2018.
26. Y. Tian, M. Mirzabagheri, S. M. H. Bamakan, H. Wang, and Q. Qu, __Ramp loss one-class support vector machine: a robust and effective approach to anomaly detection problems,‘‘ Neurocomputing, vol. 310, pp. 223–235, Oct. 2018.
27. R. Abdulhammed, M. Faezipour, A. Abuzneid, and A. AbuMallouh, __Deep and machine learning approaches for anomaly-based intrusion detection of imbalanced network traffic,‘‘ IEEE Sensors Lett., vol. 3, no. 1, pp. 1–4, Jan. 2019.
28. S. D. Anton, S. Kanoor, D. Fraunholz, and H. D. Schotten, __Evaluation of machine learning-based anomaly detection algorithms on an industrial Modbus/TCP data set,‘‘ in Proc 13th Int. Conf. Availability, Rel. Secur., Aug. 2018, vol. 41, no. 9, pp. 1–41.
29. J. Zhang, R. Gardner, and I. Vukotic, __Anomaly detection in wide area network meshes using two machine learning algorithms,‘‘ Future Gener. Comput. Syst., vol. 93, pp. 418–426, Apr. 2019.
30. A. Brown, A. Tuor, B. Hutchinson, and N. Nichols, __Recurrent neural network attention mechanisms for interpretable system log anomaly detection,‘‘ in Proc. 1st Workshop Mach. Learn. Comput. Syst., Jun. 2018, pp. 1–8.
31. L. F. Maimo, A. L. P. Gomez, F. J. G. Clemente, M. G. Perez, and G. M. Perez, __A self-adaptive deep learning-based system for anomaly detection in 5G networks, " IEEE Access, vol. 6, pp. 7700–7712, 2018.
32. B. Hussain, Q. Du, and P. Ren, __Semi-supervised learning based big data-driven anomaly detection in mobile wireless networks,‘‘ China Commun., vol. 15, no. 4, pp. 41–57, Apr. 2018.