A Survey on Fake News Detection Using Machine Learning

Akanksha Singh
Department of software engineering
Delhi Technological University
New Delhi, India
sakanksha1115@gmail.com

Sanjay Patidar
Department of software engineering
Delhi Technological University
New Delhi, India
sanjaypatidar@dtu.ac.in

2022 4th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N) | 978-1-6654-7436-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICAC3N56670.2022.10074450

***Abstract*— One of the most extensively used platforms now a days is social media. It is easily accessible sources of day today news for almost everyone throughout the globe due to its relatively low cost, ease of access and rapid dissemination. However, this news comes with a risk of being faulty or fake to mislead the readers. Fake news have a profound effect on our social life, in fact, in all fields, particularly politics and education. The spread of fake news can cause as much social ills in the same way as open damage with hazardous impacts. Spreading incorrectly data via online media to stand out enough to be noticed or monetary and political increase is common on social media these days. To avoid the disastrous consequences because of this prevailing fake news its detection is important in early stage. This survey presents a comprehensive overview of ongoing and previous studies on false news detection using various machine learning methods.**

Keywords*— *Machine learning; Detection; Fake news; social media.

I. INTRODUCTION

For a long time, social media has taken over a meaningful place in people's life. Fake news primarily prevails via social media and articles available online. Fake news indulges politics, democracy, education as well as finance and business at risk. Even while false news is not a new issue, people these days place a larger focus on social media, which leads to the acceptance of deceitful remarks and the subsequent propagation of the same wrong information. It's getting harder to tell the difference between accurate and misleading news these days, which leads to confusion and complications. Hence it is not surprising that the current research studies are not just focusing on better understanding but also working towards automating the detection of false information.

The subterfuge film "Plandemic" which was released online in May 2020 caused a huge number of fake news to prevail in public that misguided the people that wearing a mask could increase the activation of coronavirus. Not only this there exists much misinformation related to deaths caused by coronavirus as well as the symptoms, the number of people getting infected, and vaccine availability and its effectiveness. Times Now and Aaj Tak released a film depicting a burial place in Xinjiang for Chinese troops who perished during the 1962 Indo-China conflict, claiming that it showed the grave of soldiers killed in the June 20 fight. When the Nation was going through the demonetization, there appeared false news saying that the notes of five hundred and two thousand are containing some kind of chip in them.

These illustrations demonstrate how misleading news shapes our intellectual, religious, political, and other views and relationships. This misleading information are poisonous for any humankind and should be eliminated or atleast reviewed and reduced at early stages to save the harm that they could possibly do to the Nation.

II. RELATED WORK

In 2017, Arkaitz Zubiaga, Elena Kochkina, Rob Procter and Maria Liakata used sequential classifiers to classify rumour positions on social media platforms in 2017. They leverage Twitter as primary social media application, categorising tweets into four categories. On eight data sets, they employed four sequential classifiers namely hawkes algorithms, along with linear CRF, long short term memory (LSTM), and tree CRF. They discovered that sequential classifiers that exploit the reciting property outperformed non sequential classifiers in social media engagement, and that LSTM outperforms other sequential classifiers [1].

In 2017, Ahmed, Traore, and Saad identified false news, by developing an AI model using n-gram analysis. They have used five machine learning classification methods i.e. K-Nearest Neighbor, Stochastic Gradient Descent, Decision Trees, Linear Support Vector Machines and Support Vector Machines. And found that Linear Support Vector Machine (LSVM) outperforms others with the precision of 92 percent [2].

In 2018, Kotteti, Na Li, and Qian worked on using data imputation to improve the discovery of false information. A unique data preparation strategy were applied by them in order to fill the empty/missing values in the proposed data for enhancing the performance. For quantitative and hierarchical variables they used data modeling for missing values by choosing most frequent values for columns. To make up for the missing values, three things were performed: 1. Eliminate columns with the empty/missing values, 2. Replaced missing values with the blank text, and 3. applied missing values using data imitation techniques. Finally they discovered and concluded that Multi Layer Perceptron(MLP) classes enhanced the accuracy by 16 percent [3].

In 2018, Supanya and Prabhas have used naïve Bayes, SVM, and neural networks for the detection the fake news and calculated the performance measures they found that naïve Bayes has 96.08% and neural network and SVM 99.90% accuracy. Through this experiment, they found out that neural networks and support vector machines are having significant accuracy and high confidence [4].

In 2018, Akshay and Amey worked on detecting false news, and they suggested a strategy that we may use on Facebook. For predicting, he employed Naive Bayes. They utilised a dataset with 11000 articles sorted by categories (index, text,

title and label). The dataset comprises news on science and industry in addition to politics. They used the title and content as their major source for implementation, as well as some n-gram references. And finally they discovered that nave bayes had accuracy of 93.10 percent and presented the way to improve this[5].

In 2018, Manisha, Sanjana, and Tanvi, have suggested a Support Vector Machine news detection model for false or real news that has an accuracy of 87%. She had recognized comedy, negative words, ridiculousness, syntax, and punctuation using five predictive features. Its goal was to ensure that the substance of a news piece was accurate [6].

Bhowmik, Ajao, and Zargari suggested a methodology in 2019 which utilizes a mixture of (CNN) and (RNN) models to detect phony news tweets from the Twitter posts. They included rumour stories namely Ferguson Shooting Ottawa Shooting, Charlie Hebdo, Sydney Siege, Germanwing Crash. Their suggested work on a hybridization of CNN-RNN intuitively recognises crucial features connected to misleading news articles without any foreknowledge of the news and achieves an accuracy of over 80% [7].

In 2019, Riece, Correia, Murai, Veloso, and Benevuto are working on looking for a range of elements in news articles, postings, and stories that might assist identify false news with increased precision. He demonstrated the significance of these new qualities in evaluating bogus news. Discrimination, integrity, involvement, domain location, and temporal patterns are some of these characteristics. They used 2282 Buzzfeed items in their analysis (news articles). Using KNN, Nave Bayes, Random. Forest, XGBoost, Support Vector Machine(SVM), and they analysed and described the strengths and limits of this technology, and discovered that XGBoost performed better when compared with other with an accuracy of 0.86 [8].

In 2020, Yuan, Q. Ma, and W. Zhou suggested a model to identify false news. They proposed a structure-aware multi-head attention network (SMAN) based technique in their model. This strategy is based on the publications' and users' trustworthiness. Real-world datasets were employed in this technique. They tested this model against three distinct datasets and discovered that it had a high level of accuracy [9].

In 2020, Ozbay and Alatas have used AI techniques for detecting fake news. In the first phase, they preprocessed the dataset to transform unstructured data into structured data, and then they used text mining to construct about twenty-three supervised AI algorithms. They applied these algorithms to about three real-world data sets and found the accuracy and performance measures accordingly. The best average value they got was by using a decision tree, ZeroR, CVPS, and WIHW algorithms [10].

In 2020, Ankit, Sudakar, and Anil demonstrated a basic strategy for detecting false news on social media using a K-nearest neighbor classifier, which obtained an accuracy of roughly 79 percent when evaluated against a sample of Facebook news articles [11].

In 2021, Arun Nagaraja and colleagues showed in their study that false information mostly circulates through social media and is propagated further without investigating the true data. They applied various NLP techniques and two ML algorithms i.e., naïve Bayes and SVM which gives 63% and 75% accuracy respectively [12].

In 2021, Shifath, S. Islam, and F. Khan have suggested strategy that is transformer-based for identifying COVID-19 false information. They experimented using CNN and conventional lexicons. The dataset consists of COVID-19-related social media postings with classifications indicating whether they are false or legitimate. They also explored other hyper settings and explored various transformer-based models. RoBERTa shows that the maximum accuracy is 0.979. [13]

In [15] proposed a model how fake news spread on social media and how the internet affects the diffusion of false information in creating and spreading. They also discussed the solutions to reduce the dissemination of false information and provided the future research aspects in this area.

Tacchini et al. [14] produced a dataset using two different groups having news from conspiracy and science. To enhance the accuracy of their model they have leveraged the social component. They have used hypothesis and harmonic approach to divide stories into false and credible content. Harmonic algorithm exchanges the info between individuals who liked similar context.

III. Machine Learning Methods

Various stages involved in different experiments performed in the literature are given in the following flowchart **Fig.1**

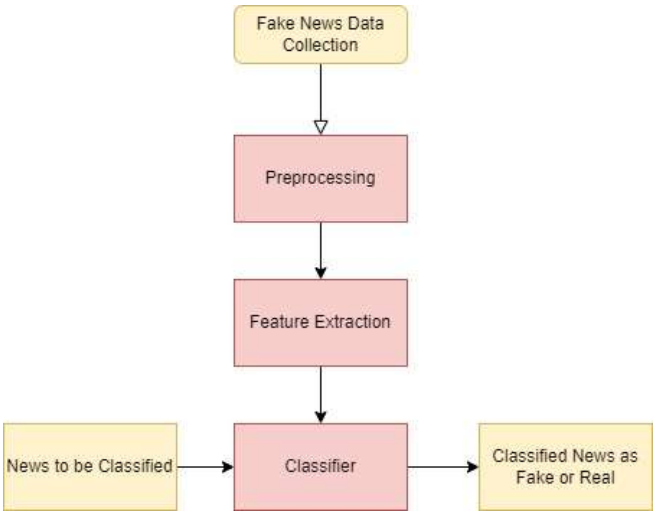

Fig.1 Flowchart for fake news detection

As discussed in related works various machine learning algorithms have been used by different researchers to detect the false news. Following we are giving a brief overview of few of those algorithms: Machine Learning Algorithms

A. Decision Tree

It is a Machine Learning supervised classification algorithm which means we have to clarify what the information is and what the relating yield is in the preparation information. It is a tree-like construction where the information is consistently parted by a specific boundary. The elements of a dataset are addressed by the inner nodes and branches that address the

decision rules and each leaf address the ultimate results or choices [2][10].

B. Naïve Bayes

The Naive Bayes approach, a supervised machine learning methodology based on the well-known Bayes theorem, is used to tackle classification issues. It's most commonly used for text classification with a big training dataset. One of the most simple and effective classification methods is the Naive Bayes Classifier. It allows for the rapid building of machine learning models as well as effective training and testing to make speedy predictions. It's a probabilistic classifier, which implies A term's frequency in a text may be determined using Term Frequency. The term n in the formula represents the number of times a phrase appears in each document or text. As a result, each term has a TF value. the algorithm's whole basis is built on probabilities that have been computed, and it predicts based on an item's likelihood [4][5][8][11].

Naïve Bayes Equation:

P(G|H) = P(H|G)*P(G) / P(H)

Where:

P(G|H) is the posterior Probability. P(H|G) is the Likelihood.
P(G) is the Class Prior Probability.
P(H) is the Predictor of Prior Probability

C. Support Vector Machine

SVM (Support vector machine) approach aims to find the hyperplane (where N is the number of characteristics) that clearly arranges the principal elements in an N-dimensional space. There is an assortment of hyperplanes from which to isolate the two kinds of informative items. Our point is to track down the plane with the biggest edge, or distance between relevant items from the two classes. Boosting the edge distance gives some support, making it more straightforward to arrange ensuing information points [2][4][6][8].

D. Random Forest

The random forest is a supervised ML algorithm. It is basically established on the outfit learning techniques where different classifiers are united to deal with an issue and to chip away at the display of the presentation of the model. Random Forest is a classifier that calculates the dataset's predicted precision by averaging the results of many decision trees applied to different subsets of the dataset [8].

E. CNN

CNN (Convolutional neural network) is the kind of ANN (artificial neural network) that analyses data using perceptrons, which are a machine learning unit method. Image processing, natural language processing, and other structured activities can all benefit from CNNs [7][12].

F. RNN

The recurrent Neural Networks (RNNs) are a sort of neural network that is both strong and reliable. It mainly uses sequential or time series data [7].

G. Evaluating Measures

Evaluation metrics are frequently used to assess categorization performance. As a result, the performance measurements are the most common. So we have used different metrics to evaluate our classifiers given as follows:

a) Accuracy The accuracy metric examines the trained model using test samples as input. The number of samples properly identified out of the total number of samples is known as accuracy. Mathematically, the following equation is used to express accuracy.

Accuracy = Correct Prediction/Total Data Points

b) Precision Precision is a proportion to tell how exact a classifier is performing. The number of false news samples projected as false news samples is the precision

Precision P can be formulated as the ratio of total true positives to total predicted instances:

$$P = TP/(TP+FP)$$

c) Recall Predicted number of positive notions from all concepts in the original class measure is Recall. It's formula is:

$$R = TP/(TP+FN)$$

d) F-Measure It is represented as a harmonic mean of Precision and recall and can be formulated as:

$$F = 2P.R/(P+R)$$

Where,

TP(True Positives) belongs to a class it actually belongs
FP(False Positives) belongs to a class it actually doesn't belong
FN(False Negative) doesn't belong to a class it actually should belong
TN(True Negatives) doesn't belong to a class it actually doesn't belong

IV. Comparative Study Of ALL ML Techniques With Accuracy

TABLE I. COMPARATIVE STUDY OF ALL ML TECHNIQUES

Author	*Year*	*Keywords used*	*Dataset*	*ML Techniques*	*Best Accuracy*
Zubiaga, Kochkina Liakata, Procter [1]	2017	Stance classification, social media, breaking news, veracity classification	Twitter dataset associated with breaking news	Linear CRF LSTM SVM Tree CRF	73% (LSTM)
H. Ahmed, I. Traore, S. Saad [2]	2017	Online fake news, online social network security, Fake news detection, text classification, N-gram analysis	Collected datasets from various sites and merged them	Stochastic Gradient Descent Support Vector Machines K-Nearest Neighbor Decision Trees Linear Support Vector Machines	92% (Linear SVM)
Madhav Kotteti, Na Li,	2018	Fake news, machine	LIAR dataset	Gradient boosting Linear SVC	45.7% (MLP

Author	*Year*	*Keywords used*	*Dataset*	*ML Techniques*	*Best Accuracy*
Lijun Qian [3]		learning, data imputation, feature extraction		MLP Classifier Decision tree Linear SVC SVC	Classifier)
Supanya Aphiwongsophon, Prabhas Chongstitvatana[4]	2018	fake news, social network, Naïve Bayes, Neural network, Support Vector Machine (SVM)	Twitter feeds	SVM Naïve Bayes Neural Networks	99.90% (SVM and Neural Networks)
Akshay Jain, Amey Kasbe[5]	2018	Machine Learning, Naïve Bayes Classifier, Web Scrapping	Github containing 11,000 news articles	Naïve Bayes	80.60 % (on title) 91.20% (on text)
Manisha Gahirwal, Sanjana Moghe, Tanvi Kulkarni [6]	2018	Stance Detection, Natural Language Processing (NLP)	URL data	Support Vector Machine	87%
O. Ajao, D. Bhowmik, S. Zargari [7]	2019	Fake News, Twitter, social media	Five rumored stories tweets	LSTM LSTM CNN LSTM DROP	82.29 %(LSTM)
J. Reis, A. Correia, F. Murai, A. Veloso, F. Benevenuto[8]	2019	------	Buzzfeed	Naïve Bayes KNN	80%(KNN)
C Yuan, Q Ma, W Zhou, Jizhong, Songlin[9]	2020	-------	Weibo, twitter 15,16	SMAN	95.60 %(SMAN)
F. Ozbay, B. Alatas[10]	2020	Fake news detection, online social media, supervised artificial intelligence algorithm	Buzzfeed, random political, ISOT fake news	23 AI algorithms	65.50 %, 96.80 % (J48 and decision tree)
A. Kesarwani, S. Singh Chauhan, A. Ramachandran Nair [11]	2020	Fake news, K-Nearest Neighbor, data mining; supervised.	Facebook news post	K-Nearest Neihbour	79%(KNN)
Nagaraja, Soumya K N, Prajwal Naik[12]	2021	Fake news, Machine Learning, Naïve Bayes, SVM, performance	Merged four datasets	Naïve Bayes SVM	75%(SVM)

V. Conclusion And Future Work

Fake news has a huge influence on our social life, as well as in other domains, such as politics and education. Fake news may create significant social and societal harm, as well as have potentially dangerous consequences. It is becoming more difficult for the citizens/consumers to obtain information that is precise and error-free and reliable because of increasing the dimensions of social media. It is important to identify these false stories first in order to save the damage they can cause worldwide.

There has been a lot of work already to identify the false information. Many researchers applied various machine learning and deep learning techniques and calculated the accuracy but there is no research where 100% accuracy is being achieved by the researchers. In the paper by Madhav Kotteti [3] they used imputation techniques to process missing values in raw data. In the future, these processing techniques could be combined with various ML techniques to reach better results.

In this survey, we have carried out the study to illustrate the work done by researchers for identifying fake news. Our future work will focus on applying ensemble techniques and optimizing them. As most of the researchers have used just one dataset to perform the experiment, we'll try to take datasets from various domains and from various URLs to achieve better results.

References

[1] A. Zubiaga et al., "Discourse-aware rumour stance classification in social media using sequential classifiers," Information Processing & Management, vol. 54, no. 2, pp. 273–290, Mar. 2018, doi: 10.1016/j.ipm.2017.11.009.

[2] H. Ahmed, I. Traore, and S. Saad, "Detection of Online Fake News Using N-Gram Analysis and Machine Learning Techniques," in Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments, vol. 10618, I. Traore, I. Woungang, and A. Awad, Eds. Cham: Springer International Publishing, 2017, pp. 127–138. doi: 10.1007/978-3-319-69155-8_9.

[3] C. M. M. Kotteti, Xi. Dong, N. Li, and L. Qian, "Fake News Detection Enhancement with Data Imputation," in 2018 IEEE 16th

Intl Conf on Dependable, Autonomic and Secure Computing, 16th Intl Conf on Pervasive Intelligence and Computing, 4th Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress(DASC/PiCom/DataCom/CyberSciTech), Athens, Aug. 2018, pp. 187–192. doi: 10.1109/DASC/PiCom/DataCom/CyberSciTec.2018.00042..

[4] S. Aphiwongsophon and P. Chongstitvatana, "Detecting Fake News with Machine Learning Method," in 2018 15th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), Chiang Rai, Thailand, Jul. 2018, pp. 528–531. doi: 10.1109/ECTICon.2018.8620051.

[5] A. Jain and A. Kasbe, "Fake News Detection," 2018 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS), 2018, pp. 1-5, doi: 10.1109/SCEECS.2018.8546944.

[6] Gahirwal Manisha et. al, "Fake News Detection," *International Journal of Advance Research, Ideas and Innovations in Technology* ISSN: 2454-132X Impact factor: 4.295

[7] O. Ajao, D. Bhowmik, and S. Zargari, "Fake News Identification on Twitter with Hybrid CNN and RNN Models," in Proceedings of the 9th International Conference on Social Media and Society, Copenhagen Denmark, Jul. 2018, pp. 226–230. doi: 10.1145/3217804.3217917.

[8] J. C. S. Reis, A. Correia, F. Murai, A. Veloso, and F. Benevenuto, "Supervised Learning for Fake News Detection," IEEE Intell. Syst., vol. 34, no. 2, pp. 76–81, Mar. 2019, doi: 10.1109/MIS.2019.2899143.

[9] C. Yuan, Q. Ma, W. Zhou, J. Han, and S. Hu, "Early Detection of Fake News by Utilizing the Credibility of News, Publishers, and Users Based on Weakly Supervised Learning." arXiv, Dec. 13, 2020. Accessed: May 26, 2022. [Online]. Available: http://arxiv.org/abs/2012.04233

[10] F. A. Ozbay and B. Alatas, "Fake news detection within online social media using supervised artificial intelligence algorithms," Physica A: Statistical Mechanics and its Applications, vol. 540, p. 123174, Feb. 2020, doi: 10.1016/j.physa.2019.123174.

[11] A. Kesarwani, S. S. Chauhan, and A. R. Nair, "Fake News Detection on Social Media using K-Nearest Neighbor Classifier," in 2020 International Conference on Advances in Computing and Communication Engineering (ICACCE), Las Vegas, NV, USA, Jun. 2020, pp. 1–4. doi: 10.1109/ICACCE49060.2020.9154997.

[12] A. Nagaraja, S. K N, A. Sinha, J. V. Rajendra Kumar, and P. Nayak, "Fake News Detection Using Machine Learning Methods," in International Conference on Data Science, E-learning and Information Systems 2021, Ma'an Jordan, Apr. 2021, pp. 185–192. doi: 10.1145/3460620.3460753.

[13] S. M. S.-U.-R. Shifath, M. F. Khan, and M. S. Islam, "A transformer based approach for fighting COVID-19 fake news." arXiv, Jan. 28, 2021. Accessed: May 26, 2022. [Online]. Available: http://arxiv.org/abs/2101.12027

[14] E. Tacchini, G. Ballarin, M. L. Della Vedova, S. Moret, and L. de Alfaro, "Some Like it Hoax: Automated Fake News Detection in Social Networks," arXiv, arXiv:1704.07506, Apr. 2017. doi: 10.48550/arXiv.1704.07506.

[15] A. Campan, A. Cuzzocrea, and T. M. Truta, "Fighting fake news spread in online social networks: Actual trends and future research directions," Dec. 2017, pp. 4453–4457. doi: 10.1109/BigData.2017.8258484.