



A Review: Machine Learning Approach and Deep Learning Approach for Fake News Detection

Sumit Kumar¹, Jyoti Tiwari²

¹ M.E. Computer Engineering Sgsits Indore, India, bhadouriasumit27@gmail.com

² Assistant Professor Comp. Tech & Appl. Dept. Sgsits Indore, India, jyotimona23@gmail.com

ABSTRACT

Increasing number of social media platforms, emerging new technologies, and population growth which results in the rate of using social media has increased rapidly. With an increasing number of users on online platforms comes to a variety of problems like fake news. The extensive growth of fake news on social media can have a serious impact on the real world and became a cause of concern for net users and governments all over the world. Distinguishing between real news and fake news becoming more challenging. The amount of fake news has become a disguise. In this paper, we have done a survey on detection techniques for fake news using Algorithms and Deep learning techniques. We have compared machine learning algorithms like Naïve-Bayes, Decision tree, SVM, Adaboost, etc. Comparing the accuracy

Key words: Convolutional neural network, support vector machine, LSTM.

1. INTRODUCTION

Fake news has become one of the biggest threats to journalism. It has weekend people trust on News agencies and government. The method of classifying news stories as false or real is known as fake news identification. To accomplish this, it must first extract specific and important features from the article's text and then use a classifier to separate fake news from genuine news... To overcome this by using some websites, to overcome this by using some websites, tools, and platforms were introduced to detect the differentiation of real news and fake news. The fakesters create a piece of news that makes the readers' minds desperate to learn what is there in the news by seeing the title of particular news. That desperate minds are giving a chance to the fakesters to create such unhealthy news. Even though fake news detection is not completely ruling the world but a lot of interest is being generated differentiation of real news and fake news. Many researchers have worked on this topic. Some of them used Traditional classification algorithms and others used Deep learning approaches. Deep learning has the advantage of having no handcrafting of rules and/or functions, instead of

finding the best feature set for a particular problem on its own. [1].

2. MACHINE LEARNING ALGORITHM APPROACH

The aim of machine learning methods is to automate the learning process of computers based on evidence. It splits the data into training and research, with the training data being used to learn the model and the testing data is used to validate the data. It is divided into three categories: supervised, semi-supervised, and unsupervised learning. For learning the model, supervised machine learning methods use labeled results. Unlabeled data, on the other hand, is analyzed using unsupervised techniques.

2.1 Naïve bayes classifier

Naive Bayes is a probabilistic machine learning algorithm, and this is a probabilistic classifier that uses the max posteriori decision rule to render classifications. The Bayes theorem is used in classification activities, and the Naive Bayes classifier is derived from it [15]. This classifier is a versatile and common conventional algorithm for practical uses. Users get a short response from these real-world apps. This algorithm is used in email spam filtering and sentiment analysis, among other things. This is one of the most basic classification problems, and it requires large datasets. ble, approximately 8 to 12 point type.

$$P(I/J)=(P(J/I)*P(I))/P(J)$$

$$\text{Posterier}=(\text{Liklihood}*\text{prior})/\text{Evidence}$$

P(I): Prior Probability of class

P(J|I): Likelihood, the probability of predictor given class

P(J): Prior probability of predictor

P(I|J): Posterior probability off class I was given the predictor J

2.2 K-Nearest Neighbor

The supervised machine learning algorithm known as the k-nearest neighbor algorithm solves classification and regression problems. Let's take a look at some labeled points, measure the distance between them, arrange the distances in decreasing order, and mark the first point as "k." The

regression and classification models' mean and mode are returned by the K marks. This is a delicate model to deal with and the outputs of classification problems are discrete values [14]. The K nearest model identifies patterns and senses intrusions.

2.3 Support Vector Machine

A supervised learning algorithm is SVM. This is one of the best-fit algorithms as compared to other algorithms. To find false news, SVM employs computational linguistics. This algorithm is used to transform learning models that are only intended for single use. The training sets and data sets are also normalized. It has a wide variety of implementations in the qualified set. A commonly used classification algorithm is the help vector machine. Often used for collecting vast amounts of data from images with a small number of dimensions.

2.4 Decision Tree

A decision tree is a flowchart representation of a supervised learning model. It's a non-parametric approach for classification and regression that's simple and easy to use. We partition the root set into subsets using a decision tree dependent on attribute value tests. Recursively, the division process is repeated on each subset. When all of the subsets at a node have the same variable, the recursion is complete. The decision tree with decision nodes and leaf nodes is the end product. A decision node has two or three nodes, with the leaf node representing either grouping or decision. It is capable of handling both categorical and numerical results.

2.5 Random Forest

Many random decision trees make up a random forest. The trees have two forms of randomness built-in. To begin, each tree is constructed from a random sample of the original data. Second, a subset of features is randomly chosen at each tree node to produce the best break.

3. DEEP LEARNING ALGORITHM APPROACH

Deep Neural Networks (DNNs) were introduced as a way to replicate the human brain's pattern recognition abilities. A DNN is a neural network with input, output, and a single hidden layer that is made up of networks. Finds mathematical manipulations to turn on non-linear or linear interactions in the input and output.

3.1 Convolutional Neural Network

Convolutional neural networks are a network of neurons that are linked in layers that are used to take inputs and outputs. CNN is a feed-forward network model that can be used to recognize objects and analyze images. The convolution layer, detector layer, and pooling layer are the three levels that make up the neural network that performs the functions. The convolutional layer's task is to build a complicated function chart. The nonlinear components of function maps are highlighted by the detector layer. The predecessor

information is reduced and the output is provided by the pooling layer. CNN looks for dormant personalities in news stories. The data size and qualified data are the most common uses of this model. Based on efficiency and speed, the CNN model is considered.

3.2 Long-Short Term Memory

A long short-term memory is a persistent recurrent neural network (LSTM). In LSTMs, individual hidden drives are used, and their typical action is to remember inputs for a long time. A memory cell, also known as an accumulator or a gated leaky neuron, is a system that works as an accumulator or a gated leaky neuron and has a weight of 1. In other words, it repeats itself.

3.3 Feedforward Neural Network

Feedforward NNs are the first and, in many ways, the most basic kind of artificial neural network. In this network, data only flows in one direction: forward from the input nodes to the output nodes, passing through any hidden nodes (if any). In the network, there are no cycles or loops. Feedforward networks (also known as associative networks) are made up of various categories of components, such as binary McCulloch–Pitts neurons (Anderson, 2006), with the perceptron becoming the most basic. In the case of backpropagation of errors, continuous neurons with sigmoid activation functions are commonly used.

3.4 Bidirectional LSTM concatenated model

Using Google's pre-trained model, the headline and article texts are translated into two different embedding layers. The two embedding vectors' results are concatenated to feed the construct. The model is made up of two CNN layers, each with 32 and 64 filters. Until moving it via a Bidirectional LSTM layer with 100 memory units, the CNN's are accompanied by max-pooling to prevent over-fitting⁵. The bidirectional LSTM's output was routed through three dense layers (512, 128, 4 units) separated by a single dropout.

4. LITERATURE REVIEW

In the field of machine learning, a large number of experiments have been performed for the identification of false news. B. M. Amine, A. Drif and S. Giordano [2] the approach that is taken to solve the problem is through a convolutional neural network. Compared to other methods based on CNN where a single CNN-based model is constructed using a text-only approach means only text metadata is used. But using more than text metadata by using another CNN show significant improvement for fake news detection.

R. K. Kaliyar et al. [3] the dataset they have used is a multiclass fake news dataset which is taken from the Fake

news challenge-based fake news dataset. Benchmark findings are available in the latest analysis focused on two groups' datasets, classifying news as false or real? There is minimal literature on multiclass estimation. There were many contents and context-related issues with the binary class dataset. So, they have used a multiclass dataset. They utilized gradient boosting with an advanced parameter that fits best for multi-class fake news datasets.

J. Lin et al. [4] have built a Long-term short-term memory (LSTM) based machine learning model with the attention-based model as LSTM suffers from the issue of long-range dependence and the encoder-decoder network output degrades rapidly as the sentence length increases. To counter the issue attention-based LSTM is used. Other machine learning models such as Random Forest, XGBoost, etc. are used as the performance measure.

Ye-Chan-Ahn, Chang-Sung Jeong [6] a method is designed to counter fake news using advantage and disadvantage of LSTM and CNN as LSTM takes a longer time to learn and is vulnerable to parallelism and CNN which require a lot of computation to connect two input and output that are located far away from position. So, they have a transformer network that uses both self-attention and feed-forward mechanisms.

Liang Wu and Huan Liu [7] States that the classification of messages was wont to find the messages spreading in social media which are not appropriate, trustworthy, etc. And proposed Trace Miner to classify messages in social media, with this an end-to-end LSTM-RNN's Classification model was also included. Trace Miner has given good results on the insufficient datasets. The longer-term work goes to be done on network mining tasks using Trace Miner.

Ahmed, Hadeer, Issa Traore, and Sherif Saad [8] two different attribute extraction techniques and six different classification techniques for machine learning have been investigated and researched. Compared to state-of-the-art techniques, they need to illustrate the efficiency of certain models and demonstrate better performance. The most significant aim of this research article was to automatically identify fake news material. For their study, they need to recognize false feedback and fake news results.

Table 1 :Work Done In Deep Learning Classification models.

Algorithms	CNN	LSTM	Feedforward neural network	Bidirectional LSTM Concatenated model
------------	-----	------	----------------------------	---------------------------------------

Authors	B. M. Amine, A. Drif and S. Giordano [2]	J. Lin, G. Tremblay-Taylor, G. Mou, D. You and K. Lee [4]	V. M. Krešňáková, M. Sarnovský and P. Butka [9]	E. Qawasmeh, M. Tawalbeh and M. Abdullah [10]
Paper	Merging deep learning model for fake news detection	Detecting Fake News Articles	Deep learning methods for Fake News detection	Automatic Identification of Fake News Using Deep Learning
Metrics	ACC:96% PREC:98% REC:94%	ACC:91% PREC:90% REC:93%	ACC:91% PREC:85% REC:99%	ACC:85% PREC:77% REC:75%
Parameter	Accuracy, Precision, Recall	Accuracy, Precision, Recall	Accuracy, Precision, Recall	Accuracy, Precision, Recall
Advantages	At once multiple classifications are done, locates objects in an image	They could solve the problem of vanishing gradients.	Can represent more complex functions very easily, It will learn more complex functions.	Solves the problem of fixed sequence to sequence prediction.
Disadvantages	Computational cost is expensive, require lot of training data	require a lot of resources and time to get trained	Prone to overfitting, Vanishing and Exploding gradient problem.	Since BiLSTM has double LSTM cells so it is costly

Table 2 Work Done in Machine Learning Classification Models.

Algorithms	Naives-Bayes	SVM	KN	Decision Tree	Random Forest
Authors	M. Granik and V. Mesyura [12]	P. B. P. Reddy, M. P. K. Reddy, G. V. M. Reddy and K. M. Mehata [13]	J. Lin, G. Tremblay-Taylor, G. Mou, D. You and K. Lee [4]	S. Lyu and D. C. -T. Lo [11]	M. P. K. Reddy, K. M. Mehata, P. B. P. Reddy and G. V. M. Reddy [13]
Paper	Fake news detection using naive Bayes classifier	Fake Data Analysis and Detection Using Ensemble Hybrid Algorithm	Detecting Fake News Articles	Fake News Detection by Decision Tree	Fake Data Analysis and Detection Using Ensemble Hybrid Algorithm
Metrics	ACC:75.40% PREC:71% REC:73%	ACC:89% PREC:85% REC:93%	ACC:79% PREC:75% REC:79%	ACC:95% PREC:92% REC:90%	ACC:86% PREC:83% REC:89%
Parameter	Accuracy, Precision, Recall	Accuracy, Precision, Recall	Accuracy, Precision, Recall	Accuracy, Precision, Recall	Accuracy, Precision, Recall

Advantages	Easily implemented, requires less training period to test the data, fast decision making.	Performant, not biased by outliers, not sensitive to over fitting.	Simple Implementation, classes don't have to be Linearly separable, Robust with regard to search space.	Interpretability, no need for feature scaling, works on both linear / non – linear problems.	Powerful and accurate, good performance on many problems, including non – linear.
Disadvantages	Loss of accuracy	Not appropriate for non-linear problems, not the best choice for large number of features.	Expensive testing of each Instance, Sensitivity to noise, Sensitivity to very Unbalanced data.	Poor results on very small datasets, over fitting can easily occur.	No interpretability, over fitting can easily occur, need to choose the number of trees manually.

5. ACKNOWLEDGMENT

My thanks to the Jyoti Tiwari mam who have contributed towards the development of the paper.

6. CONCLUSION

A comparison of different methods identified by various researchers to deal with the issue of fake news identification has been done in this article. The challenge of detecting fake news in this age is causing a lot of problems. There has been a lot of studies undertaken to find ways to prevent and reduce the issue of fake news. Various methods have been outlined in this paper to do.

REFERENCES

- [1] A. Roy, K. Basak, A. Ekbal and P. Bhattacharyya, "**A Deep Ensemble Framework for Fake News detection and classification**," arXiv preprint arXiv:1811.04670, 2018.
- [2] B. M. Amine, A. Drif and S. Giordano, "**Merging deep learning model for fake news detection**," 2019 International Conference on Advanced Electrical Engineering (ICAEE), Algiers, Algeria, 2019, pp. 1-4, doi: 10.1109/ICAEE47123.2019.9015097.
- [3] R. K. Kaliyar, A. Goswami and P. Narang, "**Multiclass Fake News Detection using Ensemble Machine Learning**," 2019 IEEE 9th International Conference on Advanced Computing (IACC), Tiruchirappalli, India, 2019, pp. 103-107, doi: 10.1109/IACC48062.2019.8971579.
- [4] J. Lin, G. Tremblay-Taylor, G. Mou, D. You and K. Lee, "**Detecting Fake News Articles**," 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 2019, pp. 3021-3025, doi: 10.1109/BigData47090.2019.9005980.
- [5] P. Shah and Z. Kobti, "**Multimodal fake news detection using a Cultural Algorithm with situational and normative knowledge**," 2020 International Conference on Advanced Computing (IACC), Glasgow, United Kingdom, 2020, pp. 1-7, doi: 10.1109/CEC48606.2020.9185643.
- [6] Y. Ahn and C. Jeong, "**Natural Language Contents Evaluation System for Detecting Fake News using Deep Learning**," 2019 16th International Joint Conference on Computer Science and Software Engineering (JCSSE), Chonburi, Thailand, 2019, pp. 289-292, doi: 10.1109/JCSSE.2019.8864171.
- [7] Wu, L., & Liu, H. (2018). "**Tracing Fake-News Footprints**". Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining - WSDM '18. doi:10.1145/3159652.3159677.
- [8] Ahmed, Hadeer, Issa Traore, and Sherif Saad. "**Detecting opinion spams and fake news using text classification**." Security and Privacy 1, no. 1 (2018): e9
- [9] V. M. Krešňáková, M. Sarnovský and P. Butka, "**Deep learning methods for Fake News detection**," 2019 IEEE 19th International Symposium on Computational Intelligence and Informatics and 7th IEEE International Conference on Recent Achievements in Mechatronics, Automation, Computer Sciences and Robotics (CINTI-MACRo), Szeged, Hungary, 2019, pp. 000143-000148, doi: 10.1109/CINTI-MACRo49179.2019.9105317.
- [10] E. Qawasmeh, M. Tawalbeh and M. Abdullah, "**Automatic Identification of Fake News Using Deep Learning**," 2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS), Granada, Spain, 2019, pp. 383-388, doi: 10.1109/SNAMS.2019.8931873.
- [11] S. Lyu and D. C. -T. Lo, "**Fake News Detection by Decision Tree**," 2020 SoutheastCon, Raleigh, NC, USA, 2020, pp. 1-2, doi: 10.1109/SoutheastCon44009.2020.9249688.
- [12] M. Granik and V. Mesyura, "**Fake news detection using naive Bayes classifier**," 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), Kyiv, Ukraine, 2017, pp. 900-903, doi: 10.1109/UKRCON.2017.8100379.
- [13] P. B. P. Reddy, M. P. K. Reddy, G. V. M. Reddy and K. M. Mehata, "**Fake Data Analysis and Detection Using Ensembled Hybrid Algorithm**," 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2019, pp. 890-897, doi: 10.1109/ICCMC.2019.8819741.
- [14] A. Kesarwani, S. S. Chauhan and A. R. Nair, "**Fake News Detection on Social Media using K-Nearest Neighbor Classifier**," 2020 International Conference on Advances in Computing and Communication Engineering (ICACCE), Las Vegas, NV, USA, 2020, pp. 1-4, doi: 10.1109/ICACCE49060.2020.9154997.
- [15] Anurag Busha 1 , Manoj Prakash.P 2 , Pelash Choudhary 3 , Vakeesh Kanna. T. I 4 "**Sentimental Analysis**" International Journal of Emerging Technologies in Engineering Research (IJETER) Volume 6, Issue 4, April (2018).