



## Application of Machine Learning Algorithms for Fake News Detection: A Review

Mohammed Ajuji<sup>1</sup>, Raihanatu Mohammed Hamid<sup>2</sup>, Datti Useni Emmanuel<sup>3</sup>, Abdulqadir Maigari Turaki<sup>4</sup>, Amina Nuhu Muhammad<sup>1</sup>

<sup>1</sup>Department of Computer Science, Faculty of Science, Gombe State University, Gombe 760214, Nigeria

<sup>2</sup>Department of Computer Science, School of Science, College of Education, Billiri, Gombe State, Nigeria

<sup>3</sup>Department of Computer Science, School of Science, Federal College of Education, Plateau, Nigeria

<sup>4</sup>Department of Computer Science, School of Science, Federal Polytechnic Bauchi, Nigeria

### ABSTRACT

Before the advent of the internet and information and communication technology. News and/or information are broadcast via television, newspapers, and magazines are also known as traditional media. The recent advances of communication technologies have considerably transcended the chronological and spatial boundaries of traditional ways of disseminating, sharing, and spreading news especially in terms of speed and ease of access. These technologies which bring people together irrespective of geographical location have created a revolution in information generation and dissemination; online people interaction, human behaviour-related data, and even real-time sharing of data. However, the abuse of social technologies such as social media platforms nowadays has become a source of concern to many stakeholders. It has given some people ample opportunity to share false, fabricated, deceitful, and hoax information also known as fake news online. In this paper, we systematically reviewed articles that applied machine learning techniques to detect fake news on social media. Various machine learning techniques were applied with the most being Naïve Bayes. Decision Tree, Support Vector Machine, and Random Forest are commonly applied. Some techniques are least utilized whereas others like Ridge Regression, Lasso Linear Regression, and Elastic Net Regression have not been explored.

### ARTICLE INFO

#### Article History

Received: August, 2021

Received in revised form: August, 2021

Accepted: September, 2021

Published online: December, 2021

### KEYWORDS

Fake news, machine learning, algorithm, social media, supervised learning

### INTRODUCTION

Social media platforms nowadays have become more popular than ever before. The use of these platforms has transcendently increased in recent times, millions if not billions of data are being generated and updated daily (Al-garadi et al., 2019). Social media has become an important data source for learning about breaking news (Zubiaga, 2019) and for following the latest trends and developments of ongoing news (Bondielli & Marcelloni, 2019). However, this has given some people ample opportunity to misuse it by sharing online; falsehood, fabricated, and deceitful information which is also known as fake news. Fake news has now grown into a big societal problem and also a major challenge for people fighting disinformation. The easy dissemination of

information by way of sharing has added to an exponential growth of its falsification. The credibility of social media contents is also at stake where the spreading of fake information is prevalent (Manzoor, 2019). Many attempts have been made and are still going on to understand, detect, and reduce if not completely eradicate the spread of fake news, but with the advent of machine learning algorithms that make it possible for machines to learn fast, detect, and predict, the field became most sought for. Many studies have been conducted that focused on fake news detection through the application of machine learning classifiers, therefore, this study is aimed at making a systematic review of the findings of those studies. (Cardoso et al., 2019) concluded that machine learning is the most used method for



automatic fake news detection. It is an essential for efficient and accurate fake news classification (Jiang & Li, 2021).

The remainder of this survey is organized as follows: Section II provides an overview of the method used to gather relevant literature. Then, Section III provides a review of Machine Learning-based approaches. An evaluation criterion is shown in section IV. Section V presents some concerns about the creation of fake news detection models. And lastly, the conclusion is drawn in section VI.

### **Rise of Citizen Journalism**

The availability of smart mobile devices has made it possible for all to be self-acclaimed journalists, it allows anyone with access to the Internet to post updates from anywhere, leading to a growing presence of citizen journalism (Zubiaga, 2019). Social media has become a go-to resource for journalists during the process of news gathering (Orellana-rodriguez & Keane, 2018). Individuals share a post, and tweet information with the people in their network without checking the veracity of the posts (Shelke & Attar, 2019), which leads to the issue of rumour diffusion in a social network. These need to be properly addressed. Therefore, machine learning models could do perfect work if correctly employed to detect and classify news as either fake or real.

### **Escalation of Fake News on Social Media**

Social interaction evolved from limited cultural boundaries, such as political, religious, and families in a locality (Zhou & Zafarani, 2020). Recently, with the development of information and communication technologies which has transcendentally transcends the temporal and spatial limitations of traditional means of disseminating information. The term fake news is used to define news content that is false, misleading, fabricated and the sources of the quoted statement of the news content cannot be verified (Han & Mehta, 2019) and (Preparation, 2019). It is easy to access, less time-consuming, user-friendly, easily conveyable socially relevant news, the possibility for obtaining the various perspectives of single news, and is being updated every minute (Classifier & Kesarwani, 2020) and (Kaliyar, 2018). The easy dissemination of

information by way of sharing has added to the exponential growth of its falsification. The credibility of social media content is also at stake where the spreading of fake information is prevalent (Manzoor, 2019). According to (Manzoor, 2019) Various forms of fake news are; visual based, user-based, knowledge-based, style-based, and stance-based. Social media platforms are one of the fastest mediums to spread information, by creating a huge impact on manipulating information which can influence readers in positive and negative aspects (Han & Mehta, 2019).

### **Fake News Detection**

Automatic classification of a text article as real or fake is a challenging task (Ahmad et al., 2020). News is being manipulated by various networking sites based on private opinions or interests. Fake news is misinformation or manipulated news that is spread across social media to damage the reputation of a person, group, organization, or political party. Because of the volume of fake news being circulated on the media in recent times, there is a need for computational methods to detect them. Fake news detection aims to help users to expose varieties of fabricated news. One can decide whether the news is true or false based on formerly witnessed fake or real news. We can use various models to access deceptive news in social media (Kaliyar, 2018). We need to be smarter at the recognition of fake or real news. A human being cannot be effective to detect all fake news. Fact-checking by a human is very laborious and expensive (Jiang & Li, 2021). So there is an urgent need to come up with a system that can automatically detect and classify news as real or fake; this could be achieved especially using machine learning classifiers (Jiang & Li, 2021).

### **METHODOLOGY**

This section presents the methodology used in this study for searching for relevant and related literature. In retrieving the relevant published papers on fake news detection models, two phases were employed. The first phase included searching for reputable academic databases and search engines. Several databases and searches are available but the



academic databases used for retrieval of relevant paper publication were as follows: DBLP Computer Science Bibliography, ACM Digital Library, ScienceDirect, SpringerLink, and IEEE Xplore. The keywords used for the literature search were coined concerning fake news as follows: fake news, survey, review, fake news detection, and fake news machine learning models. Secondly, the quality of the selected articles was ensured by excluding all the journals and conferences that were informally published. However, it should be noted that we only consider papers with a novel model, implementing an existing model would not be analyzed in this review. At the same time, the timeliness of the survey was provided by focusing on publications from the last three years.

The articles retrieved from the literature were scrutinized to ensure that the articles used met the established inclusion criteria. The criteria for an article to be included for this survey are: it must report an empirical study describing fake news detection on social media sites. Else the article would be rejected based on titles. The abstract and conclusion sections were examined to ensure that articles satisfied the screening criteria, and those that did not satisfactorily meet the designed criteria were excluded from the survey.

### MACHINE LEARNING

It is a division of Artificial Intelligence (AI) techniques that enables computer systems to learn from previous experience/practice and improve their behaviour for a given task (Nguyen et al., 2020). Machine learning techniques are twofold: supervised and unsupervised machine learning techniques. The supervised technique utilizes the training data to predict hidden future activities. Unsupervised techniques are leveraged to recognize hidden data models provided without providing training data (Khan, 2019). (Irda, 2019), (Mahid et al., 2018) Machine learning

methods have become an essential requirement for automated fake news detection. Machine learning can detect fake news accurately (Skizrqjvrskrq et al., 2018). Several machine learning algorithms, especially supervised learning that has been used by different researchers are stated below.

### Machine Learning Algorithms

Different types of machine learning algorithms exist, but almost all studies regarding fake news detection on social media used the most established and widely used type, that is, regression or supervised machine learning algorithms (Al-garadi et al., 2019). The accomplishment of the algorithms is determined by the degree to which the model accurately predicts or converts various types of prior knowledge about the task. Most of the practical machine learning algorithms consider the details of the particular problem at hand, then a model that could perform accurately is selected. Though, no machine learning algorithm can ultimately solve all problems (Cardoso et al., 2019), (Meel & Vishwakarma, 2019). As a result, researchers resort to trial and error by selecting many algorithms, building models, performing a comprehensive practical experiment and comparing their results to determine the most ideal or suitable algorithms for the detection of fake news (Oshikawa et al., 2018). Table 1 summarizes the commonly used machine learning algorithms for constructing fake news detection models. Machine Learning techniques include Support Vector Machines (SVM), decision trees, Bayes learning, k-means clustering, k-nearest neighbor, neural networks, and many more algorithms.

The section below describes the machine learning algorithms frequently used for constructing fake news detection models.

**Table 1:** Summary of machine learning algorithms tested in fake news detection literature

Study	LR	KNN	DT	SVM	NB	NN	RF	GB	XGB
[20]				Y	Y	Y			
[23]	Y		Y	Y	Y				
[13]		Y	Y		Y		Y		
[12]		Y							

Corresponding author: Ajuji, M. ✉ [majui@qsu.edu.ng](mailto:majui@qsu.edu.ng) ✉ Department of Computer Science, Faculty of Science, Gombe State University, Gombe. © 2021. Faculty of Tech. Education, ATBU Bauchi. All rights reserved



Study	LR	KNN	DT	SVM	NB	NN	RF	GB	XGB
[10]					Y				
[25]	Y		Y	Y			Y	Y	Y
[24]		Y	Y	Y	Y		Y		Y
[6]	Y	Y	Y	Y			Y		
[27]					Y				
[26]						Y	Y		

\*\*LR=Linear regression, KNN=K-Nearest Neighbor, DT=Decision tree family, SVM=Support Vector Machine, NB=Naïve Bayes, NN=Neural Network, RF=Random Forest, GB=Gradient Boosting, XGB=eXtreme Gradient Boosting, CNN=Convolutional Neural Network, GNU=Gated Recurrent Network, LSTM=Long Short-Term Memory.

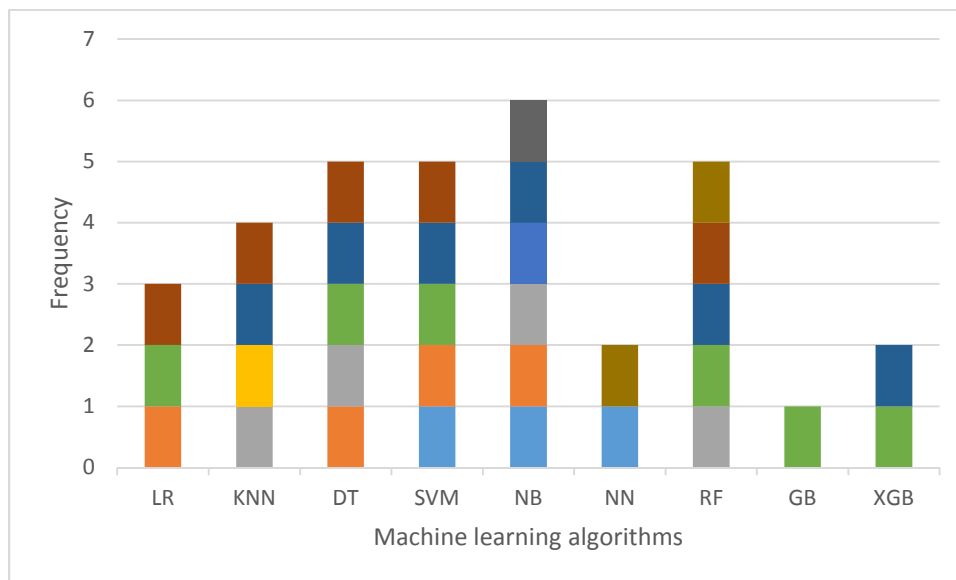


Figure 1. Machine learning algorithms applied in fake news detection.

### Linear Regression

It was developed in the field of statistics and borrowed by machine learning algorithms, it is now both a machine learning and statistical algorithm. It is a model for understanding the relationship between input and output numerical variables (Poddar & D, 2019). Linear regression is twofold, that is; simple linear regression which is a statistical method that enables users to summarize and study relationships between two continuous or quantitative variables.

$$y = B_0 + B_1 * x$$

Where  $y$  is the dependent variable (predicted result),  $x$  is an independent variable,

$B_1$  is the slope of the line and  $B_0$  is the  $y$ -intercept.

Linear regression is a linear model wherein a model assumes a linear relationship between the input variables ( $x$ ) and the single output variable ( $y$ ). When there is a single input variable ( $x$ ), the method is called a simple linear regression, otherwise, the procedure is referred to as multiple linear regression.

### K-Nearest Neighbors

The model representation for KNN is the entire dataset. KNN makes a prediction using the training dataset directly. Prediction is made for a new data point by searching through the entire training set for the  $K$  most similar instances (the



neighbours) and summarizing the output variable for those K instances. For this regression study, this might be the mean output variable. KNN can be used for both regression and classification problems.

To determine which of the K instances in the training dataset are most similar to the new input a distance measure is used. The most popular distance measure is Euclidean distance, and it is used here because it deals with real-valued input variables. Euclidean distance is calculated as the squared differences between points a and b across all input attributes  $i$ .

$$\begin{aligned} \text{Euclidean Distance } (a, b) \\ = \sqrt{\sum_{i=1}^n (a_i - b_i)^2} \end{aligned} \quad 2$$

There is much another distance measure available but Euclidean distance is used because the input variables are similar in type. For regression problems, the KNN prediction is based on the mean and the median of the K-most similar instances.

### Decision Trees

Classification and Regression Trees or CART for short is a term introduced first by Leo Breiman to refer to Decision Tree Algorithms that can be used for both classification and regression problems. The DT model is represented using a binary tree. Each node represents a single input variable ( $x$ ) and a split point on that variable (assuming the variable is numeric). The leaf nodes of the tree contain an output variable ( $y$ ) which is used to make a prediction. Predictions are made by the decision tree by traversing the binary tree given a new record (Series & Science, 2021). The tree is learned using a greedy algorithm on the training data to pick splits in the tree. Stopping criteria define how much a tree learns and pruning can be used to improve generalization on a learned tree. The decision outperformed all other algorithms tested by (Jiang & Li, 2021).

### Support Vector Machine (SVM)

Support vector machine (SVM) is a supervised machine learning algorithm that is usually used in text classification (Manzoor, 2019). (Al-garadi et al., 2019) SVM can best be explained in practice by using a hypothetical classifier called

Maximal-Margin Classifier. The numerical input variables ( $x$ ) in the data (the columns) form an  $n$ -dimensional space. A line that splits the input variable space is known as a hyperplane. A hyperplane is selected in SVM to best separate the points in the input variable space by their class, either class 0 or class 1. In  $n$ -dimension say two, you can visualize this as a line and assume that all our input points can be separated by this line. For example:

$$B_0 + (B_1 * X_1) + (B_2 * X_2) = 0 \quad 3$$

Where the coefficients ( $B_1$  and  $B_2$ ) that determine the slope of the line and the intercept ( $B_0$ ) is found by the learning algorithm, and  $X_1$  and  $X_2$  are the two input variables. You can make classifications using this line. By plugging in input values into the line equation, you can calculate whether a new point is above or below the line.

Real in practice can be messy and cannot be separated accurately with a hyperplane. The constraint of maximizing the margin of the line that separates the classes must be relaxed. This is often called the soft margin classifier. This change allows some points in the training data to violate the separating line. An additional set of coefficients are introduced that gives the margin wiggle room in each dimension. These coefficients are sometimes called slack variables. This increases the complexity of the model as there are more parameters for the model to fit the data to provide this complexity. (Smitha, 2020) compare different machine learning classifiers for fake news detection with SVM linear classification algorithm having the highest accuracy.

It is characterized by the usage of kernels, the sparseness of the solution, and the capacity control gained by acting on the margin, or several support vectors, etc. The capacity of the system is controlled by parameters that do not depend on the dimensionality of feature space. SVM algorithm operates natively on numeric attributes, as a result, it uses a z-score normalization on numeric attributes. In regression, Support Vector Machines algorithms use epsilon-insensitivity (margin of tolerance) loss function to solve regression problems.



### **Random Forest**

The idea of building many decision tree algorithms gave birth to Random Forest, after which the decision trees get a separate result (Series & Science, 2021). The outcome of prediction by a large number of the decision tree is taken up by the random forest. The random forest randomly selects a subcategory of properties from each group to ensure a variation of the decision tree. The random forest classifier was used in the construction of fake news detection models in (Cardoso et al., 2019) and (Irda, 2019). It is a machine-learning model that combines decision trees and ensemble learning (Ahmad et al., 2020), (Herrero et al., 2020). This model built several classification trees to a dataset then combines the detection from all the trees. As a result, it consists of many trees that are used randomly to select feature variables for the classifier input. Random Forest construction is achieved in steps listed in (Al-garadi et al., 2019).

### **Naïve Bayes**

Naive Bayes is a surprisingly powerful algorithm for predictive modelling (Al & Ahmed, n.d.) (Al-garadi et al., 2019). Naive Bayes is a very simple classification algorithm that makes some strong assumptions about the independence of each input variable. Nevertheless, it is effective in a large number of problem domains (Cardoso et al., 2019). Mathematically defined as (Devi et al., 2019).

$$p\left(\frac{m}{n}\right) = \frac{p\left(\frac{n}{m}\right) - p(m)}{p(n)} \quad 4$$

Where;

- ✓  $p\left(\frac{m}{n}\right)$  is the probability of hypothesis M when we have already witnessed data N, which is also called the posterior probability.
- ✓  $p\left(\frac{n}{m}\right)$  is the probability of data N when we know that hypothesis M is true.
- ✓  $p(m)$  is the probability of hypothesis M being true, which is calculated before new data is collected.
- ✓  $p(n)$  is the probability of the input data.

### **K-Nearest Neighbor**

KNN falls in the category of supervised learning (Sharma et al., 2019) and its main applications are intrusion detection, pattern recognition. It is nonparametric, so no specific distribution is assigned to the data or any assumption is made about them (Series & Science, 2021) (Hirlekar, 2020).

### **Neural Network**

It is made up of processing elements called neurons that are highly interconnected and work together to solve target problems (Poddar & D, 2019). Neural networks have a wide range of applications, usually related to pattern recognition, and they can be used to identify, evaluate, monitor, optimize, recognize, classify and filter patterns. Neural networks process input signals and produce outputs as 'black models': that is, their users do not necessarily need to understand the process taking place in the network.

The key characteristics of neural networks include their parallel computation capability and their capacity for generalization; that is, their capacity to yield a reasonable output when encountering an input never experienced before that is at the training phase. These properties allow for the approximation of very complex functions and even intractable problems.

### **EVALUATION**

The generalization of the training datasets is the main goal of building a detection model using machine learning techniques (Ahmed et al., 2019). Machine learning models should be able to perform pretty well on real data. The dataset is divided into two categories; training data and testing data. Training data is used to train machine learning classifiers whereas testing data is used to test machine learning classifiers. Some techniques can be applied for model evaluation (Nguyen et al., 2020), these include cross-validation, k-fold, holdout with various metrics such as accuracy, precision, recall, F1, Matthews correlation coefficient, receiver operating characteristic, area under the curve, mean absolute error, mean squared error, and root-mean-square error.



### Evaluation Metrics

These are tools that are used to measure the effectiveness of the proposed model, to determine whether the built model can accurately make the required prediction. Many of these tools are in existence but the most commonly used by researchers for future prediction are; accuracy, precision, recall, and f1-score.

To evaluate the performance of each model; accuracy, precision, recall, and f1 scores will be used, which are defined as follows: These were also used to evaluate cryptocurrency price prediction models (Chevallier et al., 2021). They are calculated as follows:

#### Accuracy, Precision, Recall, and F-Score

These evaluation metrics were used to evaluate fake news detection models in (Poddar & D, 2019), (Ahmad et al., 2020), and (Ghafari et al., 2020). They are calculated as follows:

**Accuracy** is simply defined as the measure of the ratio of all testing samples which is classified as correct.

$$Accuracy = \frac{tp + tn}{tp + tn + fp + fn} \quad 5$$

**Precision** means the ratio of relevant classified samples among the total retrieved samples.

$$Precision = \frac{tp}{tp + fp} \quad 6$$

**Recall** is defined as the ratio of relevant classified samples among the total amount of relevant samples.

$$Recall = \frac{tp}{tp + fn} \quad 7$$

**F1-Score** simply means the harmonic average of the precision and recall.

$$F1 - Score = 2 \left( \frac{precision * recall}{precision + recall} \right) \quad 8$$

Where the meaning of the parameters is as follows:

TP stands for the number of True Positive values  
TN stands for the number of True Negative values  
FP stands for the number of False Positive values

FN stands for the number of False Negative values

### CONCERNS ABOUT THE CREATION OF FAKE NEWS DETECTION MODELS

The major concern regarding the creation of fake news detection models are; fake news definition, data collection, and feature engineering are discussed below.

#### Fake News Definition

The term "fake news" has become popular in recent times. Though there is no generally accepted definition of fake news (Mahid et al., 2018), it keeps evolving as research reveals. Traditional fake news is generally defined as intentional behaviour to harm a person or group, repetitively, which could be difficult for the victim to defend himself or herself (Meel & Vishwakarma, 2019). By extending the definition of traditional fake news, it is the use of information technology platforms to communicate wrong information about an individual or group, either intentional or otherwise. According to (Al & Ahmed, n.d.) Fake news can be propaganda against an individual, society, organization, or political party. (Bondielli & Marcelloni, 2019) defined fake news as misinformation that is being circulated. It is also referred to as fabricated articles presented as news, that convey deliberate misinformation or hoaxes, that are designed to either attract the attention of the audience or mislead them (Orellana-rodriguez & Keane, 2018). In whatever way, it is unacceptable, in-human behaviour that is uncalled for. Humanely rejected.

#### Data Collection

Many studies on fake news detection extracted their datasets by using some particular keywords. Nonetheless, by simply tracking posts that have particular keywords, these researchers may have presented potential sampling bias (Parikh & Atrey, 2018). Such a bias would consequently introduce an over-fitting problem that affects the capability of a machine learning model to accurately make correct detections or predictions on untrained data. Limited the detection to posts that contain the predefined keywords, and overlooked many other posts relevant to fake news. Such a type of data



collection methodology limits the detection model of fake news to specified keywords. The identification of keywords for extracting posts is also subject to the author's understanding of Fake news (Mahid et al., 2018). An effective method should use a complete range of posts indicating fake news to train the machine learning classifier and ensure the generalization capability of the fake news detection model. An important goal of the machine learning approach is to generalize and not to limit the examples in a training dataset (Kaliyar, 2018). As a result, there is a need for researchers to carefully investigate whether the sampled data are extracted from data that effectively represents all possible activities on social media platforms.

### **Feature Engineering**

The effectiveness of a machine learning classifier lies in classy features. It is a vital component in improving the effectiveness of machine learning models (Ning et al., 2019). Most of the studies reviewed tried to deliver effective machine learning model solutions for the detection of fake news, especially on social media websites by providing significant features. However, these studies might have forgotten some important features. For example, those who share news online may change the way they use words and acronyms. Some websites are designed to publish fake news. Therefore, maintaining precise and accurate processes in building machine learning models from beginning to end is important in ensuring that the proposed features have significance in successful classifier performance.

### **Machine Learning Algorithm Selection**

Most studies on the machine learning approaches implemented for detecting fake news have employed a supervised machine learning technique (Bondielli & Marcelloni, 2019). In most cases, two or more machine learning algorithms should be selected and tested to determine the best machine learning algorithm for a specific dataset. The machine learning algorithms selected are to be trained on wished-for features. There is no single algorithm that solves all problems (Meel & Vishwakarma, 2019). Therefore, deciding which algorithm performs best for a specific dataset is difficult. As suggested in

(Al-garadi et al., 2019), the selection of machine learning algorithms to be tested could be done in three stages. First, specific literature on machine learning for fake news detection is important in selecting a specified classifier. Therefore, general previous research and findings on machine learning can be used as a guide to select a machine learning algorithm. Second, a text mining literature review. Lastly, a performance evaluation of comparison of comprehensive datasets can be used as the basis to select an algorithm. The stages could serve as a guide to researchers to select classifiers, nonetheless, scholars need to test many machine learning algorithms to identify the optimal classifier for an accurate machine-learning fake news detection model.

### **CONCLUSION**

This study reviewed existing literature that applied machine learning algorithms to detect fake news on social media websites. Specifically, we reviewed some aspects of detecting fake news messages by using machine learning approaches, these include, data collection, feature engineering, construction of fake news detection models, and evaluation of constructed fake news detection models. Several types of discriminative features that were used to detect fake news in online social networking sites were also summarized. In addition, the most effective supervised machine learning classifiers for classifying fake news messages in online social media platforms were identified. One of the main contributions of the current paper is the definition of evaluation metrics to successfully identify the significant parameter so the various machine learning algorithms can be evaluated against each other. Most importantly we summarized and identified the important factors for detecting fake news through machine learning techniques particularly supervised machine learning. For this purpose, we have used accuracy, precision, recall, and f1-score for modelling the behaviours in fake news.

### **REFERENCES**

- Ahmad, I., Yousaf, M., Yousaf, S., & Ahmad, M. O. (2020). *Fake News Detection Using Machine Learning Ensemble Methods*. 2020.
- Ahmed, S., Hinkelmann, K., & Corradini, F.





- (2019). *Combining Machine Learning with Knowledge Engineering to detect Fake News in Social Networks-a survey Data-Driven : Knowledge Engineering :*
- Al-garadi, M. A. L. I., Hussain, M. R., Khan, N., Murtaza, G., Nweke, H. F., Ali, I., Mujtaba, G., Chiroma, H., Khattak, H. A. L. I., & Gani, A. (2019). Predicting Cyberbullying on Social Media in the Big Data Era Using Machine Learning Algorithms : Review of Literature and Open Challenges. *IEEE Access*, 7, 70701–70718. <https://doi.org/10.1109/ACCESS.2019.2918354>
- Al, A., & Ahmed, A. (n.d.). *Detecting Fake News using Machine Learning : A Systematic Literature Review.*
- Bondielli, A., & Marcelloni, F. (2019). *A survey on fake news and rumour detection techniques.* 497, 38–55. <https://doi.org/10.1016/j.ins.2019.05.035>
- Cardoso, F., Cristina, A., & Garcia, B. (2019). *Can Machines Learn to Detect Fake News ? A Survey Focused on Social Media.* 6, 2763–2770.
- Chevallier, J., Guégan, D., & Goutte, S. (2021). *Is It Possible to Forecast the Price of Bitcoin ?* 377–420.
- Classifier, K. N., & Kesarwani, A. (2020). *Fake News Detection on Social Media using.* 0–3.
- Devi, B. L., Soni, A., Kapkoti, S. S., & Shankar, S. (2019). *Fake News Detection Based on Machine Learning by using TFIDF.* 9(11), 23969–23971.
- Ghafari, S. M., Beheshti, A., Joshi, A., Paris, C., Mahmood, A., Yakhchi, S., Orgun, M. A., & Member, S. (2020). *A Survey on Trust Prediction in Online Social Networks.* 8.
- Han, W., & Mehta, V. (2019). *Fake News Detection in Social Networks Using Machine Learning and Deep Learning : Performance Evaluation.* *Icii*, 375–380. <https://doi.org/10.1109/ICII.2019.00070>
- Herrero, Á., Cambra, C., Urda, D., Sedano, J., & Quintián, H. (2020). *Advances in Intelligent Systems and Computing 1267 13th International Conference on Computational Intelligence in Security for Information Systems ( CISIS 2020 )* (Issue Cisis).
- Hirlekar, V. V. (2020). *Natural Language Processing based Online Fake.* *Icces*, 748–754.
- Irda, Y. L. (2019). *Automatic detection of fake news on online platforms : A survey.* 3–6.
- Jiang, T. A. O., & Li, J. P. (2021). *A Novel Stacking Approach for Accurate Detection of Fake News.* 9. <https://doi.org/10.1109/ACCESS.2021.3056079>
- Kaliyar, R. K. (2018). *Fake News Detection Using A Deep Neural Network.* *2018 4th International Conference on Computing Communication and Automation (ICCCA)*, 1–7.
- Khan, A. (2019). *The Use and Abuse of Social Media for Spreading Fake News.* *2019 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS), June*, 145–148.
- Mahid, Z. I., Manickam, S., & Karuppayah, S. (2018). *Detection Techniques.* *2018 Fourth International Conference on Advances in Computing, Communication & Automation (ICACCA)*, 1–5.
- Manzoor, S. I. (2019). *Fake News Detection Using Machine Learning approaches A systematic Review.* *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), Icoei*, 230–234.
- Meel, P., & Vishwakarma, D. K. (2019). *Fake News, Rumor, Information Pollution in Social Media and Web: A Contemporary Survey of State-of-the-arts, Challenges and Opportunities.* *Expert Systems With Applications*, 112986. <https://doi.org/10.1016/j.eswa.2019.112986>
- Nguyen, G., Dlugolinsky, S., Bobák, M., Tran, V., López, Á., Ignacio, G., Peter, H., & Ladislav, M. (2020). *Machine Learning and Deep Learning frameworks and libraries for large-scale data mining : a survey.* *Artificial Intelligence Review*, 52(1), 77–124. <https://doi.org/10.1007/s10462-018-09679-z>
- Ning, M., Nyow, X., & Chua, H. N. (2019). *Detecting Fake News with Tweets ' Properties.* 24–29.



- Orellana-rodriguez, C., & Keane, M. T. (2018). Attention to news and its dissemination on Twitter : A survey ☆. *Computer Science Review*, 29, 74–94.  
<https://doi.org/10.1016/j.cosrev.2018.07.001>
- Oshikawa, R., Qian, J., & Wang, W. Y. (2018). *A Survey on Natural Language Processing for Fake News Detection*.
- Parikh, S. B., & Atrey, P. K. (2018). Media-Rich Fake News Detection : A Survey. 2018 *IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*, 436–441.  
<https://doi.org/10.1109/MIPR.2018.00093>
- Poddar, K., & D, G. B. A. (2019). *Comparison of Various Machine Learning Models for Accurate Detection of Fake News*. 1–5.
- Preparation, T. (2019). *Fake News Detection using Stance Classification : A Survey*.
- Series, I. O. P. C., & Science, M. (2021). *Fake News Detection Using Machine Learning Approaches*. <https://doi.org/10.1088/1757-899X/1099/1/012040>
- Sharma, K., Qian, F., Jiang, H. E., & Ruchansky, N. (2019). *Combating Fake News : A Survey on Identification*. 10(3).
- Shelke, S., & Attar, V. (2019). Source detection of rumour in a social network – A review. *Online Social Networks and Media*, 9, 30–42.  
<https://doi.org/10.1016/j.osnem.2018.12.001>
- Skizrqvrskrq, D., Dnh, E., Kdv, Q., Lpsdfw, L., Rxu, L. Q., & Vrflhw, P. (2018). 'hwfwlqj )dnh 1hzv zlwk Odflqh /hduqlqj Ohwkrq. 528–531.
- Smitha, N. (2020). *Performance Comparison of Machine Learning Classifiers for Fake News Detection*. 696–700.
- Zhou, X., & Zafarani, R. (2020). *A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities*.
- Zubiaga, A. (2019). *Mining social media for newsgathering : A review*. 13.  
<https://doi.org/10.1016/j.osnem.2019.100049>