

# *A Review of Fake News Detection Using Machine Learning Techniques*

Sunil Kumar<sup>1</sup>

<sup>1</sup>Student, Department of CS & IT,  
Central University of Jammu, Bagla(Rahya Suchani),  
J&K, India.  
sunilbandotra@gmail.com

Dr. Bhavna Arora<sup>2</sup>

<sup>2</sup>Assistant Professor, Department of CS & IT,  
Central University of Jammu, Bagla(Rahya Suchani),  
J&K, India  
bhavna.csit@cuammu.ac.in

**ABSTRACT**—With the growing number of users of the internet people share millions of posts, articles and videos. These posts are shared on a number of social media platforms along with twitter, facebook, youtube and other social networking web sites. It is now a well-known fact that many times any misinformation may even lead to conflicts and it also has a significant influence on public opinion. The propagation of fake news stories on social media platforms and on Internet is duping people to an extent, stopping which is the need of the hour. The research area of fake news detection is gaining interest but at the same time it involves a number of challenges due to unavailability of quality resources such as datasets, published literature etc. The existing systems are not that much efficient in the detection of fake news because of the lack of fake news datasets that are comprehensive and at the same time are community-driven datasets. This has become one of the major roadblocks in the research works related to fake news detection. At the same time there are some restrictions on the input and the news category that makes it less varied. The fake news detection system aims to use data repositories such as Buzzfeed, Politifact, CREDBANK, FakeNewsNet and various classification techniques such as Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Decision tree (DT), Logistic Regression (LR), RandomForests etc. to help to achieve maximum accuracy. This review paper provides a detailed review of various fake news detection techniques used by different researchers, the datasets they have worked upon and various evaluation parameters used by them for performance evaluation of their models. We have also discussed the difficulties and challenges faced in fake news detection.

**Keywords:** Fake News detection, Data Science, Natural Language Processing (NLP) Techniques, Machine Learning, Disinformation, Misinformation.

## **I. INTRODUCTION**

"Fake News" is fabricated news which is propagated through some traditional media platforms such as newspapers, television as well as social media platforms. The fake news is usually spread for misleading the readers which can also result in causing significant negative effects on society, for financial or political gains. The fake news problem has been there for decades but now it has become much more frequent with the increase in the use of social media platforms.

This is because with an increase in use of online social media platforms such as Twitter, Facebook and Youtube,

fake news can easily spread like a wildfire. This fake news may be written for economical or political gains. Such fake news articles can be detected by making use of a number of NLP techniques, Machine learning (ML) and Artificial intelligence (AI). The problem of fake news in India has been a particular problem during the COVID-19 pandemic.

On April 2<sup>nd</sup>, health workers, doctors and revenue officers were attacked by locals in Indore city of Madhya Pradesh. These officials had gone there for the identification of members of a 65-year-old man's family who died of COVID-19. The reason behind this attack was that some fake videos were circulated claiming that these officials were taking away the healthy Muslims and injecting them with the virus.

A study was conducted by researchers from the University of Michigan on misinformation in India. The study shows an increase in the propagation of fake news, particularly after the PM Narendra Modi's announcement of Janata curfew on March 22<sup>nd</sup>, 2020. On the basis of the study, it was noticed that the count of fake news rose from 2 in the 3<sup>rd</sup> week of January, 2020 to 60 in the 1<sup>st</sup> week of April, 2020. The study found that the number of people who were affected emotionally increased. This study makes use of 243 instances of misinformation that were collected from the records that were maintained by Tattle Civic Technology. It is a news project based in Delhi aiming to make legitimate information to be accessible to the mobile users. The record consists of all the fake news stories that are detected as fake news by six fact-checkers. These fact checkers are namely Fact Check, BOOMlive, AltNews, Quint Webqoof, IndiaToday, Factly and NewsMobile Fact Checker—certified by International Fact-Checkers Network (IFCN) between January 23<sup>rd</sup> and April 12<sup>th</sup>, 2020[1]. This paper provides a detailed review of various fake news detection techniques used by different researchers, the datasets they have worked upon and various evaluation parameters used by them for the evaluation of their model's performance. This paper also tells us about the procedure of detection of fake news and its implementation. For the classification of a news article as fake or real, firstly the text article is extracted, data is analyzed and then a number of classification algorithms are used for the classification of the news article as real or fake.

## II. WORKING OF FAKE NEWS DETECTION

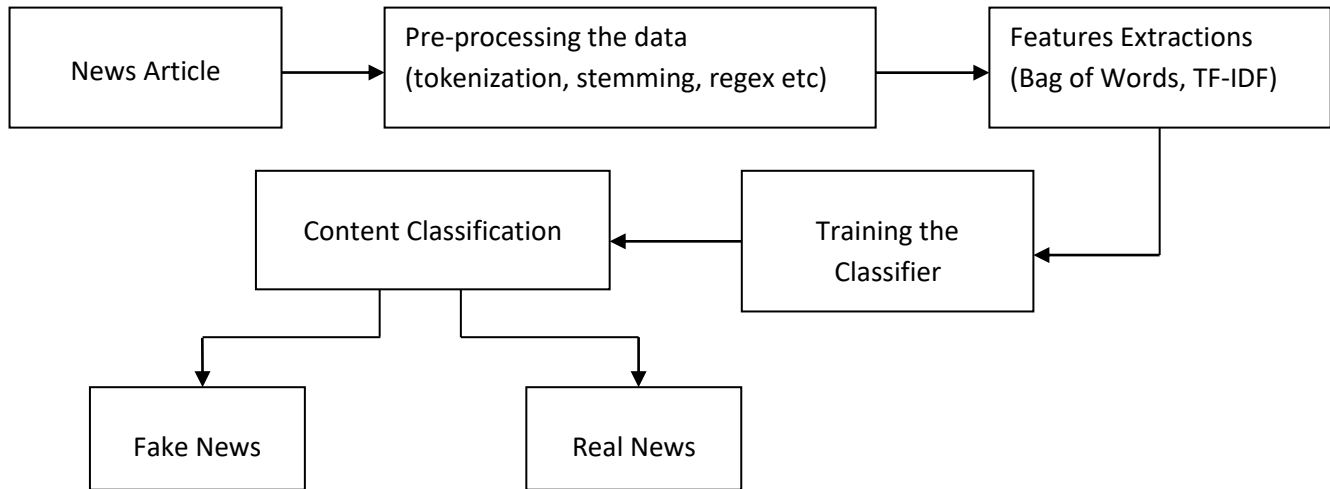


Fig. 1 - General approach for fake content detection [2]

The data preprocessing is needed to be done on the news article from the database or the Internet. This data preprocessing on the news article is necessary to train the models machine learning (ML) algorithms. Some of the techniques which are used for the conversion of text data into a form on which modelling can be performed are Regex, Tokenization, Stopwords, Stemming, Lemmatization etc [3]. These techniques are discussed as below:

### A. Data Preprocessing Techniques

Text data needs to be preprocessed before implementing machine learning algorithms on them. Some of the techniques used to preprocess the data are as follows:

- **Regex**

Many times before applying machine learning algorithms to text data, it becomes important to remove certain special characters from the text data. It is a primary importance of data preprocessing and it can be done by using a number of methods. Regex is one of the data processing technique which is used to remove punctuations from the string.

- **Tokenization**

Tokenization is a data preprocessing technique of breaking up of text into smaller units known as tokens. These smaller units called tokens can either be words, subwords or characters.

- **Stopwords**

The words which are most commonly used in any natural language are known as stopwords. (such as: "a," "the," "of," "I," "you," "it," "and"). These stopwords don't convey any subtle information necessary for the training

of a classifier. Rather it creates confusion for the classifier [4]. These stopwords needs to be removed since they do not contribute much value to the document's meaning.

- **Stemming**

Stemming is an important NLP technique which is used to convert a word into its basic form by removing its affixes. For example, if the stemming process is applied on the words 'eating', 'eats', 'eaten', then, the basic word thus formed will be 'eat'.

While performing stemming two types of errors occur namely Over-stemming, and Under-stemming. Over stemming occurs when a larger part of a word is removed, which can result in two or more words being reduced to the same root word or stemmed incorrectly. Under stemming occurs when two or more words instead of reducing to the same root word are incorrectly reduced to more than one root word.

- **Lemmatization**

Lemmatization is another form of stemming in which the affixes from a word are removed in order to extract its base form. The difference between the outputs of lemmatization and stemming is that in case of lemmatization, the base form of word is called 'lemma' whereas it is known as 'stem' in case of stemming. Lemmatization converts any word to its base form which is meaningful but in case of stemming the base form of the word may be meaningless.

After the data preprocessing is done, feature extraction is needed to be done which is also an important step towards the training of the model. Text analytics can be performed only when the raw text is converted into numerical features. Feature extraction can be performed by using

techniques such as Bag of Words (BoW), Term Frequency-Inverse Document Frequency (TF-IDF) etc. which are discussed as below:

## B. Feature Extraction Techniques

To perform text analytics, raw text is needed to be converted into numerical features. There are two main feature extraction techniques: Bag of Words and TF-IDF.

- **BoW**

A bag-of-words (BoW) is a technique which involves the representation of text that tells about the occurrence of words within a document. It gives the frequency count calculation for each word within that document. The numerical representation of the data is created using this frequency count.

- **TF-IDF**

TF-IDF is another widely used feature extraction techniques which consists of two components: Term Frequency (TF) and Inverse Document Frequency (IDF). A word's local importance by its occurrence in a document is identified by the Term Frequency (TF). On the other hand, the statistical weight that denotes the importance of a particular term in a document collection is termed as Inverse Document Frequency (IDF).

These input features play an important role for training various models of machine learning. Usually the dataset is divided into training-testing sets in the ratio of 70:30 respectively [5].

Next, we need to develop a model that can perform the classification of the output into various relevant groups/classes. This classification can be done by using a number of classification algorithms. Some of the classification algorithms which can be used for detection of fake news are discussed as below:

## C. Classification Algorithms

We can use a number of classification algorithms like Naïve Bayes(NB), Decision Tree(DT-J48), Random Forest(RF), Support Vector Machine(SVM), Gradient Boosting, K-Nearest Neighbour(KNN), Logistic Regression(LR), Multinomial Naive Bayes(MNB), Stochastic Gradient Descent(SDG) etc [6].

- **Naive Bayes(NB) Classification Algorithm**

Naive Bayes (NB) algorithm is a supervised learning classification algorithm which is based on the principle of Bayes theorem. It is an effective classification algorithms which plays an important role to build fast machine learning models that are very quick in making predictions. In Granik M et al. (2017) [7] Naive Bayes(NB) classification algorithm is used for the detection of fake news achieving an accuracy of approximately 74%.

- **Support Vector Machine (SVM) Algorithm**

Support Vector Machine (SVM) algorithms are supervised learning models for solving problems related to the binary classification [8]. The Support Vector Machine model is used to define a decision boundary which classifies data points based on some feature set [9].

- **Logistic Regression (LR) Algorithm**

Logistic regression (LR) models a binary dependent

variable by making use of a logistic function. For transformation of the output into a probability value, LR makes use of a sigmoid. Minimization of the cost function is the main objective of the LR model in order to obtain an optimal probability. The intuitive equation for classification problems involving multiple classes is given by LR model [10].

- **Decision Trees (DT) Algorithm**

A decision tree as the name suggests is a tree structure in the form of a flow-chart. It makes use of a branching method to denote each and every outcome of a decision which is possible. A specific variable's test is represented by each node within the tree and outcome of that test is represented by each branch.

- **Random Forests (RF) Algorithm**

Random forests algorithm is an ensemble learning method which is much better than simple machine learning methods. It combines various decision trees where each decision tree works separately for prediction of the outcome of a class. The majority votes received by a class give the final prediction. For getting the best outcome for classification and regression, random forests algorithms uses a combination of multiple algorithms. The Random Forest (RF) algorithm starts with a 'decision tree' having input entered at the top. Then, it goes down the tree and the data is broken into smaller sets on the basis of specific variables. The advantage of random forests is its lower error rate in comparison with other models [11].

- **K-Nearest Neighbours (KNN) Algorithm**

KNN is another simple machine learning model. It is an unsupervised model which is used for the prediction of the outcome on the basis of a specific data without the requirement of a dependent variable. The estimation done by KNN algorithm is based on the fact that how likely a data point is to be a member of one group or the other. In order to determine the group to which a data point belongs, it is required to look at various data points which are surrounding that particular data point.

So, with the final output of these algorithms we can make a conclusion of the news article being fake or real. We can also evaluate the performance of our model using various evaluation metrics which have been discussed later in this paper.

## III. LITERATURE SURVEY

Since, for few decades the fake news problem has become much more frequent with the increase the use of Internet, therefore, the research area of detection of fake news is gaining interest. Here, we have reviewed some of research works done by different researchers and we have presented a detailed analysis of these research works in Table 1. We have done analysis on the different machine learning techniques being used by these researches, different datasets used by them and various metrics used to evaluate their model's performance.

In [12], the authors have illustrated that the titles of fake news are having more verbs and nouns. They extracted a number of features such as Complexity, Psychology and Stylistic features as follows:

- Complexity features are used to calculate how much of the text is complex and readable.
- Psychological factors are used for measuring the process of human comprehension and anxiety based on writing, for ex- the number of emotional words etc.
- Stylistic features are used for defining the text's style and syntax by particular writers, usually the number of verbs and the number of nouns.

Then, Support Vector Machine (SVM) classification model was built by making use of these features. The authors have used two datasets to test their model. The first dataset is a real news dataset collected from BuzzFeed and other news websites and the second dataset is Burfoot and Baldwin's satire dataset. They achieved 91% accuracy on comparing real news against satire articles. However, it is observed that when fake news is predicted against real news the accuracy becomes 71%.

In [13], the approach used to detect fake news is based on n-gram features. N-gram features and machine learning classification techniques are used for text analysis in this approach. The authors have studied and compared different supervised classification techniques, namely, K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Stochastic Gradient Descent (SGD), Linear Support Vector Machine (LSVM), Logistic Regression (LR) and Decision tree (DT). The real and fake news websites (Kaggle) are used to compile a dataset which in turn is used for the performance evaluation of the experiment. And the evaluation results are very encouraging. They have done an investigation and comparison between two features extraction techniques and six machine learning classification techniques. The experimental evaluation shows that the best performance is achieved while using TF-IDF as feature extraction technique and classification algorithm named Linear Support Vector Machine (LSVM) (92% accuracy).

In [14], a study is presented which performs predictions related to biasness and factuality of different news media. A large number of news websites along with a rich set of features are used to perform the experiment. These features have been collected from (i) articles from the target news medium, (ii) its Wikipedia page, (iii) its Twitter account, (iv) its URL structure and (v) Web traffic information attracted by it. The results of the experiment show the necessity of each feature type.

In [15], instead of using supervised methods, the authors have used unsupervised methods for fake news detection. They have also studied a challenging problem which includes understanding profiles of users on social media platforms for fake news such as Twitter. This led to the laying down of foundation for making use of user profiles for detection of fake news. They have proposed a UFD framework which is an unsupervised learning framework. UFD models the factuality of news and user's credibility by making use of a probabilistic graphical model. For the dataset they have used the ground truth labels from BuzzFeed and Politifact, a well recognized fact-checking website. These labels are given by journalist experts. The authors have compared their proposed algorithm with four other unsupervised algorithms named MajorityVoting, TruthFinder, LTM and CRH. The experiments performed

on LIAR and BuzzFeed datasets shows that MajorityVoting achieves the worst performance as compared to all other unsupervised algorithms including the author's proposed algorithm. The author's proposed algorithm (UFD) gave the best performance on LIAR dataset achieving an accuracy of 18.4% more than the 2<sup>nd</sup> best performing algorithm. UFD becomes the best performing algorithm on BuzzFeed dataset except for recall. The authors have also done a comparison of their proposed algorithm with some supervised learning algorithms by conducting some experiments and their findings show that the performance of supervised learning methods was lower than proposed algorithm.

In [5], the authors have used an approach different from simple machine learning approach known as an ensemble approach. Ensemble learners are used for classification of news articles in an automated way. Every news article has different features that are used for distinguishing between real and fake contents. By using these features, they have trained multiple machine learning algorithms by making use of different ensemble methods. They have conducted their experiment on 4 datasets for their model's performance evaluation. These datasets are the "ISOT Fake News Dataset" (DS1), second dataset is (DS2) dataset which is publicly available dataset collected from Kaggle, the third dataset (DS3) is also collected from Kaggle and the fourth dataset is a combination of all other three datasets which consists of articles from these datasets (DS4). Since the nature of the articles is different in every dataset, therefore, the combined dataset (DS4) is created. This combined dataset (DS4) is used to evaluate the performance of classification algorithms on datasets covering a sizable array of domains within a single dataset. All these datasets includes a number of real and fake news articles from various domains. While training the models, the corpus which is to be used as an input firstly needs to be preprocessed. After the data preprocessing step is completed, the linguistic features needs to be extracted, which is accomplished by using LIWC (Linguistic Inquiry and Word Count) 2015 tool. LIWC is a process of classification of a text into different discrete and continuous variables. Different machine learning models are trained using these features as input features. For achieving the highest possible accuracy for a particular dataset, different hyperparameters are used to train the learning algorithms. A set of different parameters is used to train each model multiple times. This is done for optimizing the model in order to get the best outcome. This is done by making use of a grid search. To evaluate the performance of their model on multiple datasets, the authors have used different ensemble learners namely boosting, bagging and voting classifiers. The authors have modeled two voting classifiers each of which consists of three learning models. One of the voting classifier is an ensemble of random forest (RF), K-Nearest Neighbours (KNN) and Logistic Regression (LR). The other classifier is an ensemble of Linear Support Vector Machine (LSVM), Classification and Regression Trees (CART) and Logistic Regression (LR). The majority votes provided by all the algorithms give the final classification of voting classifiers. They have also generated a bagging ensemble which consists of a combination of 100 decision

trees. Also, XGBoost and AdaBoost are two boosting ensemble algorithms trained by them. The authors have compared the performance of their methodology with some benchmark algorithms such as 5-fold cross validation Linear Support Vector Machine (LSVM) known as Perez-LSVM, Wang-Bi-LSTM and Wang-CNN. The performance evaluation metrics used by the authors are accuracy, F1 score, precision and recall for each model. Their experiments show that Perez-LSVM and Random Forest are the algorithms achieving the maximum accuracy on dataset DS1 which is 99%. Wang-Bi-LSTM and Wang-CNN which are the benchmark algorithms are the worst performing algorithms on dataset DS1. Decision trees (bagging classifier) and XGBoost (boosting classifier) have achieved maximum accuracy of 94% on dataset DS2. LSVM, Random Forests, and Perez-LSVM are the worst performing algorithms on dataset DS2. Perez-LSVM is the algorithm whose performance was the best on dataset DS3 achieving an accuracy of 96% whereas the ensemble learners achieved an accuracy is 93.5%. The algorithm with best performance on dataset DS4 is Random Forest achieving an accuracy of 91% and Wang-Bi-LSTM is the worst performing algorithm achieving an accuracy of 62%. The overall best performance is shown by bagging classifier (decision trees) algorithm achieving an accuracy of 94% and Wang-Bi-LSTM is the worst performing algorithm with an accuracy of 64.25%.

In [16], the authors have developed a theory-based model for the detection of fake news that emphasizes on news content so that they can detect false stories before they are distributed on social media (Early Detection). The model uses a set of manual features to represent news articles, capturing both content structure and style across language levels (i.e., lexicon-level, semantic-level syntax-level and discourse-level) through conduct of a study which is interdisciplinary in nature. They have conducted experiments on two datasets which includes news articles collected from BuzzFeed and PolitiFact, respectively. The authors have compared the performance of their proposed methods with several baseline methods given in Pérez-Rosas et al.[17] which is content-based model, Castillo et al.[18] which is propagation-based model and Shu et al.[19] which is a hybrid model. Among the baseline models, the propagation-based model performed well as compared to content-based model. At the same time the best performance is delivered by the hybrid model. The method proposed by the authors outperforms all the baseline models. 5-fold cross-validation is used along with a number of supervised classification models, where XGBoost and Support Vector Machine (SVM with linear kernel), and Random Forest (RF) are the best performing algorithms in comparison to other classification algorithms such as NB and LR. Accuracy, precision, recall, and F1 scores are used as the performance evaluation metrics. The proposed algorithm is said to achieve 89.2% of accuracy with Politifact dataset and 87.9% with BuzzFeed dataset.

In [20], the authors have evaluated the performance evaluation of five machine learning models namely Support Vector Machine (SVM), K-Nearest Neighbour (KNN), Random Forest (RF), Logistic Regression (LR),

and Decision Tree (DT) and three deep learning models i.e - Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNNs) and Gated Recurrent Unit (GRU) on two real and fake news datasets. The two datasets involved are KDnugget dataset and ISOT fake news dataset. They have used feature extraction techniques such as TF, TF-IDF and embedding techniques. These techniques are used for the text representation for machine learning and deep learning models respectively. For the performance evaluation of the model, the authors have used F1-Score, Precision and Recall as evaluation metrics. A corrected version of McNemar's test have been used for making the decision of choosing a model for detection of fake news which is best as compared to other models. This is done on the basis of important differences between the performances of two model. Their experimental findings shows that on small datasets, K-Nearest Neighbours (KNN) achieved the best performance and worst performance on small datasets is shown by other models like Decision Tree, Support Vector Machine, Logistic Regression, Covolutional Neural Networks (CNN), Gated Recurrent Unit (GRU) and LSTM (Long Short-Term Memory). According to thier final experiments, the highest accuracy on the IOST dataset is achieved when the classification method used is Random Forest (RF) and features extraction technique used is TF-IDF. And when the classifier Logistic Regression (LR) is used along with TF-IDF as the feature extraction technique, then the accuracy is highest on the KDnugget dataset is achieved by with. The novel stacking model proposed by the authors has achieved 99.94% of testing accuracy on the ISOT dataset and 96.05 % of testing accuracy on KDnugget dataset.

In social media, the main aim of fake news detection is extraction of useful features and building effective models. In the future, the fake news detection can be performed by making use of these existing social media datasets. Thus, it is very important to have a comprehensive, community-driven and large-scale dataset with multi-dimension information in online fake news detection systems. The present day research works related to fake news detection makes use of a data repository named FakeNewsNet. This data repository has the potential for boosting up the research works related to the study of fake news detection problems [21].

The present day studies related to detection of fake news are based on the use of multi dimensional data respositories such as FakeNewsNet for collection, analyzation and visualization of fake news. To curb the phenomenon of fake news propagation and for the prevention of the significant influence of fake news on the society, the most important thing is that a fake news should be detected earlier. But this early detection of fake news story is a cumbersome task and challenging one and is of primary importance. It is because a fake news story which is detected earlier can be restricted from being propagated further. The task of early detection is challenging because of the unavailability and inadequacy of the information required for the detection of fake news, when its propagation phase is at the early stage [22].

## A. Evaluation Metrics

For the performance evaluation of different algorithms, a number of evaluation metrics are available. The basis for most of these evaluation metrics is the concept of confusion matrix. The confusion matrix of a classification model is a tabular representation of its performance on the testing set. It is matrix which consists of 4 parameters: true positive (TP), false positive (FP), true negative (TN), and false negative (FN) [5].

		PREDICTED CLASS	
		0	1
ACTUAL CLASS	0	TN	FP
	1	FN	TP

Fig. 2 - Confusion Matrix

0 is label for Fake News

1 is label for Real News

### • Accuracy

Accuracy is an important evaluation metric which represents the observations that are either true or false and are correctly predicted accordingly. The following equation can be used to calculate a model's accuracy:

$$\text{Accuracy} = \frac{TN + TP}{TP + FP + FN + TN} \quad (1)$$

Mostly, higher value of accuracy represents a good model. But while training a fake news detection model, if an article which is actually false is predicted as true (false positive), then its consequences may have a negative impact on the society. On the other hand, if another article which is actually true is predicted as false, then it can create trust issues. In order to avoid that, we can make use of other three evaluation metrics. These are F1-score precision and recall. These metrics take into consideration the observations that are predicted incorrectly.

### • Precision

The precision of a model is defined as its ability to not classify label a sample to be positive which is actually negative. In fake news detection, number of articles out of all the positively predicted articles which are marked as true represents the precision of a model.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

### • Recall

The recall of a model is defined as its ability to find all the positive samples. In fake news detection problem, the of news articles out of total number of news articles which are predicted as true represents the recall of a model.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

### • F1-Score

The F1-score is the metric which is used for measuring the accuracy of a model. It is used for the evaluation of the binary classification problems, which usually consists of classification classes as 'positive' or 'negative'. F1-score is also defined as the harmonic mean of the terms: precision and recall.

$$F-1 \text{ Score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (4)$$

In fake news detection systems, it is necessary that real news articles are not labeled as fake news because it can create trust issues among people.

Out of all the research papers reviewed by us, [12], [13] and [14] have used accuracy for the performance evaluation of their proposed algorithm. Whereas [15], [5], [16] and [20] have used Accuracy, F-Score, Precision and Recall as the performance evaluation metrics. We have presented detailed performance evaluation metrics for various research papers reviewed by us in Table 1.

## IV. RESEARCH GAPS AND CHALLENGES

There are a number of challenges associated with the study of detection of fake news. Firstly, the process of classifying fake news manually is very subjective in nature. Typically, it needs to verify the news article and its content's validity by comparing it with multiple sources manually. Also, the process of detecting viral fake news becomes a challenging task. Now-a-days news spread through a number of social media platforms. The proper understanding of the natural language processing (NLP) is required to solve the fake news problems using automated way. This is difficult and cumbersome process. The result of all these complexities is that the classification of text as fake news becomes a difficult task [23]. Also, the detection of fake new is a cumbersome process since there is no way for controlling the things that citizens can read and the sources that they can refer to collect particular news. It is a known fact that every social media account has power to spread any kind of news whether it is real news or fake news. The challenge here is to develop a tool which can help readers to do the prediction of news as real or fake.

Another challenge is the data source. Unstructured data has a lot of irrelevant data and garbage values that can affect the performance of the algorithm developed. Also, now fake news comprises of images and videos along with the text. So, there is a need to construct such a model which can work upon multi-modal dataset. But, the challenge here is unavailability of multi-modal dataset. Therefore, a number of researchers in the field of study of detection of fake news are working upon this problem by creating a number of multimodal datasets such as FakeNewsNet. FakeNewsNet is a data repository which is multi-dimensional in nature. This FakeNewsNet dataset is constructed using FakeNewsTracker, and it has the potential for boosting up the research works related to the study problems related to fake news detection [21]. The combination of above issues results in making the process of detection of fake news to become a big challenge for researchers in this field.

Authors	Datasets	ML Algorithms	Highest Accuracy	Precision	Recall	F-Score
Horne BD et al. (2017) [12]	Buzzfeed; Political News Data; Burfoot and Baldwin dataset	SVM	78% [Title] 71% [Contents]	-	-	-
Ahmed H, et al. (2017) [13]	Kaggle	SVM SDG LR	92%	-	-	-
Ramy Baly, et al. (2018) [14]	Twitter Wikipedia Articles	SVM	57.22% 62.29% 64.35%	-	-	-
Kai Shu, et al. (2019) [15]	LIAR Buzzfeed	UFD	75.9% 67.9%	True 0.766 0.667  Fake 0.750 0.692	True 0.783 0.714  Fake 0.732 0.643	True 0.774 0.690  Fake 0.741 0.668
Iftikhar Ahmed, et al. (2020) [5]	ISOT Dataset DS1 Kaggle DS2 Kaggle DS3 Combined Dataset DS4	LR SVM MLP KNN	99% 94% 96% 91%	0.94 0.94 0.96 0.92	1 0.95 1 0.91	0.99 0.94 0.96 0.91
X. Zhou, et al. (2020) [16]	PolitiFact BuzzFeed	SVM NB RForest LR XGBoost	89.2% 87.9%	0.877 0.857	0.908 0.902	0.892 0.879
Tao Jiang, et al. (2021) [20]	ISOT dataset KDnugget	LR RF SVM CNN KNN GRU DT LSTM	99.94% 96.05 %	1 0.97	1 0.96	1 0.96

**Table 1.** Literature Survey and its comparison on the basis of evaluation metrics

## V. CONCLUSION

Earlier the traditional media platforms such as newspapers, radio, television etc. were used information distribution. All these media platforms were having some kind of checks from higher authorities and all of the news were cross-checked for their legitimacy. But, with the increase in use of Internet in today's modern world, most of the tasks are usually performed online. Now-a-days newspapers have been replaced by various applications such as Youtube, Facebook, Twitter, online news articles and many other social media platforms. But at the same time, with the increased use of these online services, the problem of fake news is growing day-by-day which makes things more complicated and may even lead to the danger of suppressing Social Conflict. This is because there is no control of higher authorities to ensure what kind of articles are accessible to people and what kind of

news is spread by some individual on these social-media platforms. This problem of fake news which is growing day-by-day makes things more complicated and it can also cause a change of opinion of people and their attitude towards the use of technology. When a person encounters legitimate news, there is a possibility that perceptions of people about a particular topic are assumed to be true [24]. The second problem related to fake news is that even if an article is available which contradicts people's perception, but still they believe in the content of that article which suits their thinking. They don't even look into the facts involved in that particular news article. The fake news propagation problem gained most attention during the 2016 US presidential elections. During those days, people used Facebook to share these fake news stories. These fake news stories can affect people's opinions and turn their decisions which can also cause polarization of voters towards a particular

candidate in an election. This in turn can result to be a roadblock in conduct of free and fair elections. Thus, for the prevention of the propagation of fake news, Google, Facebook, Twitter etc. are also working upon this and taking important measures to detect fake news, hate speech and objectionable content.

In this review paper, we have reviewed a few research papers by various authors using a number of techniques for fake news detection. They have also used different datasets along with these techniques on different datasets. With the help of our detailed review of these research papers, we can conclude that ensemble learner methods given by X. Zhou, et al. (2020) [16] tend to have higher accuracy than other methods. This is because of the fact that in ensemble learners, a particular technique is used to train multiple models which in turn reduces errors in detection and also improves the model's overall performance. Also, the stacking method given by Tao Jiang et al. (2021) [20] achieved 99.94% of accuracy on the ISOT dataset and an accuracy of 96.05% on the KDnugget dataset. This much of accuracy is much higher in comparison to other methods. So, this method is also recommended for fake news detection. Another conclusion drawn from this review is that to facilitate the research works in field of fake news detection systems, we should make use of a multidimensional data repository. FakeNewsNet is one of the multidimensional data repository is which is having the potential to boost the research works related to the detection of fake news [21].

Also, we can make use these techniques in order to tackle the problem of fake news propagation. It is important to use machine learning (ML) techniques for detection of fake news because we cannot perform this task by ourselves. This is because the detecting fake news articles manually is very cumbersome process. So, we should rely on these machine learning techniques for the detection of fake news. Also, we should also focus our research works mainly on the early detection of fake news so that any fake news article is suppressed before its further propagation which in turn can help the society.

## REFERENCES

- [1] Steps taken towards curbing of Fake News in India, <https://www.latestlaws.com/articles/steps-taken-towards-curbing-of-fake-news-in-india/>
- [2] H. Ahmed, I. Traore, and S. Saad, "Detecting opinion spams and fake news using text classification," *Security and Privacy*, vol. 1, no. 1, 2017.
- [3] G. Geetanjali Jain and A. Mudgal, "Natural language Processing Based Fake News Detection using Text Content Analysis with LSTM," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 8, no. 11, pp. 109–115, Nov. 2019.
- [4] Elmurugi, Elsharif & Gherbi, Abdelouahed. (2018). Detecting Fake Reviews through Sentiment Analysis Using Machine Learning Techniques.
- [5] I. Ahmad, M. Yousaf, S. Yousaf, and M. O. Ahmad, "Fake News Detection Using Machine Learning Ensemble Methods," *Complexity*, vol. 2020, pp. 1–11, 2020.
- [6] [https://www.sas.com/en\\_gb/insights/articles/analytics/machine-learning-algorithms.html](https://www.sas.com/en_gb/insights/articles/analytics/machine-learning-algorithms.html)
- [7] M. Granik and V. Mesyura, "Fake news detection using naive Bayes classifier," 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), 2017, pp. 900–903, doi: 10.1109/UKRCON.2017.8100379.
- [8] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*, Cambridge University Press, Cambridge, UK, 2000.
- [9] T. Hofmann, B. Scholkopf, and A. J. Smola, "Kernel methods " in machine learning," 7e *Annals of Statistics*, vol. 36, no. 3, pp. 1171–1220, 2008.
- [10] T. M. Mitchell, *7e Discipline of Machine Learning*, Carnegie Mellon University, Pittsburgh, PA, USA, 2006.
- [11] B. Gregorutti, B. Michel, and P. Saint-Pierre, "Correlation and variable importance in random forests," *Statistics and Computing*, vol. 27, no. 3, pp. 659–678, 2017.
- [12] B. D. Horne and S. Adali, "This Just In: Fake News Packs a Lot in Title, Uses Simpler, Repetitive Content in Text Body, More Similar to Satire than Real News," *The 2nd International Workshop on News and Public Opinion at ICWSM*, Montreal, Canada, 2017.
- [13] H. Ahmed, I. Traore, and S. Saad, "Detection of Online Fake News Using N-Gram Analysis and Machine Learning Techniques," *Lecture Notes in Computer Science*, pp. 127–138, 2017.
- [14] R. Baly, G. Karadzhov, D. Alexandrov, J. Glass, and P. Nakov, "Predicting Factuality of Reporting and Bias of News Media Sources," *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 2018. )
- [15] S. Yang, K. Shu, S. Wang, R. Gu, F. Wu, and H. Liu, "Unsupervised Fake News Detection on Social Media: A Generative Approach," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, pp. 5644–5651, 2019.
- [16] X. Zhou, A. Jain, V. V. Phoha, and R. Zafarani, "Fake News Early Detection," *Digital Threats: Research and Practice*, vol. 1, no. 2, pp. 1–25, 2020.
- [17] Verónica Pérez-Rosas, Bennett Kleinberg, Alexandra Lefevre, and Rada Mihalcea. 2017. Automatic detection of fake news. *Arxiv Preprint Arxiv:1708.07104* (2017).
- [18] Carlos Castillo, Marcelo Mendoza, and Barbara Poblete. 2011. Information credibility on Twitter. In *Proceedings of the 20th International Conference on World Wide Web*. ACM, 675–684.
- [19] Kai Shu, Suhang Wang, and Huan Liu. 2019. Beyond news contents: The role of social context for fake news detection. In *Proceedings of the 12th ACM International Conference on Web Search and Data Mining*. ACM, 312–320.
- [20] T. Jiang, J. P. Li, A. U. Haq, A. Saboor, and A. Ali, "A Novel Stacking Approach for Accurate Detection of Fake News," *IEEE Access*, vol. 9, pp. 22626–22639, 2021.
- [21] K. Shu, D. Mahudeswaran, S. Wang, D. Lee, and H. Liu, "FakeNewsNet: A Data Repository with News Content, Social Context, and Spatiotemporal Information for Studying Fake News on Social Media," *Big Data*, vol. 8, no. 3, pp. 171–188, 2020.
- [22] Liu, Yang & Wu, Yi-Fang. (2018). Early Detection of Fake News on Social Media Through Propagation Path Classification with Recurrent and Convolutional Networks.
- [23] Thota, Aswini; Tilak, Priyanka; Ahluwalia, Simrat; and Lohia, Nibrat (2018) "Fake News Detection: A Deep Learning Approach," *SMU Data Science Review*: Vol. 1 : No. 3 , Article 10.
- [24] Manisha Gahirwal, Sanjana Moghe, Tanvi Kulkarni, Devansh Khakar, Jayesh Bhatia. Fake News Detection, *International Journal of Advanced Research, Ideas and Innovations in Technology*, vol 4, Issue-1, 2018. [www.IJARIT.com](http://www.IJARIT.com)