



# IJIRSET

International Journal of Innovative Research in  
**SCIENCE | ENGINEERING | TECHNOLOGY**

e-ISSN: 2319-8753 | p-ISSN: 2320-6710



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 10, Issue 7, July 2021

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

Impact Factor: 7.569



9940 572 462



6381 907 438



ijirset@gmail.com



www.ijirset.com

# Fake News Detection Using Machine Learning

Kulkarni Sampada Sham<sup>1</sup>, Nehete Sneha Pradip<sup>2</sup>, Bonde Utkarsha Yashwant<sup>3</sup>,

Prof. Rahul P. Chaudhri<sup>4</sup>

Final Year B. Tech. Students, Dept. of Computer Science and Engineering, HSM's Shri Sant Gadge Baba College of Engineering and Technology, Bhusawal. Affiliated to Dr. Babasaheb Ambedkar Technological University, Lonere, Raigad, Maharashtra, India<sup>1, 2, 3</sup>

Dept. of Computer Science and Engineering, HSM's Shri Sant Gadge Baba College of Engineering and Technology, Bhusawal. Affiliated to Dr. Babasaheb Ambedkar Technological University, Lonere, Raigad, Maharashtra, India<sup>4</sup>

**ABSTRACT:** Fake News has immense impact in our modern society. The rise of fake news has become a global problem that even major tech companies like Facebook and Google are struggling to solve.

We used different types of algorithms like Passive-Aggressive algorithm, Support Vector Machine, etc. to solve classification problem and also We used multiple performance metrics to compare the results for each algorithm. This work can be used as a significant building block for determining the veracity of Fakenews. With the success of machine learning technologies we can save our society from fake news. We need to develop a machine learning programme to identify fake/unreliable news based on content acquired.

**KEYWORDS:** Fake News Detection, Machine Learning, ML Algorithms, TF-IDF.

## I. INTRODUCTION

With the advancement of technology, information is freely accessible to everyone. Internet provides a huge amount of information but the credibility of information depends upon many factors. Enormous amount of information is published daily via online and print media, but it is not easy to tell whether the information is a true or false. It requires a deep study and analysis of the story, which includes checking the facts by assessing the supporting sources, by finding original source of the information or by checking the credibility of authors etc. The fabricated information is a deliberate attempt with the intent in order to damage/favor an organization, entity or individual's reputation or it can be simply with the motive to gain financially or politically. "Fake News" is the term coined for this kind of fabricated information, which misleads people. During the Indian election campaigns, we find many such fabricated posts, news articles and morphed pictures circulating on the social media. .

The purpose of this concept is to create a classification that is able to predict whether a user's claim is fake or real. The "Fake News Detection System" project uses machine learning algorithms and natural language processing techniques. Machine learning is a subset of artificial intelligence in the field of computer science that uses statistical techniques to give a computer the ability to learn with data without having to program it explicitly.

## II. LITERATURE REVIEW

Anjali Jain[1] has demonstrates a model and the methodology for fake news detection. With the help of Machine learning and natural language processing, author tried to aggregate the news and later determine whether the news is real or fake using Support Vector Machine.

Syed Ishfaq Manzoor[2] has played a vital role in classification of the information although with some limitations.

Karishnu Poddar [3] aims to tackle this issue using a computational model of probabilistic and geometric machine learning models. Both, logistic regression and SVM models give better scores with larger datasets, therefore, it is difficult to predict which one will perform better in the future.

### III. RELATED WORK

According to various researches conducted in this area, Fake news detection methods comprise of four basic types – Knowledge Based, Style based, Stance based and Visual based. This section elucidates research in all these types of detection methods and a few other important researches that have received higher recognition. It also presents some of the important features that were used recently in various research papers to determine the credibility of news. The feature extraction is the crucial phase of Machine learning.

#### 1. Fake News Detection Methods:

##### a) Knowledge Based Detection:

Its purpose is to use external sources to fact-check the claims made in the news material. Two typical external sources are the open web and knowledge graphs. Open web sources are compared in terms of consistency and frequency of claims, while knowledge graphs are used to test whether claims can be inferred from existing facts in the graph.

##### b) Style Based Detection:

Style based identification focuses on how the content is presented to the users. Fake news is not usually written by a journalist, that being said, writing styles can vary. In some of article's headlines, there is just enough information to make readers curious to go to a certain webpage or video. This type of eye-catching headlines or web links is called as click-bait headlines, which can be a source of Fake news.

##### c) Stance based detection:

This method compares how a series of posts on social media or a group of reputable sources feels about the claim - Agree, Disagree, Neutral or is Unrelated.

##### d) Visual Based Detection on Social Media:

Digitally altered images are circulating like wildfire everywhere on social media. Photoshop can be used freely these days to modify images enough to fool people into thinking they are looking at a real picture. The field of multimedia forensics has developed a number of methods for detecting tampering in videos and images. There are also some basic techniques on the web for common people to spot photo-shopped images viz. Google's reverse image search, get image metadata etc.

### IV. METHODOLOGY

#### 1. Response Based Detection

Fake news generally carries strong sentiments and thus circulates in no time on social media. Response based technique takes into consideration the collected responses on tweets/posts to determine the credibility of the news.

The project was developed in Python with the help of a Sci-Kit library. Python has a large set of libraries and extensions, which can be easily used in machine learning. Sci-Kit Learn Library is the best resource for machine-learning algorithms where almost all types of machine learning algorithms are readily available for Python, so a simple and quick evaluation of the ML algorithm is possible

#### 2. Data Collection And Analysis

In the first phase of this project, which was style based detection on the content/body of the news article; I used two different datasets of varying length and trained the model on each of them. Below are two datasets, which I used



- **LIARA Benchmark Dataset for Fake News Detection**

The original dataset contained 13 columns for train, test and validation files. The training set included 12,386 human-label short statements, sampled from news releases, TV or radio interviews, campaign speeches etc. The data was collected from a Fact-checking website PolitiFact through its API.

For implementation of First phase of the project, I chose only training set file and 2 columns from this file for classification.

The other columns can be added later to enhance the performance. Below were the columns that were used:

Column1: Statement

Column2: Label (True/False)

Classes (labels) were grouped such that in the newly created dataset you find only two labels (True/False) as compared to six present in the original. They were grouped as below:

True	-	True
Mostly-true	-	True
Half-true	-	True
Barely-true	-	False
False-	-	False
Pants-fire	-	False

The dataset contained four columns:

a) URL   b) Headline   c) Body   d) Label

This dataset included 4336 news articles with longer body text than the shorter text in the previous dataset. The average word count of bodies in the dataset was 576 words per article. The label was mentioned as 0/1; 0 for fake news and 1 for real news.

## V. STEPS OF METHOD IMPLEMENTATION

### 1. Text Preparation

Social media data is highly unstructured - most of them informal communication with typos, slang and bad-grammar etc. In order to get better insights, it is necessary to clean the data before it can be used for predictive modeling. For this purpose basic pre-processing was done on the news training data.. This step was comprised of

- Conversion to Lower case:
- Removal of Punctuations:.
- Stop-words removal:
- Tokenization:
- Lemmatization:



Text pre-processing was a necessary step before the data was ready for analysis. Features in a noise-free corpus have a smaller sample space size leading to increased accuracy.

## 2. Feature Generation

We can use text data to generate many features such as word count, frequency of large words, frequency of unique words, n-grams, etc. By representing words that capture their meanings, semantic relationships, and the variety of contexts they are used in, we can enable computers to understand the text and perform clustering, classification, etc.. For this purpose, Word Embedding techniques are used to convert text into numbers or vectors, so that computer can process them.

- **Word Embedding**

A word-embedding format generally tries to map a word to a vector using a dictionary. The following frequency based word embedding vectors was used for training the data. They are also categorized into Linguistic based features.

- **Count Vector As a Feature**

Count Vector is a matrix notation of the dataset, in which rows represent the documents in the corpus, columns represent a term from the corpus, and cells represent the count of that particular term in a particular document.

- **TF-IDF vectors as a feature**

TF-IDF weight represents the relative importance of a term in the document and entire corpus.

*TF stands for Term Frequency:* It counts the number of times a word appears in a document. Since the size of each document is different, a word may appear more in a long sized document than a shorter one. Thus, the length of the document often divides the term frequency.

$$TF(t, d) = \frac{(\text{Number of times } t \text{ occurs in a document } 'd')}{(\text{Total word count of document } d)}$$

*IDF stands for Inverse Document Frequency:* A word is of no use if it is present in all the documents. Some words like "a", "an", "the", "on", "off" etc. appear multiple times in a document but are of little importance. IDF reduces the importance of these words and increases the importance of rare words. The higher the value of IDF, the more unique the word is.

$$IDF(t) = \log_e \left( \frac{\text{Total number of documents}}{\text{Number of documents with term } t \text{ in it}} \right)$$

*TF-IDF – Term Frequency-Inverse Document Frequency:* TF-IDF works by penalizing the most commonly occurring words by assigning them less weightage while giving high weightage to terms, which are present in the proper subset of the corpus, and has high occurrence in a particular document. It is the product of Term Frequency and Inverse Document Frequency.

$$TFIDF(t,d) = TF(t,d) * IDF(t)$$

## 3. Algorithms Used For Classification

This section deals with the training of classifiers. Various classifiers were examined to predict the class of the text. I specifically explored five different machine-learning algorithms - the Multinomial Naive Bayes Passive Aggressive Classifier, Logistic Regression, Linear Support Vector Machines, and Stochastic Gradient Descent. These classifiers were implemented using the Python library Sci-Kit Learn

## Introduction to the algorithms

### a) Naïve Bayes

It is a classification technique based on Bayes' Theorem with an assumption of independence among predictors. In simple terms, a Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature.

$$P(c/x) = \frac{P(x/c)P(c)}{P(x)}$$

$P(c|x)$  = posterior probability of class given predictor

$P(c)$  = prior probability of class

$P(x|c)$  = likelihood (probability of predictor given class)

$P(x)$  = prior probability of predictor

Naive Bayes model is easy to build and particularly useful for very large data sets. Along with simplicity, Naive Bayes is known to outperform even highly sophisticated classification methods.

### b) Passive Aggressive Classifier

The Passive Aggressive Algorithm is an online algorithm; ideal for classifying massive streams of data (e.g. twitter). It is easy to implement and very fast. It works by taking an example, learning from it and then throwing it away. Passive-Aggressive algorithms are generally used for large-scale learning. It is one of the few 'online-learning algorithms'.

Passive-Aggressive algorithms are called so because:

*Passive:* If the prediction is correct, keep the model and do not make any changes. i.e., the data in the example is not enough to cause any changes in the model.

*Aggressive:* If the prediction is incorrect, make changes to the model. i.e., some change to the model may correct it.

### c) Logistic Regression

Logistic regression is a supervised learning classification algorithm, used to predict the probability of occurrence of an event (0/1, True/False, Yes/No). It uses sigmoid function to estimate probabilities. It is one of the simplest ML algorithms that can be used for various classification problems.

## Types Of Logistic Regression

Logistic regression can be divided into following types –

- i. Binary or Binomial
- ii. Multinomial
- iii. Ordinal

### d) Support Vector Machine

A "Support Vector Machine" (SVM) is a supervised machine learning algorithm that can be used for both classification or regression challenges. However, it is mostly used in classification problems. In the SVM algorithm, we plot each data item as a point in an n-dimensional space (where n is the number of features you have) in which the value of each feature is the value of a particular coordinate. Then, we perform the classification by finding the hyper-plane that separates the two classes very well (see snapshot below).

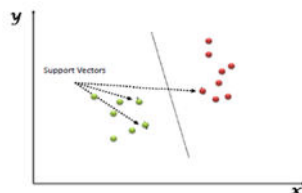


Figure 1: SVM 1

#### e) Stochastic Gradient Descent

A SGD algorithm starts at a random point, updates the cost function with each of the iteration using one data point at a time and builds a classifier with progressively higher accuracy given a large dataset. In SGD, a sample of training set or one training value is used to calculate parameters, which are much faster than other gradient descent. This algorithm is useful in cases where the optimal points cannot be found by equating the slope of the function to 0.

#### f) Random Forest Algorithm

Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems in ML. It is based on the concept of ensemble learning, which is a process of combining multiple classifiers to solve a complex problem and to improve the performance of the model

#### g) Decision Tree

Decision trees are a supervised learning technique that can be used for both classification and regression problems. It is a tree-structured classifier. The decision tree consists of two nodes, which are the decision node and the leaf node. Decision nodes are used to make any decisions and have multiple branches, while leaf nodes are the outputs of those decisions and contain no more branches.

### 4. Metrics Used To Access The Performance Of Model

In this section, I've explored some of the most important metrics by which the performance of a machine learning model is measured. These metrics measure how well our model is able to classify or evaluate predictions.

#### a) Classification Accuracy

It is the most common evaluation metric for classification problems. It is defined as the number of correct predictions versus the total number of predictions. However, this metric alone may not provide enough information to decide whether the model is a good fit.

#### b) Area under ROC-curve

The area under the ROC curve is a performance metric used for binary classification. It describes the ability of a model to spread between two classes. If the area under the curve or AUC is 1.0, it means that it has made all the predictions correctly while an AUC of 0.5 is as good as random predictions. ROCs can be further classified into sensitivity and specificity.

#### c) Sensitivity

It is called as "Recall" and Recall is the ratio of correctly predicted positive instances to the all instances in actual class - Yes. In this work, "Fake" was selected as positive class and "Real" as negative.

#### d) Specificity

It is the number of instances in the negative class that are actually predicted correctly. It is called as True negative rate.

**e) Confusion Matrix**

It is also known as Error matrix, which is a table representation that shows the performance of the model. It is special kind of Contingency table having two dimensions- “actual”, labeled on x-axis and “predicted” on y-axis. The cells of the table are the number of predictions made by the algorithm.

Total Instances		Predicted	
		Yes	No
Actual	Yes	True Positive	Fake Negative
	No	True Negative	True Negative

Table 1: Confusion Matrix

True Positives: It is correctly predicted positive values.

True Negatives: It is correctly predicted negative values.

False Positives: It is incorrectly predicted negative values as positive values.

False Negatives: It is incorrectly predicted negative values as positive values.

**Classification Report**

Scikit-learn provides a convenience report when working on classification problems which outputs precision, recall, F1 score and support for each class.

**Precision:**

Precision is the ratio of correctly predicted positive instances to the total predicted positive instances. High precision means low False Positive rate.

$$Precision = \frac{TP}{TP + FP}$$

**Recall (Sensitivity):**

Recall is the ratio of correctly predicted positive instances to the all instances in actual class - Yes.

$$Recall = \frac{TP}{TP + FN}$$

**F1-Score:**

It is the weighted average of Precision and Recall. Therefore, it takes into consideration both false positives and false negatives. F1 score is usually more useful than accuracy, especially when there is uneven class distribution.

$$F1\ Score = 2 * \frac{(Recall * Precision)}{(Recall + Precision)}$$

**VI. RESULT**

Figure 2 is a graphical representation of accuracy of learning algorithms on all datasets. It can be seen that there is no much difference between the performance of learning algorithms using various performance matrices except for decision tree and naïve bayes. Overall, the best performing algorithm is SVM (accuracy 94.7%), whereas the worst performing algorithm is decision tree (accuracy 80.74%).



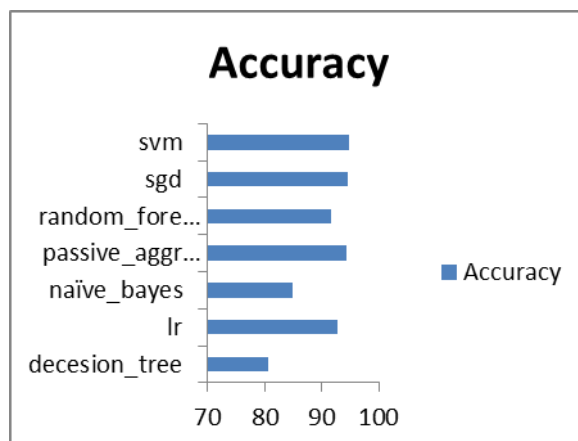


FIGURE 2:ACUURACY

## VII. CONCLUSION

The task of classifying news manually requires in-depth knowledge of the domain and expertise to identify anomalies in the text. The primary aim of the research is to identify patterns in text that differentiate fake articles from true news. We extracted different textual features from the articles used the feature set as an input to the models. The learning models were trained and parameter-tuned to obtain optimal accuracy. Some models have achieved comparatively higher accuracy than others. We used multiple performance metrics to compare the results for each algorithm.

Fake news detection has many open issues that require attention of researchers. For instance, in order to reduce the spread of fake news, identifying key elements involved in the spread of news is an important step. Graph theory and machine learning techniques can be employed to identify the key sources involved in spread of fake news.

## REFERENCES

- [1]. Anjali Jain<sup>1</sup>, Avinash Shakya<sup>2</sup>, Harsh Khatter<sup>3</sup>, Amit Kumar Gupta<sup>4</sup> (2019) "A SMART SYSTEM FOR FAKE NEWS DETECTION USING MACHINE LEARNING "
- [2]. Syed Ishfaq Manzoor, Dr Jimmy Singla, Nikita Former Assistant Professor (ICOEI 2019) "Fake News Detection Using Machine Learning approaches
- [3]. Karishnu Poddar, Geraldine Bessie Amali D, Umadevi K S(2019) "Comparison of Various Machine Learning Models for Accurate Detection of Fake News"
- [4]. Aayush Ranjan (2016) "FAKE NEWS DETECTION USING MACHINE LEARNING "
- [5]. Terry Traylor, Jeremy Straub, Gurmeet, Nicholas Snell(2019) "Classifying Fake News Articles Using Natural Language Processing to Identify In-Article Attribution as a Supervised Learning Estimator"
- [6]. [https://drive.google.com/file/d/1er9NJTLUA3qnRuyhfzuN0XUsoIC4a-\\_q/view](https://drive.google.com/file/d/1er9NJTLUA3qnRuyhfzuN0XUsoIC4a-_q/view)



**Impact Factor:  
7.569**



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details