

# Fake News Detection Using Machine Learning

Anusha Acharya<sup>1</sup>, Ashwitha Jathan<sup>1</sup>, Deepa Anchan<sup>1\*</sup>, Krithi Dinesh Kottary<sup>1</sup>, Mr. Sunil B.N<sup>2</sup>

<sup>1</sup> Computer Science and Engineering, Sahyadri College of Engineering and Management, Mangaluru, India – 575007

<sup>2</sup> Assistant Professor, Department of Computer Science and Engineering

\*Contact email address: deepa.anchan97@gmail.com

## ABSTRACT

**Fake news has been a social issue and in the era of digitization, a common man has easy access to all the data shown online. Fake news, one of the biggest problem in new era, is so powerful that it can change ones opinion and can make wrong impact while taking decisions. There are few plugins available on web browsers which give real time information regarding authenticity of news. Selection of algorithms to build these plugins make a huge impact on them. So, the motive is to build a project that would suggest which could be the best algorithm among the three implemented algorithms for further development by the developer. Algorithms being implemented for the comparison are Support Vector Machine, Naive Bayes Algorithm and Logistic Regression Algorithm.**

## *Keywords:*

SVM, Naïve Bayes, Logistic Regression

## 1. INTRODUCTION

Maybe a decade ago, there were days where people relied on sources like newspapers and television for news. Doordarshan and All India Radio were almost used by all. These news were from reliable sources and delivered to people only at a standard timing. Due to globalization and digitization, there started the era of smart devices. As the technology enhanced, people could afford to buy devices in which they could send and receive digitized data. With this development in technology, dependability on digitized data increased. There emerged innovative ideas where in social media became a medium for easier communication. Gaps between people and nations started diminishing and folks were connected to each other globally through social networks. It was a platform for networking and business. Finding customers and reaching out to right people started happening after the introduction to new communication media. As a result, new business ideas and advertising techniques came up. Back then people walked miles to spend quality time in the libraries, subscribed to weekly and monthly magazines. But now, people read blogs and articles online and subscribe to email newsletters. Even standard newspapers have a domain where they keep uploading daily e-papers.

Along with enormous advancement in communication technology, there were also weeds in the farm. Here, weeds

indicate the unrealistic data or information that circulate and deceive common people. There could be many reasons as to

why these fake news are increasing among people day by day. One such reason could be to put an enterprise or a public figure down so that they lose their image and suffer loss. In these cases, the fake news are generated by haters in such a way that it is too appealing to a common man and anybody would believe in the misinterpretation that was spread intentionally. Individuals in the media companies like journalists face a great challenge as their high quality vocabulary and work are questioned due to fake news. Some hoaxes generated bring in unnecessary discussions among public causing fear and riots.

The trustworthiness of data on the web has risen as a paramount issue of present day society. The challenge is to identify such deceptive news. Since the topic isn't new, many scientists believed that this problem could be tackled with the help of Machine Learning[1]. The usage of Artificial Intelligence has been increased in problems that deal with classification due to availability of required resources such as affordable hardware components and datasets of larger size[1]. Platforms like Whatsapp have evolved in such a way that the forwarded information are indicated to the user and notified that they cannot be trusted.

The main aim here is to develop a system that takes in a dataset and educates the users as to which classifier among Support Vector Machine, Logistic Regression, and Naive Bayes Classifier, accurately classifies the data as fake or true. The analysis shows the confusion matrices with respect to TF and TF-IDF features for both Liar-Liar and ISOT datasets for the algorithms. Precision, recall and F1-scores for the corresponding confusion matrices are intended to be displayed.

## 2. METHODOLOGY

As the end result of the project, it is expected to obtain the most accurate algorithm, among SVM, LR, Naïve Bayes classifier, that gives the accuracy, confusion matrix. The algorithms used are not friendly with noisy, incomplete and inconsistent data. Hence there is a need to modify the content into which the classifiers can tackle. The first step that the system has to undergo is pre-processing. There are many ways that add value to the performance of a classifier. They involve having more data for training, usage of bigger length news articles, removal of stop words and stemming process getting indulged in pre-processing[1]. In pre-processing step, stop word removal, punctuation removal and conversion of text to lowercase is done. Thus, by doing so the useless contents are left behind. Once this is done, extraction of features is initiated to make it easy for algorithm for classification with respect to authentication of news. There are many features that could be extracted for detection of fabricated news amongst which there could be three features that would serve up to best of their abilities. They are: TF-IDF which uses bi-gram frequency, probabilistic context free grammars (PCFGs) or syntactical features and TF feature. Usage of every feature independently yields different accuracy values. Out of these, TF-IDF outperforms amongst all these features[4]. The extracted features in our project are TF and TF-IDF[8]. Term-frequency tells about the word that has arisen maximum number of times in a sentence. In case of TF, the weight for word is equally distributed. Thus, to have knowledge about the word that has more weight in the sentence, TF-IDF is used. These two features play a main role for detection of deceptive news.

### 2.1 Datasets

ISOT\_Fake\_News\_Dataset\_ReadMe and Liar-Liar dataset are datasets that are used throughout the analysis. Two sets of datasets with varying size where used to compare the outcome of the machine learning models. The first dataset used here is named as 'Liar Liar Dataset' [17]. This dataset contains about 12.8K news articles. The second dataset used here is named as 'ISOT Fake News Dataset' [18] [19]. It is about 31K in size. Both the above mentioned datasets have statement and label columns.

### 2.2 Classifiers

#### 2.2.1 Naive Bayes Classifier

It is a reliable and a straightforward approach in the field of classification. The algorithm is well suited for natural language processing. It is based on the concept of probability and Bayes theorem and is used to predict the unknown class. The features are considered to be independent of each other. This assumption is represented mathematically as follows:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \cdot P(B|A)}{P(B)}$$

where the prior probabilities  $P(A)$  and  $P(B)$  represent the probability of the hypothesis  $A$  being true and probability of the data respectively, the posterior probability  $P(A|B)$  and  $P(B|A)$  represent the probability of hypothesis  $A$  given the data  $B$  and the probability of data  $B$  given that the hypothesis  $A$  was true. Here, the prior probabilities of the dependent classes and then of the features are calculated, then the posterior probability is calculated. Finally, the article belongs to the dependent class with the highest posterior probability.

#### 2.2.2 Logistic Regression

It is a simple model to perform supervised binary classification in machine learning. The model tries to create a connection between the features (independent variables) and the variables that needs to be predicted (dependent) . It takes a weighted combination of the input features, and passes it through a sigmoid function, which performs a probabilistic estimation of the outcome by smoothly mapping any real number to a number between 0 and 1. The values obtained between 0 and 1 as the final outcome will then be converted into exactly 0 or 1.

$$p = \frac{e^{(\beta_0 + \beta(x))}}{e^{(\beta_0 + \beta(x))} + 1}$$

$$p = \frac{e^y}{1 + e^y} \quad (1)$$

where  $p$  is the probability of success. This is the sigmoid Function. If  $p$  is the probability of success,  $1-p$  will be the probability of failure which can be written as:

$$q = 1 - p = 1 - \left( \frac{e^y}{1 + e^y} \right) \quad (2)$$

where  $q$  is the probability of failure. On dividing (1) by (2), we get

$$\frac{p}{1-p} = e^y$$

After taking log on both side, we get,

$$\log \left( \frac{P}{1-P} \right) = y$$

$$\log \left( \frac{P}{1-P} \right) = \beta_0 + \beta(X)$$

(3) is the logistic regression equation.

### 2.2.3 Support Vector Machine

It is a supervised classification algorithm used in machine learning. The advantage of using support vector machine is that the algorithm works well with the feature spaces having high dimensions. It also works well where the count of dimensions is higher than the number of samples. The algorithm works on the concept of finding a separator (hyper-plane) line between the data of different classes. Using the hyper plane and the features of the new data, the category of the data could be found out.

The extracted features are then fed to classifiers, that was earlier mentioned, separately. The output from each can be viewed and the best result obtained can be used for development of the fake news detector.

## 3. RESULTS

In this project, the performance of the classification models are compared by means of the precision, recall and F1 score on two different datasets for two different features i.e TF and TF-IDF. Precision is the ratio of true positive to the union of true positive and false positive. Recall is the ratio of true positive to the union of true positive and false negative values in the confusion matrix. F1 score is also called F-measure or F-score which gives the combined balanced score between precision and recall. It is also called harmonic mean of precision and recall.

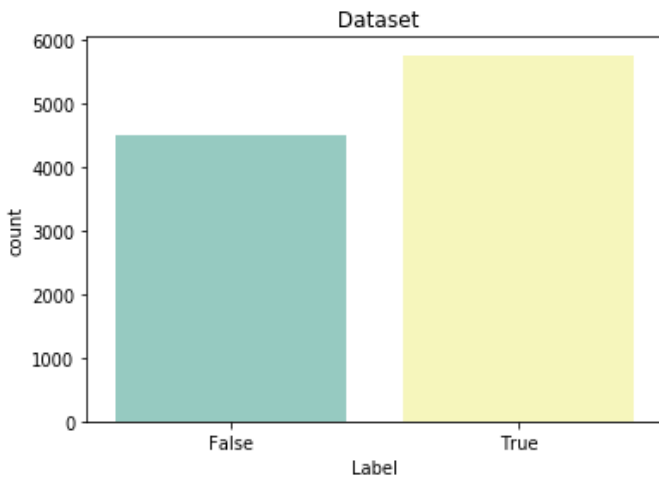


Figure 1. Distribution of data in Liar-Liar Dataset

The above graph depicts the data distribution in 'Liar-Liar Dataset'.

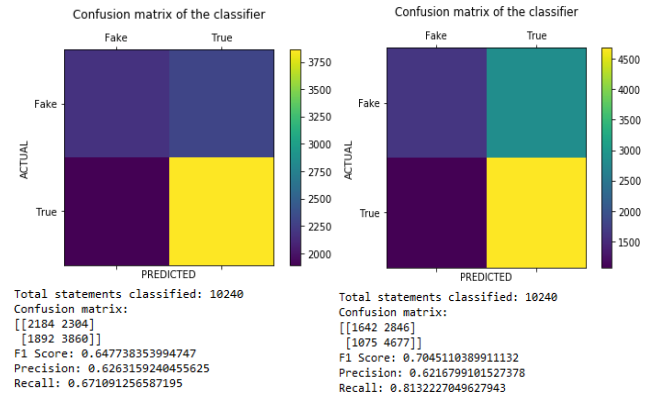


Figure 2. Confusion matrix for Logistic Regression

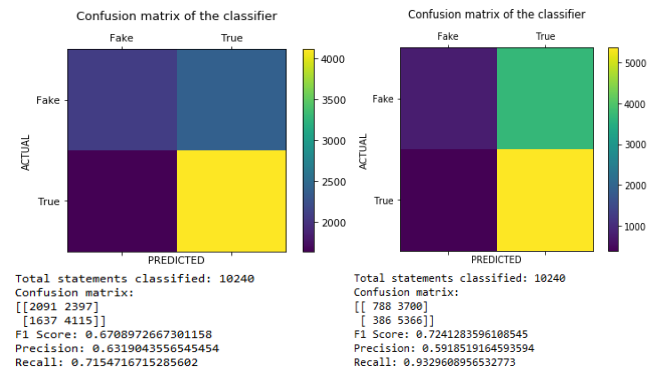


Figure 3. Confusion Matrix for Naive Bayes

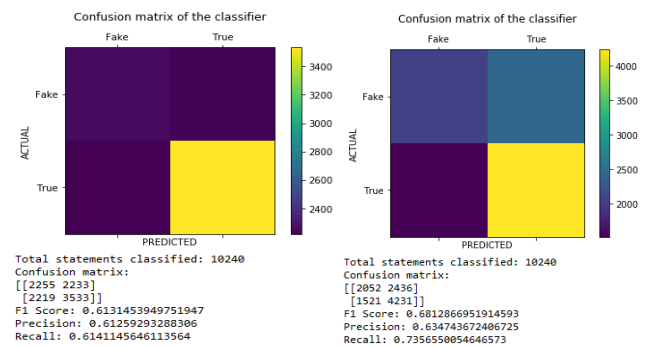


Figure 4. Confusion Matrix for SVM

Fig. 2, Fig. 3 and Fig. 4 depicts the confusion matrices obtained when the features TF and TF-IDF are extracted and

given to the classifier. The F1 score, precision and recall values are displayed.

		Logistic Regression	Support Vector Machines	Naive Bayes
tf	Precision	62.63%	61.25 %	63.19%
	Recall	67.10%	61.41%	71.54%
	F1 Score	64.77%	61.31%	67.08%
tf - idf	Precision:	62.16%	63.47%	59.18%
	Recall:	81.32%	73.56%	93.29%
	F1 Score:	70.45%	68.12%	72.41%

Table I. Classification report for Liar-Liar Dataset

Table I depicts the precision, recall and f1-score of three different classifiers for TF and TF-IDF features.

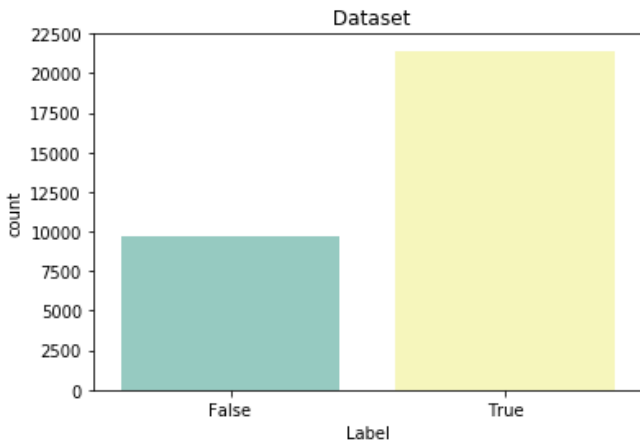


Figure 5. Distribution of ISOT dataset

The above graph shows the data distribution in ISOT dataset.

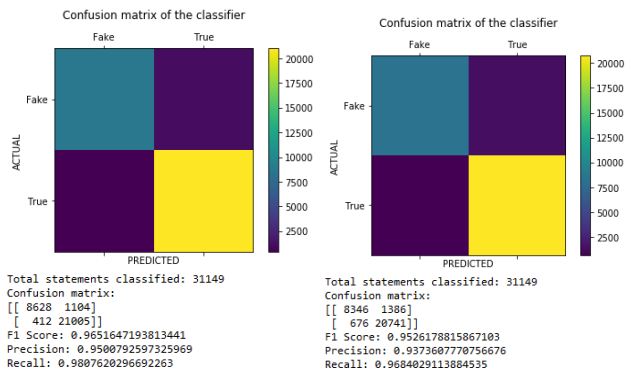


Figure 6. Confusion matrix of Logistic regression

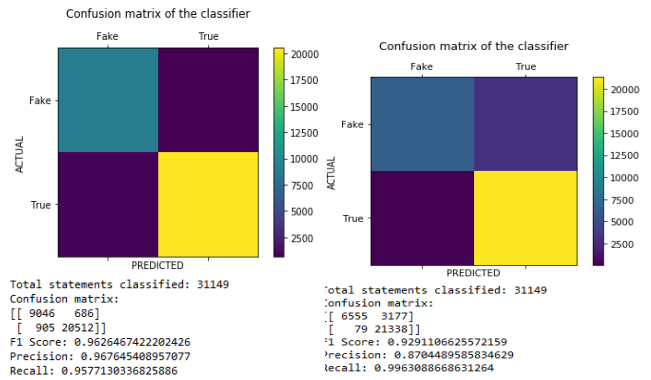


Figure 7. Confusion Matrix of Naive Bayes

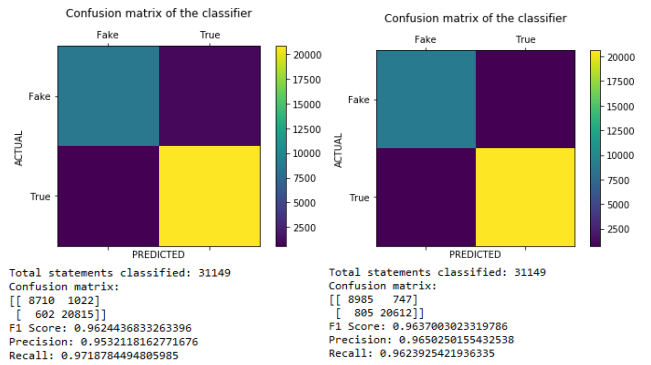


Figure 8. Confusion Matrix of SVM

Fig. 6, Fig. 7 and Fig. 8 depicts the confusion matrices obtained when the features TF and TF-IDF are extracted and given to the classifier. The F1 score, precision and recall values are displayed.

		Logistic Regression	Support Vector Machines	Naive Bayes
tf	Precision	95.00%	95.32%	96.76%
	Recall	98.07%	97.18%	95.77%
	F1 Score	96.51%	96.24%	96.26%
tf - idf	Precision:	93.73%	96.50%	87.04%
	Recall:	96.84%	96.23%	99.63%
	F1 Score:	95.26%	96.37%	92.91%

Table II. Classification report for ISOT Dataset

Table II depicts the precision, recall and f1-score of three different classifiers for TF and TF-IDF features.

As per the analysis made on Liar-Liar dataset, Naïve Bayes gave better accuracy for TF feature, whereas Logistic Regression gave better accuracy for TF-IDF feature. The second dataset – ISOT gave better results compared to Liar-

Liar dataset for the same features. This is because the size of ISOT dataset is almost thrice the size of Liar-Liar dataset.

#### 4. CONCLUSION

Fake news, which is one of the main problem of the present world has the potential to influence people's decision and mould their opinions. Detailed analysis of the implemented algorithms is carried out which is used for classification of true and fake news. Here the system will successfully recommend the researcher about the accuracy of the algorithms and gives output as to which is the best algorithm among the given algorithms. It will plot the graph of accuracy with respect to the Naive Bayes, Logistic Regression and SVM algorithms.

#### REFERENCES

- [1] M. Granik and V. Mesyura, "Fake news detection using naive Bayes classifier" 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering(UKRCON), Kiev, 2017, pp. 900-903. doi: 10.1109/UKRCON.2017.8100379
- [2] Pérez-Rosas, Verónica & Kleinberg, Bennett & Lefevre, Alexandra & Mihalcea, Rada. "Automatic Detection of Fake News" 2017.
- [3] S.Gupta, R. Thirukovalluru, M. Sinha and S. Mannarswamy, "CIMTDetect: A Community Infused Matrix-Tensor Coupled Factorization Based Method for Fake News Detection" 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Barcelona, 2018, pp. 278-281.
- [4] S.Gilda, "Evaluating machine learning algorithms for fake news detection" 2017 IEEE 15th Student Conference on Research and Development (SCORED), Putrajaya, 2017, pp. 110-115. doi: 10.1109/SCORED.2017.8305411
- [5] V.Rubin, N.Conroy, Y.Chen, S.Cornwell, "Fake News or Truth? Using Satirical Cues to Detect Potentially Misleading News", Proceedings of NAACLH-LT, pp. 7-17, 2016.
- [6] V.Rubin, N.Conroy and Y.Chen. "Towards News Verification: Deception Detection Methods for News Discourse" HICSS2015 (2015)
- [7] M.Aldwairi, A.Alwahedi, "Detecting Fake News in Social Media Networks", Procedia Computer Science, Volume 141, 2018, Pages 215-222, ISSN 1877-0509, <https://doi.org/10.1016/j.procs.2018.10.171>
- [8] B.Riedel, I.Augenstein, G.Spithourakis, S.Riedel, "A simple but tough-to-beat baseline for the Fake News Challenge stance detection task" arXiv:1707.03264v2 [cs.CL]
- [9] S.Mone, D.Choudhary, A.Singhania, "FAKE NEWS IDENTIFICATION" Retrieved from <http://cs229.stanford.edu/proj2017/final-reports/5244348.pdf>
- [10] E.Tacchini, G.Ballarín, M.Vedova, S.Moret, and L.Alfaro, "Automated Fake News Detection in Social Networks" arXiv:1704.07506 [cs.LG]
- [11] M.Gahirwal, S.Moghe, T.Kulkarni, D.Khakhar, J.Bhatia. "Fake News Detection" International Journal of Advance Research, Ideas and Innovations in Technology 4.1 (2018). www.IJARIT.com.
- [12] M.Vedova, E.Tacchini, S.Moret, G.Ballarín, M.DiPierro and L.Alfaro, "Automatic Online Fake News Detection Combining Content and Social Signals" 2018 22<sup>nd</sup> Conference of Open Innovations Association (FRUCT), Jyväskylä, 2018, pp.272279. doi:10.23919/FRUCT.2018.8468301.
- [13] Chiu, J., Gokcen, A., Wang, W., Yan, X.: "Classification of fake and real articles based on support vector machines" (n.d.). Retrieved from <https://pdfs.semanticscholar.org/05be/61a7726916644e49e578378d4faa7aa5f4ce.pdf>.
- [14] Chaudhry, A.K., Baker, D., Thun-Hohenstein, P.: "Stance detection for the fake news challenge: identifying textual relationships with deep neural nets (n.d.)". Retrieved from <https://web.stanford.edu/class/cs224n/reports/2760230.pdf>
- [15] R.Oshikawa, J.Qian, W.Wang 2018 ."A Survey on Natural Language Processing for Fake News Detection". arXiv:1811.00770v1 [cs.CL]
- [16] C.Kotteti, X.Dong, N.Li and L.Qian, "Fake News Detection Enhancement with Data Imputation" 2018 IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing, 16th Intl Conf on Pervasive Intelligence and Computing, 4th Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech), Athens, 2018, pp. 187-192.
- [17] W.Yang Wang. 2017. "Liar, liar pants on fire": A new benchmark dataset for fake news detection Proc. of ACL.
- [18] Ahmed H, Traore I, Saad S. "Detecting opinion spams and fake news using text classification", Journal of Security and Privacy, Volume 1, Issue 1, Wiley, January/February 2018.
- [19] Ahmed H, Traore I, Saad S. (2017) "Detection of Online Fake News Using N-Gram Analysis and Machine Learning Techniques". In: Traore I., Woungang I., Awad A. (eds) Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments. ISDDC 2017. Lecture Notes in Computer Science, vol 10618. Springer, Cham (pp. 127138).