# Enhancing fake news detection with ensemble-based machine learning model

Girirajula Yogananda
Master of Computer Application
Chaitanya Bharathi Institute of Technology (A)
Hyderabad, Telangana, India
yogananda521991@gmail.com

B. Srinivasa S P Kumar
Master of Computer Application
Chaitanya Bharathi Institute of technology (A)
Hyderabad, Telangana, India

Abstract - People rely on social media platforms like Facebook, WhatsApp, Twitter, and Telegram as significant channels for the distribution of information in today's society without checking the accuracy or source of the information. Social media have drawn the attention of people all around the world who use them to distribute false information because they are easily accessible, inexpensive, and convenient for exchanging information. It is feasible to fabricate fake news to deceive the public and make money for oneself or others. They can also be used for other personal objectives like slandering public figures or altering laws. Due to this, much research has been done to precisely detect false news to lessen its harmful impacts and prevent its fatal conclusion. By selecting different machine learning models (ML) such as logistic regression, random forest (RF), and passive-aggressive algorithms, they were trained to detect the Fake News articles based on the dataset. By using the Voting Classifier algorithm, we created an ensemble model that trains the other algorithms and predicts and forecasts an output (class) based on the highest probability of the selected class as the output using a voting technique. The model is evaluated using machine learning metrics such as accuracy, precision, and recall, with LR having the highest accuracy of 98%.

Keywords: Fake News, Machine Learning, Ensemble Model, Voting Classifier.

## I. INTRODUCTION

The problem of false news has grown more pervasive and influential in the modern digital era. False or inaccurate material that is presented as news is known as fake news, and it is frequently spread through a variety of media channels, including social networking sites, websites, and even established news agencies. It encompasses a wide range of misinformation, including fabricated stories, doctored images or videos, misleading headlines, and biased narratives. The rapid spread and easy accessibility of information on the internet have amplified the spread of fake news, posing significant challenges to individuals, societies, and democratic processes worldwide.

The stock price of Apple Inc. saw significant changes in 2008 as a result of false claims that Steve Jobs was ill (allegedly having a heart attack). Fake news can be intentionally created and shared to deceive, manipulate public opinion, or generate financial gains for those involved. False information may have negative effects on society, including polarisation, public confusion, loss of faith in media institutions, and even physical damage.

ML is a subfield of AI that gives computers the ability to learn from data and predict the future. Large datasets are analysed by ML algorithms, which find patterns that would be hard for people to spot on their own. By leveraging these patterns, ML algorithms can make predictions about future outcomes or behaviours.

## II. LITERATURE SURVEY

Verónica Pérez-Rosas [1] et al. have developed two brand-new datasets that include seven distinct news domains and are intended to be used in the identification of false news. They performed many exploratory analyses on the detection of linguistic distinctions in false and authentic news information and provided a detailed description of the collecting, annotation, and validation procedure. The second thing they did was run a series of learning tests to create reliable false news detectors. Furthermore, they offered comparisons of the automatic and manual detection of bogus news.

Kai Shu [2] et al.'s study includes a comprehensive investigation of the challenge of recognizing false news on social media, including definitions of fake news based on psychology and sociological theories, existing data mining tools, assessment criteria, and sample datasets. They also discussed pertinent study subjects, unsolved challenges, and prospective future research avenues for social media fake news detection.

 Copyright @ 2023 Author

Karishma Sharma [3] et al. have discussed the issue of false news in today's society and, in particular, have underlined the technological difficulties it presents. They have spoken about the approaches and strategies that are now in use for both identification and mitigation, focusing on the key developments in each methodology and their benefits and drawbacks. In addition, they have highlighted fresh research directions to aid in the future creation of successful and multidisciplinary solutions.

Soroush Vosoughi et al. [4] examined factual and false news stories that were circulated on Twitter between 2006 and 2017 for differences in their distribution. Their data was comprised of 126,000 tales that 3 million consumers tweeted 4.5 million times. They utilized information from six different factchecking agencies to categorize news as true or false, and there was 95–98% agreement between the categories. They observed that false news had a larger novelty value than actual news, demonstrating that people were more eager to distribute fresh information.

V. Agarwal et al. [5] addressed the NLP and ML techniques. They trained the data on five classifiers and used bag-of-words, n-grams, count vectorizer, TF-IDF, and bag-of-words to evaluate which one performed best for their specific dataset of tagged news statements. To determine which model works best, they used ML measures like as accuracy, recall, and f1 scores.

Shailender Kumar [6] et al concentrated on analysing publications from 2017 to 2021 and examining several methods for identifying false news. This study provides a comprehensive overview of historical and present studies on the identification of fake news using various ML algorithms.

To investigate and evaluate methods for identifying fake news, Xinyi Zhou [7] et al. looked at fake news from four different perspectives: the false information it includes, the way it is written, how it is disseminated, and the credibility of its source. The survey also offers a few intriguing research opportunities in light of the review. They also identified and discussed pertinent underlying theories from many domains to encourage interdisciplinary research on misleading news.

Julio C. S. Reis et al.'s [8] research was centred on comprehending and identifying bogus news articles that circulate on social media. To do this, they investigated a variety of features taken from news reports, including the source and social media posts. They examined the key characteristics put out in the literature for fake news identification, presented a fresh set of features, and assessed how well existing techniques and features for automatic false news detection performed at making predictions. Their research on the value and significance of traits for spotting bogus news produced some fascinating conclusions. Finally, they highlighted the potential and problems associated with using false news detection systems in practise.

Xishuang Dong [9] et al. proposed a novel deep two-way semisupervised learning model with a supervised learning path and an unsupervised learning path. These two paths were built with convolutional neural networks, and their aggregate performance was tuned for detection. Additionally, they build a shared convolutional neural network between these two pathways to exchange the low-level properties. Their experimental results using Twitter datasets demonstrate that their proposed technique may successfully identify fake news with a little annotated data.

Arjun Roy [10] et al. have developed a number of deep learning models for classifying fake news into pre-established, finegrained categories and detecting it. Bi-directional Long ShortTerm Memory (Bi-LSTM) and CNN networks were first used to generate the models. The representations from these two models are fed into a Multi-layer Perceptron Model (MLP), which classifies the data. They outperformed the state-of-theart with an overall accuracy of 44.87%, showing encouraging results.

## III. METHODOLOGY

### A. Fakenews dataset

The Fakenews dataset was utilized in this research study, which was taken from an online website called Kaggle. The dataset has 5 features: an id, a title, an author, a text, and a label. The dataset consists of 25000 different news articles. The size of the dataset in terms of memory usage is 94.0 MB.

### B. Data preprocessing

In this study the dataset has gone under several preprocessing methods. To enrich the dataset, the missing values has been handled by using Machine Learning (ML) approaches. The textual data is transformed into vectors using pre-processing methods. To Convert the textual data into numerical representations that can be fed into machine learning algorithms Natural Language Processing (NLP) techniques are used namely Count Vectorizer and TF-IDF Vectorizer. With count vectorization, sometimes referred to as the "bag of words" approach, each document is represented as a vector, with each element indicating the frequency of a specific word within the text. TF-IDF stands for "Term Frequency-Inverse Document Frequency" is a numerical statistic that reflects the importance of a word in a document within a larger collection of documents. To create numerical representations, TF-IDF vectorization combines the ideas of term frequency and inverse document frequency. The count feature vectors acquired by the count-


Copyright @ 2023 Author

vectorizer are reweighted using the TF-IDF transform algorithm. The classifier receives the input to provide improved predictions and classification outcomes.

## IV. EMPLOYED ML TECHNIQUES

In this study, four different ML algorithms were utilized to analyze the FAKENEWS dataset. These algorithms, namely Logistic Regression, Random Forest, Passive Aggressive, and Voting Classifier were selected based on their distinct abilities to handle classification tasks effectively.

### A. Logistic regression

A statistical model called logistic regression is employed in binary classification tasks to estimate the likelihood of an occurrence or a binary result from a set of input characteristics. It is a particular kind of generalized linear model that enhances linear regression to address classification issues. Any realvalued input is converted by the logistic function into a number between 0 and 1, which represents the chance that the event will occur. Logistic regression may therefore be used to produce probabilities.

### B. Random forest

Both classification and regression applications can benefit from Random Forest, a potent ensemble learning methodology. It consists of many decision trees, each of which was constructed using a randomly chosen subset of the attributes and data. With the random forest approach, more accurate predictions than those provided by a single decision tree can be achieved by merging the predictions of multiple decision trees. Using a technique called bagging or bootstrap aggregating, every decision tree in the random forest is trained separately on a random sample of the training data. By ensuring variation among the trees, this random sample helps to prevent overfitting. The ultimate forecast in classification-related issues is determined by the class that obtains the most votes from the individual trees.

### C. Passive aggressive

The PA algorithm belongs to a family of online learning algorithms that is particularly helpful in situations where the data is non-stationary, which means that the underlying distribution of the data may vary over time and the algorithm has to react fast to these changes. For binary classification jobs, it is often employed. The objective of the PA algorithm is to update its model while minimizing the loss function and making the changes as little as feasible. This characteristic makes it appropriate for online learning when real-time model updating is required and new data points are received sequentially. Since they update the model progressively without necessitating complete retraining on the entire dataset, passive-aggressive techniques are memory-friendly and computationally efficient.

### D. Voting classifier

The VC is a method of ensemble learning that combines the results of several individual classifiers to provide final predictions. It is especially helpful when multiple classifiers have varied strengths and weaknesses, and pooling their judgements can enhance performance in general. Each individual classifier in the VC is trained independently on the same training dataset and produces predictions for the input samples. For each specific classifier, other classifiers can be utilized, such as LR, decision trees, SVM and neural networks. They can also have distinct hyperparameters and be trained on separate data subsets. The VC uses a majority vote or weighted voting mechanism to integrate the predictions of the several classifiers during prediction.

## V. IMPLEMENTATION RESULTS

The efficiency of the different ML models was evaluated using metrics such as accuracy, precision, recall, F1 score, and confusion matrix. In terms of classification accuracy, precision in forecasting positive instances, recall in capturing real positive instances, total predictive power, and the trade-off between true positive and false positive rates, these measures offer a thorough overview of the models' performance. It is crucial to the assessment process to understand the phrases TP, TN, FP, and FN.

True Positive (TP): The number of correctly predicted positive instances.

True Negative (TN): The number of correctly predicted negative instances.

False Positive (FP): The number of incorrectly predicted positive instances (also known as Type I error). False Negative (FN): The number of incorrectly predicted negative instances (also known as Type II error).

The performance metrics are as follows:

Accuracy: The total accuracy of the model's predictions is measured by accuracy.

Accuracy = (TP + TN) / (TP + TN + FP + FN)

Precision: Precision is the ability of a model to correctly identify positive examples among all cases that were expected to be positive.

Precision = TP / (TP + FP)

Recall: The model's capacity to accurately identify every occurrence of positivity is measured by recall.

Recall = TP / (TP + FN)

F1 score: F1 score is the combination of both precision and recall which gives a fair evaluation of the model's performance.

F1 score = 2 * (Precision * Recall) / (Precision + Recall) The news articles are given labels in the dataset of 0 and 1, where 0 denotes false news and 1 denotes actual news.

### A. Logistic regression

 Copyright @ 2023 Author

The LR was implemented using two preprocessing techniques, count and TF-IDF vectorizer. The LR model has achieved an accuracy of 98.2%. The precision for detecting the fake news is 97.8% and for real news it is 98.5%. The recall for detecting the fake news is 98.5% and real news is 97.8%. Similarly, the F1score for detecting fake news is 98.1% and for real news 98.2%.

The confusion matrix for LR model is shown below in FIGURE 1.

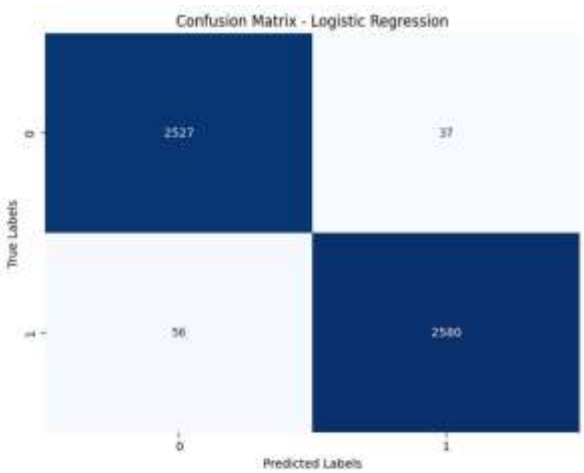


FIGURE 1. confusion matrix for LR

B. Random forest

RF was implemented same way as LR using count and TFIDF vectorizer. The RF model has achieved an accuracy of 92.3%. The precision for detecting the fake news is 90.9% and for real news it is 93.5%. The recall for detecting the fake news is 93.6% and real news is 90.9%. Similarly, the F1-score for detecting fake news is 92.2% and for real news 92.2%. The confusion matrix for RF model is shown below in FIGURE 2.

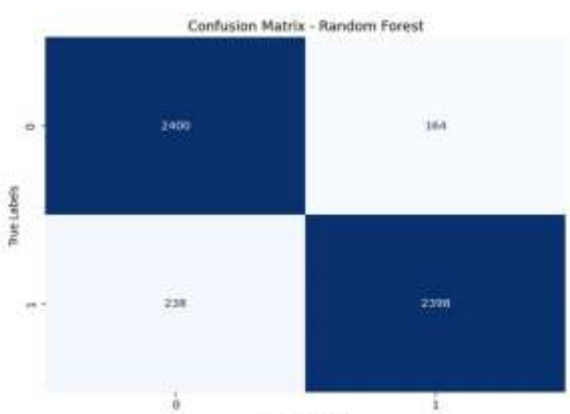


FIGURE 2. confusion matrix for RF

C. Passive aggressive

Utilizing the Count Vectorizer and TF-IDF Vectorizer, two pre-processing methods, the PA model was developed. The PA model has achieved an accuracy of 98.4%. The precision for detecting the fake news is 98.0% and for real news it is 98.7%. The recall for detecting the fake news is 98.7% and real news is 98.1%. Similarly, the F1-score for detecting fake news is 98.3% and for real news 98.4%. The confusion matrix for PA model is shown below in FIGURE 3.

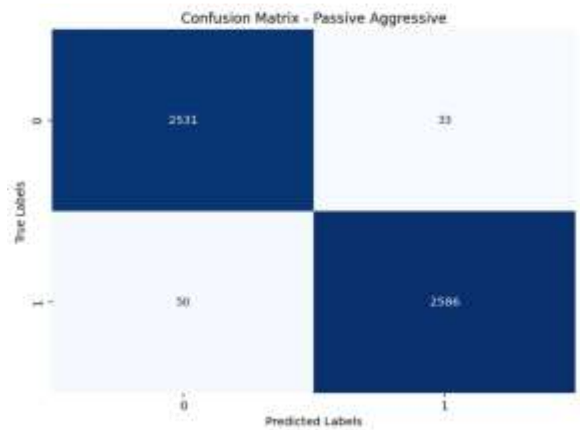


FIGURE 3. confusion matrix for PA

D. Voting classifier

The VC has chosen the “random forest” model as the best model for the dataset. The VC model has achieved an accuracy of 98.4%. The precision for detecting the fake news is 98.0% and for real news it is 98.7%. The recall for detecting the fake news is 98.7% and real news is 98.1%. Similarly, the F1-score for detecting fake news is 98.3% and for real news 98.4%. The confusion matrix VC model is shown below in FIGURE 4.

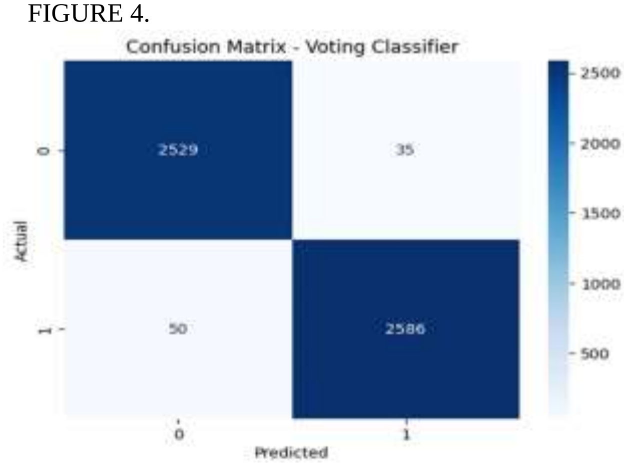


FIGURE 4. confusion matrix for VC


Copyright @ 2023 Author

TABLE 1. Summary of the performance metrics for the machine learning models:

| Metrics\Models | Logistic Regression | Random Forest | Passive Aggressive | Voting Classifier |
|---|---|---|---|---|
| Accuracy | 0.982 | 0.922 | 0.981 | 0.983 |
| Precision (fake news) | 0.978 | 0.909 | 0.980 | 0.980 |
| Precision (real news) | 0.985 | 0.935 | 0.987 | 0.986 |
| Recall (fake news) | 0.985 | 0.936 | 0.987 | 0.986 |
| Recall (real news) | 0.978 | 0.909 | 0.981 | 0.981 |
| F1–score (fake news) | 0.981 | 0.922 | 0.983 | 0.983 |

## VI. CONCLUSION AND FUTURE SCOPE

In conclusion, using ensemble models to detect false news has produced encouraging results in terms of increasing accuracy and reliability. Ensemble models can generate more robust conclusions by merging numerous separate models, each with its strengths and flaws. With the use of hard voting strategy, the voting classifier made the final prediction model.

Despite the positive outcomes this study showed, many changes can be made for future enhancements such as taking a dataset that is bigger and more relevant to real-time news articles. Other ML algorithms can be employed to increase the model prediction accuracy in place of the ones utilised in this study. The ensemble model may be enhanced to become more reliable and accurate by including deep learning techniques such as Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), and boosting algorithms such as CatBoost (Category Boosting), and AdaBoost (Adaptive Boosting).

## REFERENCES

[1] V. Pérez-Rosas, B. Kleinberg, A. Lefevre, and R. Mihalcea, ‘‘Automatic detection of fake news,’’ 2017, arXiv:1708.07104.

[2] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, ‘‘Fake news detection on social media: A data mining perspective,’’ ACM SIGKDD Explor. Newslett., vol. 19, no. 1, pp. 22–36, 2017.

[3] K. Sharma, F. Qian, H. Jiang, N. Ruchansky, M. Zhang, and Y. Liu, ‘‘Combating fake news: A survey on identification and mitigation techniques,’’ ACM Trans. Intell. Syst. Technol., vol. 10, no. 3, pp. 1–42, May 2019.

[4] S. Vosoughi, D. Roy, and S. Aral, ‘‘The spread of true and false news online,’’ Science, vol. 359, pp. 1146–1151, May 2018.

[5] V. Agarwal, H. P. Sultana, S. Malhotra, and A. Sarkar, ‘‘Analysis of classifiers for fake news detection,’’ Proc. Comput. Sci., vol. 165, pp. 377–383, Jan. 2019.

[6] S. Kumar, S. Kumar, P. Yadav, and M. Bagri, ‘‘A survey on analysis of fake news detection techniques,’’ in Proc. Int. Conf. Artif. Intell. Smart Syst. (ICAIS), Mar. 2021, pp. 894–899.

[7] X. Zhou and R. Zafarani, ‘‘A survey of fake news: Fundamental theories, detection methods, and opportunities,’’ 2018, arXiv:1812.00315.

[8] J. C. S. Reis, A. Correia, and F. Murai, A. Veloso, and F. Benevenuto, ‘‘Supervised learning for fake news detection,’’ IEEE Intell. Syst., vol. 34, no. 2, pp. 76–81, Mar./Apr. 2019

[9] X. Dong, U. Victor, S. Chowdhury, and L. Qian, ‘‘Deep twopath semi supervised learning for fake news detection,’’ 2019, arXiv:1906.05659.

[10] Roy, K. Basak, A. Ekbal, and P. Bhattacharyya, ‘‘A deep ensemble framework for fake news detection and classification,’’ 2018, arXiv:1811.04670


Copyright @ 2023 Author