

# Detection of Fake and Real Messages using Machine Learning Techniques

Akshatha T, Akshatha S.A

<sup>1</sup>Professor, Dept. of Artificial Intelligence and Data Science Engineering, Bearys Institute of Technology, Mangalore, Karnataka

<sup>2</sup>Professor, Dept. of Computer Science and Engineering, Bearys Institute of Technology, Mangalore, Karnataka

\*\*\*

**Abstract** - The objective of this paper is to identify real or fake message by using machine learning. Machine learning algorithms play a crucial role in differentiating between authentic and false news articles. Here we have used Logistic Regression, Gradient Boosting Classifier (GBC), and Random Forest Classifier because of their effectiveness in handling text-based classification tasks. Two datasets containing 1,000 real and fake messages are preprocessed and then split into training and testing datasets. The testing dataset is used to evaluate the trained models. After four models were trained classification report was generated and manual test was performed to check whether message is real or fake.

**Key Words:** Machine learning, fake and real news, Logistic Regression, Gradient Boosting Classifier (GBC), Random Forest Classifier.

## 1. INTRODUCTION

Detecting real and fake news is crucial to preventing misinformation, safeguarding democracy, and minimizing negative societal impacts. False information can deceive people, influence opinions, and cause unnecessary fear, particularly in areas like politics, health, and finance. It also erodes trust in media and institutions while enabling fraud and cybercrimes. By distinguishing between genuine and misleading news, we can ensure the spread of accurate information, promote critical thinking, and help individuals make well-informed decisions based on credible sources. There exist various methods of detecting fake and original news.

### a. Manual Fact-Checking

One of the most effective ways to verify information is through manual fact-checking. Websites like Snopes, PolitiFact, and FactCheck.org analyze claims using credible sources and expert evaluations. Additionally, individuals can compare the information with well-established news sources to determine its reliability. If a message lacks confirmation from trusted organizations, it is likely misleading or false.

### b. AI and Machine Learning-Based Detection

Artificial intelligence (AI) and machine learning (ML) play a key role in identifying misinformation. Natural Language

Processing (NLP) examines text structure, sentiment, and word patterns to detect inconsistencies. Deep learning models analyze large datasets to recognize characteristics common in fake news. These automated systems enhance the speed and accuracy of misinformation detection compared to manual methods.

### c. Metadata and Source Verification

Evaluating the source of a message is essential in determining its authenticity. Trusted sources are typically transparent about their authorship and maintain a strong track record of accuracy. Examining metadata, including publication date, author details, and domain credibility, helps in assessing reliability. Websites that provide little information about their origins or frequently publish exaggerated content may be unreliable.

### d. User Behavior Analysis

The spread of misinformation is often linked to bots and coordinated campaigns. By analyzing social media activity, unusual trends can be detected, such as new accounts repeatedly posting the same message. Another sign of misinformation is abnormal engagement, where content receives many shares but limited meaningful discussion. Identifying these patterns helps in recognizing and limiting the spread of fake messages.

### e. Image and Video Forensics

Misleading messages frequently contain altered images or videos to distort facts. Reverse image search tools, like Google Reverse Image Search and TinEye, help verify whether an image has been edited or used out of context. Additionally, deepfake detection technology examines inconsistencies in video elements, such as unnatural facial movements or lighting, which may indicate digital manipulation.

In our research we have used machine learning methods to identify the fake and original news. The methodology involves first collecting the data, classifying the data, data preprocessing, feature extraction, Model training and Evaluation, Model testing and then result.

## 2. PROPOSED METHODOLOGY:

### Data loading and Labeling:

The fake and real news datasets were downloaded from GitHub in .CSV format, with each dataset having a size of approximately 60 MB. The number of real and fake news messages in the dataset are balanced, with an equal number of real and fake articles. The 'real.csv' and 'fake.csv' datasets are stored in a folder and then loaded into the program using pandas by specifying the correct file paths on the local system. These two datasets are then stored in two variables, true and fake, as pandas DataFrames, which allow for manipulation, processing, and analysis of the data. A new column named 'label' is created in each dataset, where the value '0' is assigned to the fake news dataset and '1' is assigned to the real news dataset.

### Data Preprocessing:

In the data preprocessing step, the first operation is to concatenate the two datasets using concat(), creating a single dataset named 'news' that combines both the fake and real news datasets. In the next step, the 'news' dataset is cleaned by eliminating missing values and dropping columns that are not required. In further purification, the wordopt function converts the text to lowercase and removes URLs, HTML tags, punctuation, numbers, and newline characters in the text column of the 'news' data frame. The processed text and labels are stored in two variables, x and y.

### Feature Extraction:

We use the train\_test\_split function to divide the data into training and testing sets. The dataset is split into x\_train, x\_test, y\_train, and y\_test, with half of the data (50%) allocated to the test set (test\_size=0.5). Next, the TfidfVectorizer is imported to transform the text data (x\_train and x\_test) into numerical feature vectors using the Term Frequency-Inverse Document Frequency (TF-IDF) method.

### Model training and Evaluation:

The models were developed using four algorithms: Logistic Regression, Decision Tree Classifier, Random Forest Classifier, and Gradient Boosting Classifier. These algorithms were trained on the text data transformed using TF-IDF to evaluate their performance in text classification. Once each model was trained, we produced a classification report to assess their performance.

### Model testing:

The Logistic Regression, Gradient Boosting Classifier (GBC) and Random Forest Classifier are used to predict whether news is fake or genuine, and it also allows for manual testing by inputting a new article.

## 3.Result:

```
true.head()
```

	title	text	subject	date
0	As U.S. budget fight looms, Republicans flip t...	WASHINGTON (Reuters) - The head of a conservat...	politicsNews	December 31, 2017
1	U.S. military to accept transgender recruits o...	WASHINGTON (Reuters) - Transgender people will...	politicsNews	December 29, 2017
2	Senior U.S. Republican senator: 'Let Mr. Muell...	WASHINGTON (Reuters) - The special counsel inv...	politicsNews	December 31, 2017
3	FBI Russia probe helped by Australian diplomat...	WASHINGTON (Reuters) - Trump campaign adviser ...	politicsNews	December 30, 2017
4	Trump wants Postal Service to charge 'much mor...	SEATTLE/WASHINGTON (Reuters) - President Donal...	politicsNews	December 29, 2017

Fig 1: True news dataset

```
fake.head()
```

	title	text	subject	date
0	Donald Trump Sends Out Embarrassing New Year...	Donald Trump just couldn't wish all Americans ...	News	December 31, 2017
1	Drunk Bragging Trump Staffer Started Russian ...	House Intelligence Committee Chairman Devin Nu...	News	December 31, 2017
2	Sheriff David Clarke Becomes An Internet Joke...	On Friday, it was revealed that former Milwauk...	News	December 30, 2017
3	Trump Is So Obsessed He Even Has Obama's Name...	On Christmas day, Donald Trump announced that ...	News	December 29, 2017
4	Pope Francis Just Called Out Donald Trump Dur...	Pope Francis used his annual Christmas Day mes...	News	December 25, 2017

Fig 2: Fake news dataset

```
manual_testing(news_article)

'\n\nLR Prediction: it is a fake news

\nGBC Prediction: it is a fake news

\nRFC Prediction: it is a fake news'
```

Fig 3: Final result

New\_article=str(input()) function will take the input from the user and manual\_testing(news\_article) function displays the output after prediction. In figure-3 we can see the result obtained from three different trained models. LR, GBC and RFC are identified the input message as fake.

## 4. CONCLUSIONS

The code successfully handles basic model training and testing, but there are multiple opportunities for improvement. These include addressing edge cases, optimizing hyper parameters, comparing model performance, and incorporating more advanced text preprocessing techniques. By refining these aspects, the

model's overall performance and reliability can be significantly improved.

## REFERENCES

- 1) Michael Crawford, Taghi M, Khoshgoftaar, Joseph D. Prusa Aaron N. Richter and Hamzah Al Najada "Study of audit spam identification utilizing AI strategies" *Journal of Big Data* (2015) 2:23, Springer diary, DOI 10.1186/s40537-015-0029-9.
- 2) Kolli Shivagangadhar, Sagar H, Sohan Sathyan, Vanipriya C.H " Fraud Detection in Online Reviews utilizing Machine Learning Techniques" *International Journal of Computational Engineering Research*, Volume, 05, Issue 05, May - 2015, ISSN (e): 2250 - 3005.
- 3) Elmurngi, E., Gherbi, A. (2018). Fake reviews detection on movie reviews through sentiment analysis using supervised learning techniques. *International Journal on Advances in Systems and Measurements*, 11(1): 196-207.
- 4) Karupusamy, S., Maruthachalam, S., Mayilswamy, S., Sharma, S., Singh, J., Lorenzini, G. (2021). Efficient computation for localization and navigation system for a differential drive mobile robot in indoor and outdoor environments. *Revue d'Intelligence Artificielle*, 35(6): 437-446. <https://doi.org/10.18280/ria.350601>.
- 5) Al-Adhaileh, M.H., Alsaade, F.W. (2022). Detecting and analysing fake opinions using artificial intelligence algorithms. *Intelligent Automation & Soft Computing*, 32(1): 644-655. <https://doi.org/10.32604/iasc.2022.02225>