



# AI Based Fake News Detection System Using Machine Learning And NLP

Name of 1<sup>st</sup> Author : Reenu Bala, Name of 2<sup>nd</sup> Author : Ms Bhawna Sharma

Designation of 1<sup>st</sup> Author : M.Tech Student, Designation of 2<sup>nd</sup> Author : HOD of CSE Dept.

Department of 1<sup>st</sup> Author : CSE

City : Pataudi, State : Haryana

## Abstract

In the digital age, spreading false information online has become a major problem that affects public opinion and has negative social, political, and economic effects. Because so much information is produced every day, it is impossible and inefficient to manually identify bogus news. This work introduces an artificial intelligence (AI)-based method for detecting fake news that automatically classifies news articles as authentic or fraudulent using machine learning and natural language processing (NLP) techniques. Text preprocessing, feature extraction using Term Frequency–Inverse Document Frequency (TF-IDF), and classification using machine learning methods like Random Forest and Naïve Bayes are all done by the suggested system. The suggested method detects bogus news stories with excellent accuracy and dependability, according to experimental results. News organisations and social media sites can use the system to help stop the spread of misleading information.

**Keyword:** Artificial Intelligence, Machine Learning, NLP, TF-IDF, Social Media, Fake News Detection

## 1. Introduction

The way that information is produced and consumed has changed dramatically due to the quick development of social media and digital news outlets. These platforms make it simple to obtain news, but they have also grown to be a significant source of false or misleading information. The term "fake news" describes purposefully false or misleading information that is passed off as news in order to trick the reader. Political instability, societal fear, and a decline in media trust can result from the quick spread of fake news. Due to the enormous amount of online data, human verification is the foundation of traditional false news detection techniques, which is both time-consuming and ineffective. Thus, artificial intelligence (AI)-based automated fake news detection systems are needed. When artificial intelligence (AI) and natural language processing (NLP) are coupled, news information can be effectively classified by analysing linguistic traits and textual patterns. This paper suggests a machine learning and natural language processing (NLP)-based AI-based fake news detecting system. The following are this study's main contributions: The creation of an automated system for identifying false information

Using strong NLP preprocessing methods Comparative evaluation of several machine learning classifiers

## 2. Review of literature

Using machine learning and deep learning approaches, some researchers have put forth strategies for identifying bogus news. Shu et al. identified bogus news using linguistic and social context factors, however their method required a lot of metadata. Using TF-IDF features, Kumar et al. employed SVM and Naïve Bayes classifiers, although their accuracy on complicated datasets was limited. Although deep learning-based methods like CNN and LSTM have demonstrated encouraging outcomes, they need substantial computer resources and datasets.

The majority of current systems have either great computing complexity or poor accuracy. An effective and lightweight AI-based fake news detection algorithm that attains high accuracy at a fair computational cost is thus required.

### 2.1 Evaluation of Current Research

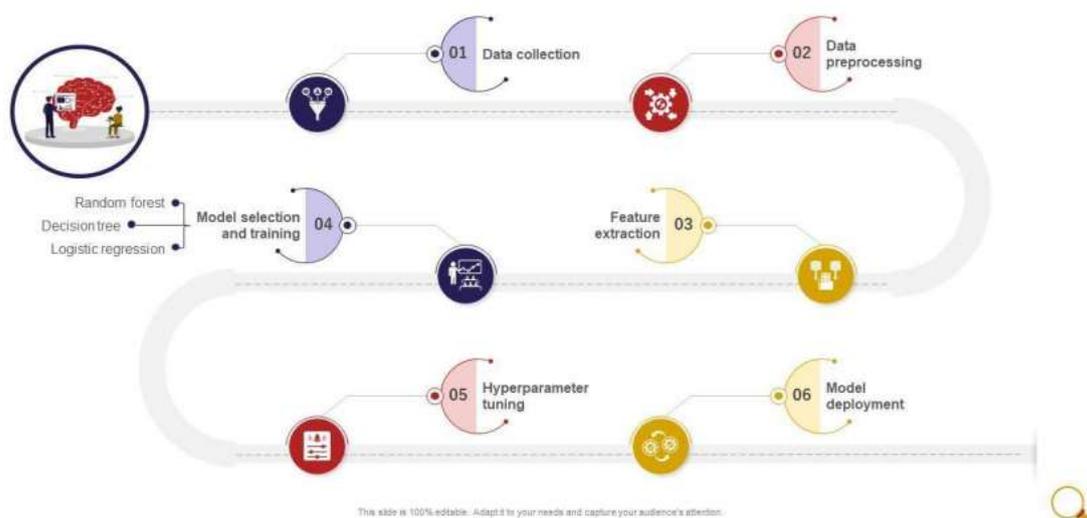
Numerous academics have used various methods, including machine learning and deep learning, to work on the detection of fake news. The majority of current research employs algorithms such as Naïve Bayes, Random Forest and Logistic Regression are used to determine whether news is authentic or fraudulent. On clean, tiny datasets, these techniques have demonstrated good accuracy. Nevertheless, the current research has a number of shortcomings. Large, well-labeled datasets are essential to most models, but they are not always accessible. Many systems only work well on particular datasets; when evaluated on fresh or real-world data, they fall short. Additionally, the majority of studies do not provide regional or multilingual data and primarily concentrate on English language content.

## 3. The suggested system

The structured workflow of the suggested AI-based fake news detection system includes feature extraction, data collection, preprocessing, model training, and categorisation.

### Machine learning workflow for fake news detection

This slide showcases steps for machine learning based fake detection that are data collection, data preprocessing, feature extraction, model training, hyperparameter tuning etc.



This slide is 100% editable. Adapt it to your needs and capture your audience's attention.

## 4. Research techniques

### 4.1 Compiling datasets

The study's dataset comprised news stories with labels that were classified as either real or fake. The information came from publicly accessible datasets that are frequently utilised in studies on false news. Every record had news items along with the label that went with them. There were training and testing sets in the dataset.

### 4.2 Text Preprocessing

The following NLP preprocessing approaches were used to increase classification accuracy:

- Changing the text to lowercase
- Tokenization
- Elimination of stop words
- Stemming, or lemmatization

The noise and dimensionality are decreased by these procedures.

### 4.3 Extraction of Features

The TF-IDF approach is used to transform the processed text into numerical characteristics. While lessening the impact of commonly used phrases, TF-IDF draws attention to key concepts.

### 4.4 Algorithms for Classification

Three machine learning algorithms are used:

- **Naïve Bayes:** A probabilistic classifier suitable for text classification
- **Logistic Regression:** A supervised learning algorithm for binary classification
- **Random Forest:** An ensemble learning method used to improve predictive accuracy and control over-fitting

The models are trained on the training dataset and tested on unseen data.

### 4.5 Assessment of Performance

The performance of the system is evaluated using standard metrics:

- Accuracy
- Precision
- Recall
- F1-score

These metrics provide a comprehensive assessment of the system.

## 5. Algorithm description

This section presents the algorithmic steps followed by the proposed AI-based fake news detection system. The algorithm processes textual news data, extracts meaningful features using Natural Language Processing techniques, and applies machine learning classifiers to determine whether the news is real or fake.

### Algorithm 1: AI-Based Fake News Detection

#### Input:

- News dataset  $D = \{d_1, d_2, \dots, d_n\}$
- Corresponding labels  $L = \{\text{real}, \text{fake}\}$

#### Output:

- Classified news label (Real or Fake)

#### Steps:

1. Collect labeled news articles from the dataset.
2. Preprocess the text by converting it to lowercase, removing stop-words, and applying stemming or lemmatization.
3. Transform the cleaned text into numerical features using TF-IDF.
4. Split the dataset into training and testing sets.
5. Train machine learning models (Naïve Bayes, Logistic Regression and Random Forest) using the training data.
6. Test the trained models on the testing data.
7. Evaluate the performance using accuracy, precision, recall, and F1-score.
8. Output the predicted class label for each news article.

## 6. Pseudocode for suggested system

The pseudocode below describes the logical flow of the fake news detection system.

Algorithm: Fake News Detection System

Input: News Dataset D

Output: Classification result (Real / Fake)

Begin

Load dataset D

For each news article d in D do

Convert text to lowercase

Tokenize the text

Remove stop-words

Apply stemming or lemmatization

End For

Apply TF-IDF to extract feature vectors

Split dataset into Training\_Set and Testing\_Set

Train Naïve Bayes model using Training\_Set

Train Logistic Regression model using Training\_Set

Train Random Forest model using Training\_Set

For each instance t in Testing\_Set do

Predict label using trained models

End For

Calculate Accuracy, Precision, Recall, and F1-score

Display classification results

End

## 7. Results and discussion

The experimental results indicate that all three classifiers are effective in detecting fake news. Performance was evaluated based on accuracy, precision, recall, and F1-score.

### Algorithm

Naïve Bayes

Logistic Regression

Random Forest

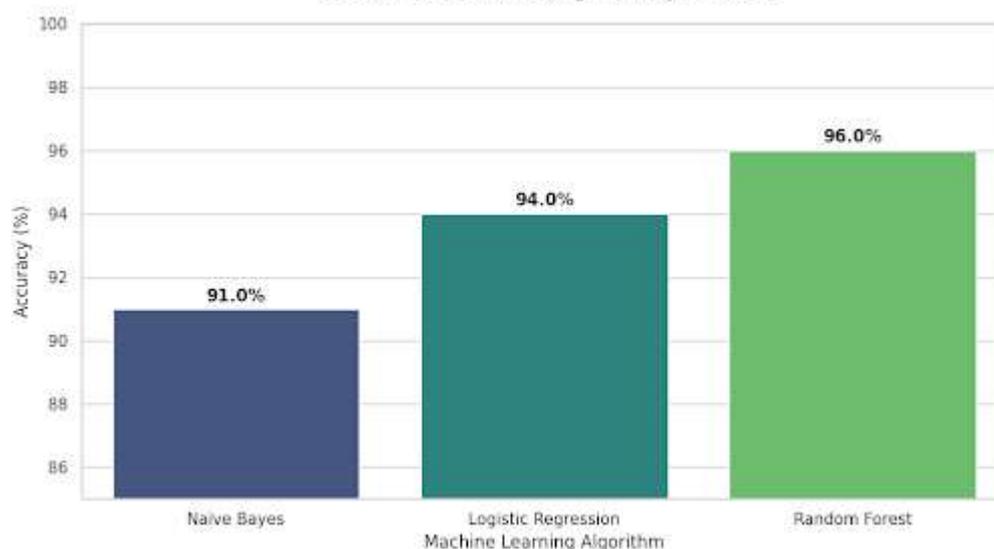
### Accuracy

91%

94%

96%

**Classifier Accuracy Comparison**



The results confirm that Logistic Regression and Random Forest provide superior performance for this dataset, with Random Forest achieving the highest overall accuracy.

## 8. Conclusion and future work

Using machine learning and natural language processing, this study demonstrated an AI-based fake news identification system. The algorithm effectively pulls meaningful elements from textual input, preprocesses it, and determines whether news stories are authentic or fraudulent. Experimenting

According to evaluation, the Random Forest classifier attains the maximum accuracy. Future research can incorporate deep learning models like CNN, BERT, and LSTM to increase detection accuracy even more. It is also possible to investigate language support and real-time fake news identification.

## 9. References

- [1] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake news detection on social media: A data mining perspective," *ACM SIGKDD Explorations Newsletter*, vol. 19, no. 1, pp. 22–36, 2017.
- [2] X. Zhou and R. Zafarani, "A survey of fake news: Fundamental theories, detection methods, and opportunities," *ACM Computing Surveys*, vol. 53, no. 5, pp. 1–40, 2020.
- [3] S. Kumar and N. Shah, "False information on web and social media: A survey," *IEEE Access*, vol. 6, pp. 69486–69502, 2018.
- [4] J. Thorne and A. Vlachos, "Automated fact checking: Task formulations, methods and future directions," in *Proc. 27th Int. Conf. on Computational Linguistics (COLING)*, Santa Fe, NM, USA, 2018, pp. 3346–3359.
- [5] D. M. Lazer *et al.*, "The science of fake news," *Science*, vol. 359, no. 6380, pp. 1094–1096, 2018.
- [6] T. Q. Nguyen, T. T. Nguyen, and H. T. Nguyen, "Applying machine learning algorithms for fake news detection," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 6, pp. 545–551, 2020.
- [7] A. Jain and A. Kasbe, "Fake news detection," in *Proc. IEEE Int. Students' Conf. Electrical, Electronics and Computer Science (SCEECS)*, Bhopal, India, 2018, pp. 1–5.
- [8] M. Granik and V. Mesyura, "Fake news detection using naive Bayes classifier," in *Proc. IEEE First Ukraine Conf. Electrical and Computer Engineering (UKRCON)*, Kyiv, Ukraine, 2017, pp. 900–903.
- [9] R. K. Kaliyar, A. Goswami, and P. Narang, "FakeBERT: Fake news detection in social media with a BERT-based deep learning approach," *Multimedia Tools and Applications*, vol. 80, no. 8, pp. 11765–11788, 2021.
- [10] Y. Wang, "Liar, liar pants on fire: A new benchmark dataset for fake news detection," in *Proc. 55th Annu. Meeting of the Assoc. for Computational Linguistics (ACL)*, Vancouver, BC, Canada, 2017, pp. 422–426.

[11] I. Rish, “An empirical study of the naive Bayes classifier,” in *Proc. IJCAI Workshop on Empirical Methods in Artificial Intelligence*, Seattle, WA, USA, 2001, pp. 41–46.

[12] C. Cortes and V. Vapnik, “Support-vector networks,” *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

[13] N. Tabassum, S. Chakraborty, and A. Das, “Transformer-based fake news detection in low-resource languages,” in *Proc. 4th Workshop on Dravidian Languages in Technology (DravidianLangTech)*, ACL, 2024, pp. 245–252.

[14] R. Anggrainingsih, A. Suryani, and M. R. Faisal, “Transformer-based models for combating rumours on social media: A comparative study,” *Artificial Intelligence Review*, vol. 57, no. 3, pp. 1–29, 2024.

[15] J. Rout, S. Dash, and D. Mishra, “An enhanced attention-based transformer model for fake news detection,” *AI*, vol. 5, no. 3, pp. 623–641, 2025.

[16] E. Papageorgiou, I. Kanellos, and G. Stamou, “Harnessing large language models and deep neural networks for fake news detection,” *Information*, vol. 16, no. 4, Art. no. 297, 2025.

