

# Fake News Detection Using Artificial Intelligence

Dr.G.Jayamurgan 

Professor, Department of CSE

Sengunthar Engineering College (Autonomous), Tiruchengode, India

[gjayamurugan.cse@scteng.co.in](mailto:gjayamurugan.cse@scteng.co.in)

<https://orcid.org/0009-0004-7797-3094>

Gayathri K, Sujith Priya S

UG Students, Department of CSE

Sengunthar Engineering College (Autonomous), Tiruchengode, India

[gayu29gayathri@gmail.com](mailto:gayu29gayathri@gmail.com), [sujithpriya39@gmail.com](mailto:sujithpriya39@gmail.com)



## Publication History

Manuscript Reference: IRJCS/RS/Vol.13/Issue03/CSMR26.MRCS10111

Research Article | Open Access | Double-Blind Peer Reviewed Article ID: IRJCS/RS/Vol.13/Issue03/CSMR26.MRCS10111

Received: 30, January 2026, Revised: 13, February 2026, Accepted: 28 February 2026 Published Online: 25 March 2026

<https://www.irjcs.com/volumes/Vol13/iss-03/32.CSMR26.MRCS10111.pdf>

**Article Citation:** Dr. Jayamurgan, Gayathri, Sujith Priya (2026), Fake News Detection Using Artificial Intelligence, IRJCS: International Research Journal of Computer Science, Volume 13, Issue 03 of 2026 pages 286-291

**Doi:** > <https://doi.org/10.26562/irjcs.2026.v1303.32> **BibTeX Key** Dr.Jayamurgan@2026Fake

**Orcid:** <https://orcid.org/0009-0004-9398-7488>

IRJCS papers should be cited as IRJCS (International Research Journal of Computer Science, AM Publications, India 2026, ISSN 2393-9842, <https://doi.org/10.26562/irjcs.2025.v1303.32> The journal's official abbreviation is IRJCS.

About the License: Copyright © 2026 copyright by the authors. This article is an open access and license under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** In recent years, due to the booming development of online social networks, fake news for various commercial and political purposes has been appearing in large numbers and widespread in the online world. With deceptive words, online social network users can get infected by this online fake news easily, which has brought about tremendous effects on the offline society already. An important goal in improving the trustworthiness of information in online social networks is to identify the fake news timely. This paper aims at investigating the principles, methodologies and algorithms for detecting fake news articles, creators and subjects from online social networks and evaluating the corresponding performance. Information preciseness on Internet, especially on social media, is an increasingly important concern, but web-scale data hampers, ability to identify, evaluate and correct such data, or so called "fake news," present in these platforms. In this paper, we propose a method for "fake news" detection and ways to apply it on Face book, one of the most popular online social media platforms. This method uses NaiveBayes classification model to predict whether a post on Facebook will be labeled as real or fake. The results may be improved by applying several techniques that are discussed in the paper. Received results suggest, that fake news detection problem can be addressed with machine learning methods.

**Keywords:** Fake News Detection, Artificial Intelligence, Machine Learning, Natural Language Processing, Naive Bayes.

## I. INTRODUCTION

This days" fake news is creating different issues from sarcastic articles to a fabricated news and plan government propaganda in some outlets. Fake news and lack of trust in the media are growing problems with huge ramifications in our society. Obviously, a purposely misleading story is "fake news" but lately blathering social media's discourse is changing its definition. Some of them now use the term to dismiss the facts counter to their preferred viewpoints. The importance of disinformation within American political discourse was the subject of weighty attention, particularly following the American president election. The term 'fake news' became common parlance for the issue, particularly to describe factually incorrect and misleading articles published mostly for the purpose of making money through page views. In this paper, it is sought to produce a model that can accurately predict the likelihood that a given article is fake news. Face book has been at the epicenter of much critique following media attention. They have already implemented a feature to flag fake news on the site when a user sees it, they have also said publicly they are working on to distinguish these articles in an automated way. Certainly, it is not an easy task. A given algorithm must be politically unbiased– since fake news exists on both ends of the spectrum and also give equal balance to legitimate news sources on either end of the spectrum. In addition, the question of legitimacy is a difficult one. However, in order to solve this problem, it is necessary to have an understanding on what Fake News is. Later, it is needed to look into how the techniques in the fields of machine learning, natural language processing help us to detect fake news. The objective of this project is to examine the problems and possible significance related with the spread of fake news. We will be working on different fake news dataset in which we will apply different machine learning algorithms to train the data and test it to find which news is the real news or which one is the fake news. As the fake news is a problem that is heavily affecting society and our perception of not only the media but also facts and opinions themselves. By using the artificial intelligence and the machine learning, the problem can be solved as we will be able to mine the patterns from the data to maximize well defined objectives. So, our focus is to find which machine learning algorithm is best suitable for what kind of text data set. Also, which dataset is better for finding the accuracies as the accuracies directly depends on the type of data and the amount of data.

The more the data, more are your chances of getting correct accuracy as you can test and train more data to find out your results. However, on the other hand, social media provides the ideal place for the creation and spread of fake news. Fake news can become extremely influential and has the ability to spread exceedingly fast. With the increase of people using social media, they are being exposed to new information and stories every day. Misinformation can be difficult to correct and may have lasting implications. Therefore we need to make sure the news we read are correct and real so, there should be fake news detection in our applications so that we don't fall for those fake news spreading around us.

## **II. LITERATURE REVIEW**

The rapid growth of social media platforms has significantly increased the spread of misinformation and fake news across the internet. Researchers have proposed various techniques to detect fake news using machine learning, natural language processing (NLP), and data mining methods. These approaches analyze textual patterns, user behavior, and information propagation to identify misleading content. Conroy, Rubin, and Chen<sup>1</sup> introduced an automatic deception detection approach that uses linguistic features and machine learning algorithms to identify fake news articles. Their research focused on analyzing writings styles and linguistic cues that often appear in deceptive content. They categorized fake news detection methods into two major groups: linguistic-based detection and network-based detection. Linguistic methods analyze the textual content of news articles, while network-based approaches examine how news spreads through social networks. Chen, Conroy, and Rubin<sup>2</sup> further investigated the problem of misleading online content through click bait detection. Click bait headlines are designed to attract user attention but often contain exaggerated or misleading information. Their research demonstrated that analyzing headline structures and textual features can effectively detect misleading content. They proposed classification models that combine textual and contextual features to improve fake news detection accuracy. Najafabadi et al.<sup>3</sup> explored the application of deep learning techniques in big data analytics. Their study highlighted how deep neural networks can learn complex representations from large datasets. They emphasized that deep learning models can analyze large-scale textual data and automatically extract meaningful patterns. These capabilities make deep learning suitable for fake news detection, where large amounts of textual data must be analyzed efficiently.

Tacchini et al.<sup>4</sup> proposed a fake news detection system that analyzes user interaction patterns on social media platforms such as Facebook. Instead of focusing only on the textual content, their approach considered the behavior of users who interact with posts. Their findings showed that fake news often spreads within specific communities and receives rapid engagement through likes and shares. Using machine learning classification methods, their model achieved high accuracy in distinguishing fake news from legitimate information.

Shao et al.<sup>5</sup> investigated the role of social bots in spreading misinformation across social media platforms. Their research analyzed millions of online messages and discovered that automated accounts play a major role in the early spread of fake news. Bots are capable of rapidly sharing misleading content, which increases the visibility of fake information. Their study suggests that identifying bot activity is an important step in controlling the spread of fake news.

Parikh and Atrey<sup>6</sup> conducted a comprehensive survey on media-rich fake news detection, focusing on news articles that contain multimedia elements such as images and videos. Their research emphasized that fake news often includes manipulated images or misleading visual content to support false claims. They suggested that combining textual analysis with multimedia verification techniques can significantly improve detection accuracy.

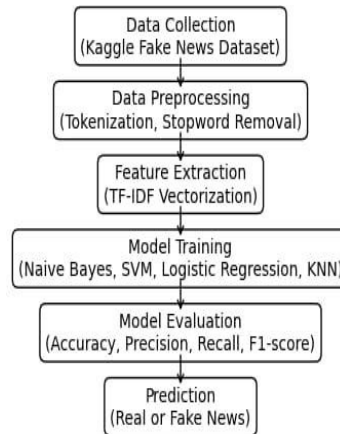
Helmstetter and Paulheim proposed a weakly supervised learning approach for detecting fake news on Twitter. Their method automatically collects large datasets of tweets and labels them based on the credibility of the source. Although the dataset may contain noisy labels, their machine learning model was able to achieve a high F1-score. This study demonstrated that large-scale datasets can improve the effectiveness of fake news detection systems.

Shu et al.<sup>8</sup> presented a data mining perspective on fake news detection, highlighting the importance of integrating content analysis with social context features. Their research proposed a framework that combines textual features, user profiles, and network propagation patterns. The results showed that incorporating social context information significantly improves the performance of fake news detection models.

Although many studies have proposed effective methods for fake news detection, several challenges still remain. One major challenge is the availability of high-quality labeled datasets required for training machine learning models. Another challenge is the rapid evolution of misinformation strategies, where fake news creators continuously modify their content to avoid detection systems. Therefore, developing intelligent and adaptive fake news detection systems remains an important research area. The proposed system in this study addresses these challenges by applying machine learning algorithms and natural language processing techniques to analyze textual features of news articles and classify them as real or fake.

## **III. PROPOSED METHODOLOGY**

The proposed fake news detection system uses machine learning and natural language processing techniques to automatically classify news articles as real or fake. The system processes textual data from news articles, extracts meaningful features, and applies classification algorithms to determine the authenticity of the information. The overall methodology consists of several stages including data collection, data preprocessing, feature extraction, model training, and prediction.



### A. Data Collection

The first step in the proposed system is collecting a dataset containing both real and fake news articles. The dataset used in this project is obtained from publicly available sources such as the Kaggle Fake News Dataset, which contains labeled news articles. Each record in the dataset includes information such as the title of the article, textual content, and label indicating whether the news is real or fake. The dataset serves as the training and testing data for the machine learning model. A large dataset improves the accuracy of the detection system because the model can learn patterns from a greater amount of information.

### B. Data Preprocessing

Raw textual data often contains noise such as punctuation, special characters, and stop words that do not contribute to meaningful analysis. Therefore, preprocessing is performed to clean and normalize the dataset.

The preprocessing stage includes the following steps:

1. Lower case Conversion–All text is converted to lowercase to maintain consistency.
2. Removal of Punctuation and Special Characters–Unnecessary symbols are removed.
3. Stop Word Removal – Common words such as “the”, “is”, and “and” are removed because they do not contribute to classification.
4. Tokenization–The text is divided into individual words or tokens.
5. Stemming or Lemmatization – Words are reduced to their base forms to improve text analysis.

These preprocessing techniques help improve the quality of the dataset and enhance the performance of the machine learning model.

### C. Feature Extraction

Machine learning algorithms cannot directly process textual data. Therefore, the cleaned text must be converted into numerical features using feature extraction techniques. In this project, TF-IDF (Term Frequency–Inverse Document Frequency) vectorization is used to transform textual data into numerical vectors. TF-IDF measures the importance of a word in a document relative to the entire dataset. The TF-IDF score is calculated based on two factors:

- Term Frequency (TF): Frequency of a word in a document.
- Inverse Document Frequency (IDF): Importance of the word across all documents.

This method helps identify words that are significant in distinguishing fake news from real news.

### D. Machine Learning Model Training

After feature extraction, the dataset is divided into training and testing sets. The training set is used to train the machine learning model, while the testing set is used to evaluate the model's performance.

Several machine learning algorithms are applied in this project, including:

- Naive Bayes
- Logistic Regression
- Support Vector Machine (SVM)
- K-Nearest Neighbor (KNN)

These algorithms learn patterns from the training data and create a classification model capable of predicting whether a news article is real or fake.

### E. Model Evaluation

The performance of the trained models is evaluated using different evaluation metrics such as:

- Accuracy
- Precision
- Recall
- F1-Score

These metrics help determine the effectiveness of the model in detecting fake news.

## F. Prediction and Classification

In the final stage, the trained model is used to classify new input news articles. When a user enters a news headline or article, the system performs preprocessing and feature extraction on the input text. The processed text is then passed to the trained machine learning model, which predicts whether the news is real or fake. The result is displayed to the user through a graphical user interface, allowing users to verify the authenticity of the news article.

## IV. TECHNOLOGIES USED

The proposed fake news detection system is developed using several technologies and tools that support data processing, machine learning model development, and visualization. These technologies enable efficient analysis of textual data and accurate classification of news articles.

### A. Programming Language

The system is implemented using the Python programming language, which is widely used in machine learning and artificial intelligence applications. Python provides a simple syntax and a large number of libraries that support data analysis, machine learning, and natural language processing.

### B. Machine Learning Libraries

Several Python libraries are used for implementing machine learning algorithms and data processing tasks.

#### 1. Scikit-learn

Scikit-learn is a powerful machine learning library used for implementing classification algorithms such as Naive Bayes, Logistic Regression, Support Vector Machine (SVM), and K-Nearest Neighbor (KNN). It also provides tools for model evaluation, feature extraction, and dataset splitting.

#### 2. Pandas

Pandas is a data analysis library used for handling structured datasets. It allows efficient loading, manipulation, and preprocessing of the news dataset. Pandas provides functions to remove missing values, filter data, and analyze dataset statistics.

#### 3. NumPy

NumPy is a numerical computing library used for performing mathematical operations on arrays and matrices. It is commonly used with machine learning libraries to perform efficient numerical computations.

### C. Natural Language Processing Tools

Natural Language Processing (NLP) techniques are used to analyze and process textual data in the news articles.

#### 1. NLTK (Natural Language Toolkit)

NLTK is a widely used NLP library that provides tools for text preprocessing tasks such as tokenization, stop-word removal, stemming, and lemmatization. These techniques help convert raw textual data into meaningful information that can be used by machine learning models.

#### 2. TF-IDF Vectorizer

TF-IDF (Term Frequency–Inverse Document Frequency) is used to convert textual data into numerical features. It measures the importance of words in a document relative to the entire dataset and helps machine learning algorithms identify key patterns in news articles.

### D. Dataset

The dataset used for this project is obtained from the Kaggle Fake News Dataset, which contains labeled news articles categorized as real or fake. The dataset includes features such as the title of the article, the textual content, and the label indicating the authenticity of the news.

### E. Development Environment

The project is developed using Jupyter Notebook and Visual Studio Code. These development environments provide interactive tools for writing, testing, and debugging Python code. Jupyter Notebook is particularly useful for data analysis and visualization tasks.

### F. Data Visualization Tools

Visualization libraries such as Matplotlib and Seaborn are used to analyze the distribution of real and fake news in the dataset. These tools help generate graphical representations such as bar charts, histograms, and accuracy comparison graphs.

## V. SYSTEM IMPLEMENTATION AND RESULTS

The implementation of the proposed fake news detection system is carried out using Python programming language along with several machine learning and natural language processing libraries. The system processes textual news data, trains classification models, and predicts whether the given news article is real or fake. The implementation includes dataset preparation, preprocessing, feature extraction, model training, and result evaluation.

### A. System Implementation

The system is implemented using Python libraries such as Pandas, NumPy, Scikit-learn, and NLTK. These libraries support data preprocessing, feature extraction, machine learning model development, and result visualization. Initially, the dataset containing real and fake news articles is loaded into the system using the Pandas library. The dataset is then analyzed to check for missing values, duplicates, and inconsistencies. After cleaning the dataset, preprocessing techniques are applied to prepare the textual data for machine learning algorithms. The preprocessing stage includes lowercase conversion, removal of punctuation marks, tokenization, stop-word removal, and stemming. These techniques help remove irrelevant information and improve the quality of the textual dataset.

After preprocessing, the textual data is converted into numerical form using the TF-IDF (Term Frequency–Inverse Document Frequency) feature extraction technique. This method assigns a weight to each word based on its frequency and importance in the dataset. The dataset is then divided into training and testing datasets. The training dataset is used to train machine learning models, while the testing dataset is used to evaluate the model's performance.

Several machine learning algorithms are implemented in the system, including:

- Naive Bayes
- Logistic Regression
- Support Vector Machine(SVM)
- K-Nearest Neighbor(KNN)

Each model learns patterns from the dataset and creates a classification model capable of predicting whether a news article is real or fake.

### B. Prediction Interface

The developed system provides a simple interface where users can enter a news headline or article. The system preprocesses the input text and applies the trained machine learning model to classify the news. The output of the system displays whether the news article is Real News or Fake News, enabling users to verify the authenticity of online information.

### C. Performance Evaluation

The performance of the machine learning models is evaluated using several evaluation metrics. These metrics measure the effectiveness and reliability of the classification model.

The evaluation metrics include:

- Accuracy–Percentage of correctly classified news articles.
- Precision – Measures how many predicted fake news articles are actually fake.
- Recall–Measures how many fake news articles are correctly identified.
- F1-score–Provides a balanced evaluation of precision and recall.

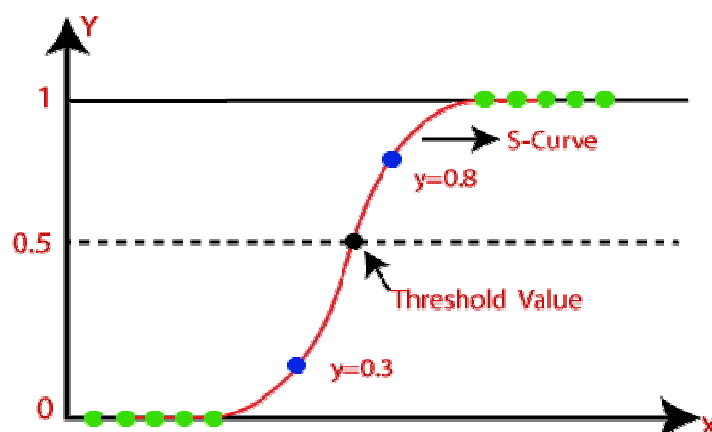
A confusion matrix is used to visualize the classification performance of the model. The confusion matrix consists of four components:

- True Positive (TP)
- True Negative(TN)
- False Positive(FP)
- False Negative(FN)

These values help determine the accuracy of the prediction model.

### D. Experimental Results

The experimental results demonstrate that machine learning algorithms can effectively detect fake news based on textual patterns and linguistic features. Among the algorithms tested, Support Vector Machine and Logistic Regression achieved higher classification accuracy compared to other models. The models were able to identify fake news with high reliability and minimal error rates. The confusion matrix results indicate that the system correctly classifies a large percentage of news articles. The graphical results also show that the trained models perform well in distinguishing between real and fake news articles. The implemented fake news detection system can therefore be used as a practical tool to verify online information and reduce the spread of misinformation on digital platforms.



## VII CONCLUSION

Many people consume news from social media instead of traditional news media. However, social media has also been used to spread fake news, which has negative impacts on individual people and society. In this paper, an innovative model for fake news detection using machine learning algorithms has been presented. This model takes news events as an input and based on twitter reviews and classification algorithms it predicts the percentage of news being fake or real.

In conclusion, the AI-based Fake News Detection system successfully achieves its primary goal of identifying and classifying misleading or false information. By leveraging natural language processing (NLP) techniques and machine learning algorithms, the system can effectively analyze textual content and detect patterns indicative of fake news. The preprocessing steps, including tokenization, stop-word removal, and feature extraction, significantly enhanced the quality of input data, leading to improved model accuracy. The system's evaluation demonstrates promising results in terms of precision, recall, and overall accuracy, indicating its reliability for practical applications. Its modular design ensures ease of maintenance and scalability, allowing the integration of more advanced models or additional datasets in the future. The project underscores the potential of AI in addressing critical challenges posed by misinformation in digital media, providing a robust tool to support users, journalists, and organizations in verifying information.

## REFERENCES

1. Chen, Y., Conroy, N. J., & Rubin, V. L. (2015, November). Misleading online content: Recognizing clickbait as false news. In Proceedings of the 2015 ACM Workshop on Multimodal Deception Detection (pp. 15-19). ACM.
2. Conroy, N. J., Rubin, V. L., & Chen, Y. (2022, November). Automatic deception detection: Methods for finding fake news. In Proceedings of the 78<sup>th</sup> ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community (p. 82). American Society for Information Science.
3. Della Vedova, M. L., Tacchini, E., Moret, S., Ballarin, G., Di Pierro, M., & de Alfaro, L. (2022, May). Automatic Online Fake News Detection Combining Content and Social Signals. In 2018 22nd Conference of Open Innovations Association (FRUCT) (pp. 272-279). IEEE.
4. Haiden, L., & Althuis, J. (2018). The Definitional Challenges of Fake News.
5. Helmstetter, S., & Paulheim, H. (2018, August). Weakly supervised learning for fake news detection on Twitter. In 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM) (pp. 274-277). IEEE.
6. Najafabadi, M. M., Villanustre, F., Khoshgoftar, T. M., Seliya, N., Wald, R., & Muharemagic, E. (2015). Deep learning applications and challenges in big data analytics. *Journal of Big Data*, 2(1), 1.
7. Parikh, S. B., & Atrey, P. K. (2018, April). Media-Rich Fake News Detection: A Survey. In 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR) (pp. 436-441). IEEE.
8. Shao, C., Ciampaglia, G. L., Varol, O., Flammini, A., & Menczer, F. (2017). The spread of fake news by social bots. arXiv preprint arXiv:1707.07592, 96-1