

An NLP-Based Approach to Fake News Identification Using TF-IDF and Logistic Regression

Vidyasri G.¹; Dr. K. Priya²

¹Department of Computer Applications (MCA), Marudhar Kesari Jain College for Women (Autonomous), Vaniyambadi, Tirupattur District – 635751

²PG & Research Department of Computer Science, Marudhar Kesari Jain College for Women (Autonomous), Vaniyambadi, Tirupattur District – 635751

Publication Date: 2026/10/22

Abstract: Online news services and social networks enable information to circulate rapidly, but they also make it easier for inaccurate stories to reach large audiences. Checking every item manually is difficult when the volume of published content is high. This paper examines a text-classification pipeline that labels news articles as FAKE or REAL. Article titles and body text are cleaned using Natural Language Processing (NLP), represented as Term Frequency–Inverse Document Frequency (TF-IDF) vectors, and passed to a Logistic Regression classifier. Performance is assessed with accuracy, precision, recall, F1-score, and a confusion matrix. On the selected test set, the model obtained 75.85% accuracy and a 77.95% F1-score. The results illustrate both the usefulness and the limitations of a lightweight, text-based approach: the classifier identified REAL articles more reliably than FAKE articles.

Keywords: Fake News Classification, NLP, TF-IDF, Logistic Regression, Text Mining, Machine Learning.

How to Cite: Vidyasri G.; Dr. K. Priya (2026) An NLP-Based Approach to Fake News Identification Using TF-IDF and Logistic Regression. *International Journal of Innovative Science and Research Technology*, 11(10), 1271-1274. <https://doi.org/10.38124/ijisrt/26oct569>

I. INTRODUCTION

Digital publishing has changed the speed and scale at which people encounter news. A story can be reposted across websites and social platforms within minutes, regardless of whether its claims have been verified. When misleading content is presented in a news-like format, readers may have difficulty distinguishing it from trustworthy reporting. This creates a need for practical methods that can help screen large collections of online articles.

Human fact-checking remains important, yet it requires time and specialist attention. Text analytics can provide an initial automated assessment by learning language patterns associated with labelled examples. In a typical text-classification workflow, unstructured documents are cleaned and encoded as numerical features before being processed by a learning algorithm.

This study evaluates a straightforward pipeline based on TF-IDF representations and Logistic Regression. The article title and main text are combined as the model input, then assigned to one of two classes: FAKE or REAL. The objective is to measure how well this approach performs on a held-out subset of the selected dataset, rather than to claim

that text classification alone can establish the factual truth of every news report.

➤ Objectives

- Prepare article titles and text for computational analysis using text-cleaning operations.
- Represent the cleaned documents with TF-IDF features.
- Train a Logistic Regression classifier to distinguish the FAKE and REAL labels.
- Measure predictive performance using accuracy, precision, recall, F1-score, and a confusion matrix.
- Examine class-wise errors and identify practical directions for further improvement.

II. LITERATURE REVIEW

Research on automated fake-news analysis spans feature engineering, conventional machine learning, and more recent language-modelling approaches. The survey by Alghamdi, Luo, and Lin [1] reviews machine-learning methods used in this area and provides a broader view of the classification strategies explored in prior work.

Document representations such as Bag-of-Words and TF-IDF are commonly used to convert text into vectors that can be processed by conventional classifiers. The conference paper cited in [2] discusses Natural Language Processing and machine-learning techniques for fake-news detection. Such representations are useful baselines because they are relatively straightforward to implement and can handle sparse, high-dimensional text data.

The study listed in [3] specifically considers Logistic Regression with TF-IDF for fake-news classification. Logistic Regression is a reasonable baseline for binary classification because it estimates class probabilities from the input features and is generally efficient to train. Nevertheless, lexical features may not capture context, satire, source credibility, or the truth of a claim. Building on this line of work, the present study implements the TF-IDF–Logistic Regression combination and reports its performance on the selected dataset.

III. METHODOLOGY

The experiment follows a sequence of data preparation, text normalization, dataset splitting, feature construction, model training, and evaluation. The same prepared text is represented numerically before it is supplied to the binary classifier.

➤ Dataset Preparation

A JSON-formatted subset of FakeNewsNet was processed in Python with Pandas in Google Colab. The dataset used in this experiment contains 2,048 records and nine fields: text, keywords, title, summary, date, label, user, spread, and tweets. There are 1,024 records in each class. Empty values in the title and text fields were replaced with empty strings, after which the two fields were joined to create the document supplied to the model. Records with no usable combined text were removed. Labels were encoded as FAKE = 0 and REAL = 1.

➤ Text Preprocessing

Text normalization was applied to reduce formatting variation and remove material that was not required for the chosen representation. The processing sequence included:

- Converting letters to lowercase so that case differences do not create separate features.
- Removing web addresses and HTML markup.
- Discarding non-alphabetic characters according to the implemented cleaning rule.
- Collapsing repeated whitespace to produce consistently formatted text.

➤ Training and Test Sets

The prepared records were divided into training and test subsets using an 80:20 split. The resulting sets contained 1,638 training records and 410 test records. Stratification maintained the class balance in the test subset, which included 205 FAKE and 205 REAL articles. The split used

random_state = 42 so that the experiment could be reproduced with the same data and software configuration.

➤ TF-IDF Feature Representation

Each document was represented using Term Frequency–Inverse Document Frequency. The term-frequency component reflects a term's presence within a particular document, while the inverse-document-frequency component reduces the influence of terms that occur throughout many documents. Their combination gives a weighted representation of the text.

$$\text{TF-IDF}(t,d) = \text{TF}(t,d) \times \text{IDF}(t)$$

English stop words were excluded and the vocabulary was capped at 10,000 features. To prevent information from the test subset influencing vocabulary construction, the vectorizer was fitted on the training documents only; the learned transformation was then applied to the test documents. The resulting matrices had dimensions of 1,638 × 10,000 for training and 410 × 10,000 for testing.

➤ Logistic Regression Model

The TF-IDF vectors were used as input to a Logistic Regression classifier. For each document, the classifier predicts one of the two encoded categories: FAKE (0) or REAL (1). The model was configured with max_iter = 1000 and random_state = 42. The trained classifier was then applied to the held-out test data.

➤ Evaluation Measures

The predictions were assessed using accuracy, precision, recall, and F1-score, with a confusion matrix used to inspect class-level errors. In the reported calculation, REAL (label 1) is treated as the positive class.

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN})$$

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$$

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$$

$$\text{F1-score} = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$$

Here, TP and TN represent correctly predicted positive and negative instances, while FP and FN represent the two types of incorrect prediction. Precision reflects the proportion of positive predictions that are correct; recall measures the proportion of actual positive instances recovered. The F1-score combines precision and recall into one measure.

➤ Processing Workflow

The workflow begins by loading and preparing the dataset. Titles and article text are cleaned and joined, and the data is split into training and testing subsets. TF-IDF is fitted on the training text and applied to both subsets. Logistic Regression learns from the training vectors, after which predictions for the test vectors are compared with the known labels using the selected evaluation measures.

IV. RESULTS AND DISCUSSION

➤ Overall Performance

Evaluation was carried out on the 410 held-out articles. The measured accuracy was 75.85%, meaning that

approximately three quarters of the test examples received the correct label. The summary measures are listed in Table 1.

Table 1 Test-Set Performance of the TF-IDF and Logistic Regression Model.

Metric	Result
Accuracy	75.85%
Precision	71.72%
Recall	85.37%
F1-score	77.95%

The reported recall of 85.37% applies to the REAL class, which was designated as positive. Precision was 71.72%, while the F1-score was 77.95%. Taken together,

these values indicate useful but imperfect classification performance and show that accuracy alone does not describe every aspect of the model's behaviour.

➤ Class-wise Results

Table 2 Class-Wise Classification Report.

Class	Precision	Recall	F1-score
FAKE	0.82	0.66	0.73
REAL	0.72	0.85	0.78
Overall accuracy	—	—	0.76

The class-level figures show that REAL articles were recovered more consistently than FAKE articles: recall was 0.85 for REAL and 0.66 for FAKE. Thus, a portion of fake articles were labelled as real by the model. This error is important in a fake-news screening application, because a system that misses misleading articles may provide an incomplete warning to users.

➤ Confusion Matrix

The confusion matrix compares the actual class of each test article with the category predicted by the classifier. It separates correct predictions from false alarms and missed cases, making it possible to inspect which class contributes most to the model's errors.

[Insert Fig. 1: Confusion matrix generated from the experiment.]

➤ Visualization of Evaluation Metrics

A bar chart of accuracy, precision, recall, and F1-score can be used to compare the four reported measures at a glance. The values to plot are 75.85%, 71.72%, 85.37%, and 77.95%, respectively. The visualization should be generated from the same experimental output as Table 1.

[Insert Fig. 2: Performance-metric chart generated from the reported results.]

The results indicate that the selected text representation and classifier can separate some FAKE and REAL articles, but the class-wise report also shows that the performance is not uniform across labels.

➤ Discussion

The experiment provides a baseline for classifying news from textual content. TF-IDF supplies a sparse numerical representation, and Logistic Regression offers a comparatively simple model for learning the distinction between the two labels. The reported results are promising as a baseline, but they should not be interpreted as proof that the system can independently verify news claims.

The recall difference between the two classes is a key limitation: the model's recall for FAKE articles is 0.66, compared with 0.85 for REAL articles. The approach also relies on a relatively small, balanced dataset and lexical features, which may not represent contextual meaning or changing patterns in misleading content. Future experiments could test larger and more diverse datasets, evaluate additional linguistic and source-related features, compare alternative classifiers, and validate performance on data collected from different time periods or platforms.

V. CONCLUSION

This work implemented a baseline fake-news classification pipeline using text preprocessing, TF-IDF feature extraction, and Logistic Regression. The experiment used 2,048 records from a JSON-formatted FakeNewsNet subset and evaluated the model on 410 held-out samples. The reported test results were 75.85% accuracy, 71.72% precision, 85.37% recall, and 77.95% F1-score. Class-wise results showed stronger recall for REAL articles (0.85) than for FAKE articles (0.66), indicating that some misleading articles were not detected. The model is computationally straightforward, but its performance and reliability require further assessment. Future work should investigate larger and more varied datasets, richer contextual features, comparisons

with other learning methods, and evaluation on independent data.

ACKNOWLEDGEMENT

The authors acknowledge Marudhar Kesari Jain College for Women (Autonomous), Vaniyambadi – 635751, Tamil Nadu, for providing access to the computational facilities in the FIST Instrumentation Lab under Grant No. SR/FST/COLLEGE/2023/1511 (FIST Program 2023, TPN-88688). Please verify the official laboratory name and grant details before submission.

AUTHOR CHECKLIST BEFORE JOURNAL RESUBMISSION

- Re-run the experiment and verify all metrics against the saved classification report and confusion matrix.
- Confirm that the subset contains 2,048 usable records after removing empty documents; the stated split counts should add up consistently.
- Insert the actual confusion matrix and metric chart; do not submit placeholder captions.
- Verify every reference against the original publication and ensure each cited work supports the related statement.
- Review the journal's similarity policy and inspect the revised report. Rewriting does not guarantee a particular similarity score.
- Ensure the acknowledgement's laboratory name, grant number, and funding details match the institution's official record.

REFERENCES

- [1.] J. Alghamdi, S. Luo, and Y. Lin, "A comprehensive survey on machine learning approaches for fake news detection," *Multimedia Tools and Applications*, vol. 83, pp. 51009–51067, 2024.
- [2.] "Fake News Detection Using Natural Language Processing and Machine Learning Techniques," 2023 Seventh International Conference on Image Information Processing (ICIIP), IEEE, 2023, doi: 10.1109/ICIIP61524.2023.10537640.
- [3.] R. Hudaini, M. Hendriansyah, and E. Gautama, "Fake News Classification Using Logistic Regression with TF-IDF Feature Extraction," *Perbanas International Seminar on Economics, Business, Management, Accounting and IT*, 2026.