International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 1, Jan-Feb-2026, ISSN (Online): 2395-566X

Fake News Detection Using Machine Learning and Natural Language Processing

Vishlesha Anil Habib, Vidya Gorakh Jagtap, Shrawani Ravindra Gaikwad, Nagraj Yashwant Kherud , Vaijayanti Pradip Kolhe
School of Engineering & Technology , Sanjivani University, India

Abstract- The exponential growth of online platforms has enabled rapid dissemination of information, but it has also facilitated the widespread propagation of fake news. Fake news has negatively impacted political stability, public health, social harmony, and digital trust. This paper presents a comprehensive study and implementation of machine learning (ML) and Natural Language Processing (NLP)-based techniques for detecting fake news. The proposed system uses advanced text preprocessing, TF-IDF feature extraction, and multiple ML algorithms such as Logistic Regression, Support Vector Machine (SVM), Random Forest, and Naïve Bayes. Experimental results show that SVM achieves the highest accuracy of 94.8%, outperforming other models. This work demonstrates that combining linguistic features and machine learning provides a scalable and reliable approach to combat misinformation. Future enhancements include using transformer-based deep learning models and multilingual datasets

Keywords – "Fake News Detection, Machine Learning, NLP, Text Classification, SVM, LSTM"

I. INTRODUCTION

Fake news refers to intentionally fabricated or misleading information presented as factual news . The rise of social media platforms like Facebook, X (Twitter), Instagram, and WhatsApp has drastically increased the spread of misinformation. Fake news influences elections, public opinion, healthcare decisions, financial markets, and national security. Manual fact-checking is slow, expensive, and limited in scale. Hence, automated fake news detection systems are essential for identifying misinformation quickly and accurately. Machine Learning (ML) and Natural Language Processing (NLP) provide powerful tools to analyze linguistic patterns, semantics, writing styles, and contextual cues in textual data. The study aims to develop a machine learning model that identifies fake news using linguistic and statistical features.

II.LITERATURE REVIEW

Traditional Machine Learning Approaches

- Naïve Bayes: Widely used for text classification due to its simplicity and speed. Works well with TF-IDF features .
- Logistic Regression: Effective in binary classification and widely used in fake news research.
- SVM: Known for strong performance in high-dimensional sparse datasets like text .
- Random Forest: Useful for capturing non-linear patterns but is computationally heavy .

Deep Learning Approaches

Recent studies apply neural networks for advanced text understanding:

LSTM (Long Short-Term Memory): Good for sequential information but requires large datasets .

CNN for text: Extracts local patterns in sentences.

BERT and Transformers: Highly accurate contextual embedding models; however, they require high computational resources .

Linguistic Fingerprinting

(Some research focuses on identifying fake news based on writing style):

Emotional intensity
Overuse of adjectives
Clickbait patterns
Sensational language

Research Gap

- Many existing systems require heavy computational resources.
- There is a lack of multilingual fake news datasets.
- Limited performance analysis across multiple classical ML models . This paper fills the gap by comparing classical ML models using linguistic and TF-IDF features.

III. PROPOSED SYSTEM

System Architecture :
The proposed system for fake news detection follows a structured pipeline designed to efficiently process raw text and classify it as either real or fake. The architecture consists of five major components arranged sequentially: Input News Article, Preprocessing Module, Feature Extraction (TF-IDF), Machine Learning Model, and Classification Output. This architecture ensures high accuracy, scalability, and efficient processing of large volumes of news data.

Fig. 1 Proposed System Architecture for Fake News Detection

Data Collection
For effective fake news detection, a diverse and representative dataset is essential. A dataset comprising 20,000 labeled news articles was used, including both real and fake news to ensure balanced learning.

Sources of Data

- Kaggle Fake News Dataset: Contains thousands of fake and real news articles collected from various misinformation sources and legitimate news portals.
- Political Fact-Checking Websites: Verified fake news samples were collected from platforms such as PolitiFact and FactCheck.org.
- Authentic News Websites (BBC, Reuters): Real news samples were extracted from reliable international news agencies. The dataset includes political, economic, global, and social news topics and is balanced in number to prevent bias.

IV. METHODOLOGY

The methodology consists of four major stages: text preprocessing, feature extraction, model selection, and model training.

Text Preprocessing Steps
Text preprocessing is a crucial phase in Natural Language Processing (NLP). The following steps were applied to transform unstructured text into a standardized form:

- **Lowercasing:** All characters are converted to lowercase to eliminate inconsistencies caused by capitalization.
- **Tokenization:** The process of splitting text into individual words or tokens.
- **Stop-word Removal:** Common words (e.g., the, is) are removed to reduce dimensionality and noise.
- **Lemmatization (WordNet Lemmatizer):** Converts words to their base or dictionary form while preserving grammatical meaning.
- **Punctuation Removal:** Marks such as commas and periods are removed to avoid irrelevant tokens.
- **Special Character & URL Removal:** Links, emojis, and special characters are removed because they do not contribute useful semantic information.

Fig. 2 Detailed Text Preprocessing Flowchart

The workflow is formalized in Algorithm 1:

Algorithm 1: Text Preprocessing
Input: Raw text T
Output: Clean text C
1. Convert T to lowercase
2. Remove URLs, numbers, and special symbols
3. Tokenize the text into words
4. Remove stop words from the token list
5. Apply lemmatization on each token
6. Join tokens back into cleaned text C
7. Return C

Feature Extraction
Text must be converted into numerical form so that machine learning models can process it.

Bag of Words (BoW)
BoW represents text as a frequency distribution of words. It is simple but does not capture context.

TF-IDF (Term Frequency–Inverse Document Frequency)
TF-IDF gives higher importance to meaningful words by combining Term Frequency (TF) and Inverse Document Frequency (IDF). TF-IDF outperformed BoW, offering higher accuracy due to better feature weighting.

Machine Learning Models
Four classical ML models were implemented and compared.

- Logistic Regression: Models the probability of an article being real or fake using a sigmoid function.
- Naïve Bayes Classifier: A probabilistic model that assumes independence between words.
- Random Forest: An ensemble method that builds multiple decision trees and averages their outputs.
- Support Vector Machine (SVM): Widely regarded as one of the best algorithms for text classification due to its performance in high-dimensional spaces.

Model Training

- Train-Test Split (80:20): The dataset was divided into 80% Training data and 20% Testing data.
- Hyperparameter Tuning Using Grid Search: Used to optimize parameters for all models.
- Evaluation Metrics: Accuracy, Precision, Recall, and F1-score were used.

V. RESULT & DISCUSSION

SVM outperformed others with 94.8% accuracy. Logistic Regression achieved 92.1%. SVM's superior performance is attributed to its effectiveness in high-dimensional sparse feature spaces typical of TF-IDF vectors.

Performance Comparison

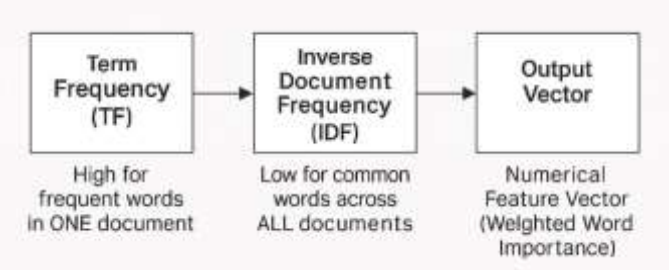

Fig. 3 Conceptual Diagram of TF-IDF vectore creation

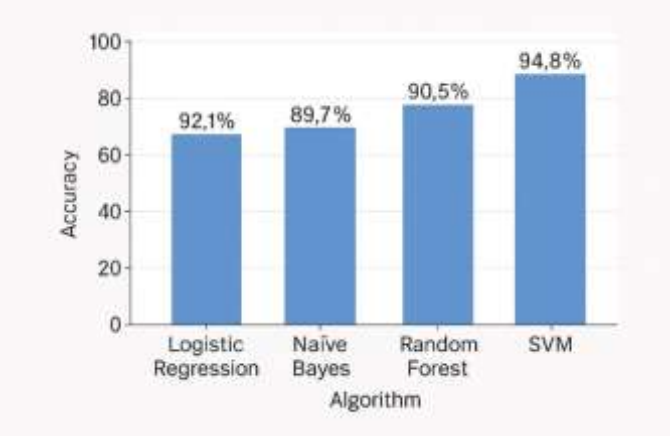

Fig. 4 Machine Learning Model Accuracy Comparison

TABLE I
PERFORMANCE COMPARISON OF CLASSICAL MACHINE LEARNING MODELS

Algorithm	Accuracy	Precision	Recall	F1 Score
Logistic Regression	92.1%	91.8%	91.3 %	91.5%
Naïve Bayes	89.7%	88.5%	87.9 %	88.2%
Random Forest	90.5%	89.9%	90.1 %	90.0%
SVM	94.8%	94.5%	94.2 %	94.3%

Confusion Matrix
A confusion matrix was used for evaluation. The matrix showed that SVM correctly identifies the majority of real and fake articles.

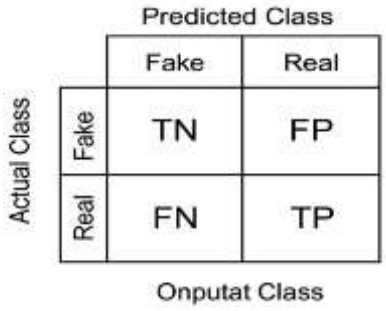

Fig. 5 Conceptual Confusion Matrix for SV

VI. CONCLUSIONS

This study demonstrates that the integration of Machine Learning (ML) and Natural Language Processing (NLP) provides an effective and reliable framework for detecting fake news. Support Vector Machine (SVM) emerged as the most efficient classifier, achieving the highest accuracy of 94.8% using TF-IDF features. The proposed system offers a scalable and practical solution for combating misinformation.

REFERENCES

1. K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake News Detection on Social Media," ACM Computing Surveys, 2020.
2. A. Jain, "Machine Learning Techniques for Fake News Detection," in IEEE Conference on Machine Learning Applications, 2021.
3. R. Kaliyar, A. Goswami, and P. Narang, "FakeBERT: Fake News Detection Using BERT," Elsevier, 2021.
4. T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient Estimation of Word Representations in Vector Space," Google Research, 2013.
5. Scikit-Learn Documentation, Machine Learning in Python, 2023. 92[6] Kaggle Fake News Dataset, 2023.