© MAY 2026 | IRE Journals | Volume 9 Issue 11 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I11-1717820

Fake News Detection using NLP and Deep Learning

PRADIP SUBHASH AHIRE[1], NIRAJ MAHENDRA LOKHANDE[2], PROF GADEKAR M. J.[3]
[1, 2, 3] *A.M. college, Hadapsar, Pune Savitribai Phule Pune University*

Abstract*—The continuous growth of digital communication platforms has made information sharing easier and faster than ever before. At the same time, the spread of false and misleading news through social media, blogs, and online news portals has become a major global concern. Fake news can influence public opinion, create social panic, and negatively affect political, economic, and healthcare systems. To address this issue, researchers have focused on developing intelligent systems capable of automatically identifying fake information. This research paper presents a detailed study of fake news detection using Natural Language Processing (NLP) and Deep Learning techniques. The proposed approach applies multiple NLP preprocessing methods such as tokenization, stop-word elimination, stemming, lemmatization, and text vectorization to prepare news articles for analysis. Different deep learning models including Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), Bidirectional LSTM (Bi-LSTM), and transformer-based BERT models are used for classification. The study also compares traditional machine learning algorithms with modern deep learning techniques. Experimental observations show that transformer-based architectures provide better contextual understanding and higher prediction accuracy. The proposed system aims to support automated misinformation detection and improve the reliability of online content.

Keywords*—Fake News Detection, NLP, Deep Learning, CNN, LSTM, BERT, Machine Learning, Artificial Intelligence.

I. INTRODUCTION

Digital technology has completely transformed the way people consume and share information. Platforms such as Facebook, Twitter, Instagram, YouTube, and online news websites allow millions of users to access real-time information from anywhere in the world. Although these technologies provide significant advantages, they have also increased the circulation of false information and manipulated news content.

Fake news refers to intentionally misleading or fabricated information presented in the format of genuine news. Such content is often designed to manipulate public opinion, gain financial profit, influence elections, or create social instability. During major global events such as elections, pandemics, and international conflicts, the spread of misinformation becomes extremely rapid.

Manual fact-checking methods are effective but difficult to scale due to the huge volume of online content generated every second. As a result, automated fake news detection systems based on Artificial Intelligence (AI) have become increasingly important. Natural Language Processing enables computers to analyse human language, while Deep Learning models help identify hidden semantic patterns within textual data.

This paper explores different NLP and deep learning approaches used for fake news detection and evaluates their effectiveness in identifying misleading information.

II. PROBLEM STATEMENT

The increasing amount of misinformation shared through social media and online news platforms has created serious challenges in identifying reliable information sources. Traditional manual verification techniques are time-consuming and cannot efficiently handle large-scale digital data.

Major issues related to fake news detection include:

- Rapid spread of misleading information
- Difficulty in identifying trustworthy sources
- Human bias during manual verification
- Lack of scalable automated systems
- Challenges in multilingual content analysis
- AI-generated synthetic news articles

Therefore, there is a strong need for an intelligent and automated fake news detection framework using NLP and Deep Learning techniques.

III. OBJECTIVES

The main objectives of this research work are:

1. To study various fake news detection methods.
2. To preprocess textual data using NLP techniques.

© MAY 2026 | IRE Journals | Volume 9 Issue 11 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I11-1717820

3. To extract meaningful linguistic features from news content.
4. To implement deep learning models for classification.
5. To compare the performance of multiple algorithms.
6. To improve detection accuracy.
7. To reduce the spread of misinformation.
8. To support real-time news verification.

IV. LITERATURE REVIEW

Researchers across the world have proposed several machine learning and deep learning approaches for identifying fake news.

Shu et al. introduced the FakeNewsNet dataset, which combines news content with social media engagement information. Their research highlighted the importance of user interaction analysis in fake news detection.

Wang proposed the LIAR dataset, which contains political statements labelled according to truthfulness categories. This dataset became one of the most widely used benchmark datasets in fake news research.

Traditional machine learning approaches such as Naive Bayes, Support Vector Machine (SVM), and Logistic Regression showed moderate success in text classification tasks. However, these methods often struggle to understand contextual meaning.

Deep learning techniques such as CNN and LSTM improved fake news detection by automatically learning semantic relationships from text. More recently, transformer-based architectures such as BERT achieved superior performance due to their contextual language understanding capabilities.

Recent studies also focus on multimodal fake news detection using text, images, videos, and social network behaviour. Despite these advancements, challenges such as explainability, multilingual detection, and adversarial manipulation still remain.

V. PROPOSED METHODOLOGY

The proposed fake news detection system consists of multiple stages including data collection, preprocessing, feature extraction, model training, and classification.

5.1 Data Collection

The following datasets are used for experimentation:

Dataset Number of Samples Type

Kaggle Fake News Dataset 20,800 News Articles
LIAR Dataset 12,800 Political Statements
FakeNewsNet 23,000+ News + Social Data

5.2 Data Preprocessing

Before training the models, textual news data is cleaned and transformed using NLP techniques.

NLP Preprocessing Steps

1. Tokenization
2. Stop-word Removal
3. Stemming
4. Lemmatization
5. Vectorization

Example

Original Sentence:

“The fake news spreads rapidly on social media platforms.”

After Preprocessing:

“Fake news spread rapidly social media platform”

VI. SYSTEM ARCHITECTURE

The proposed fake news detection framework follows the sequence below:

Input News Article
↓
Text Cleaning
↓
Tokenization
↓
Stop-word Removal
↓
Feature Extraction
↓
Word Embedding
↓
Deep Learning Model
↓
Classification
↓
Fake / Real News

VII. NLP TECHNIQUES USED

7.1 Tokenization

Tokenization divides textual data into smaller units known as tokens.

Example:

Sentence: “Fake news spreads rapidly”

Tokens: [“Fake”, “news”, “spreads”, “rapidly”]

7.2 Stop-word Removal

Frequently used words such as “the”, “is”, and “and” are removed to improve processing efficiency.

7.3 Stemming
Stemming converts words into their root forms.
Examples:
- playing → play
- running → run

7.4 Lemmatization
Lemmatization transforms words into meaningful base forms.
Example:
- better → good

7.5 Vectorization
Textual data is converted into numerical form using methods such as:
- TF-IDF
- Word2Vec
- GloVe
- FastText
- BERT Embeddings

VIII. DEEP LEARNING MODELS

8.1 Convolutional Neural Network (CNN)
CNN models identify important local textual features and keywords.
Advantages
- Faster training process
- Efficient feature extraction
- Good classification accuracy

8.2 Long Short-Term Memory (LSTM)
LSTM networks are designed to capture long-term dependencies in sequential textual data.
Advantages
- Handles sequence information effectively
- Suitable for language-based tasks

8.3 Bidirectional LSTM (Bi-LSTM)
Bi-LSTM processes information from both forward and backward directions.

Advantages
- Better contextual understanding
- Improved prediction capability

8.4 BERT Model
BERT uses transformer architecture and contextual embeddings for language understanding.

Advantages
- High prediction accuracy
- Strong semantic understanding
- Better handling of contextual ambiguity

IX. MATHEMATICAL MODEL

For binary classification:
$P(y=1|x)=\sigma(Wx+b)$
Where:
- x = Input feature vector
- W = Weight matrix
- b = Bias value
- σ = Sigmoid activation function
- y = Predicted output

Binary Cross Entropy Loss Function
$Loss = -[y \log(p) + (1-y) \log(1-p)]$

X. ALGORITHM

Fake News Detection Algorithm

Step 1: Data Collection
- Download datasets
- Store data in CSV format

Step 2: Data Cleaning
- Remove punctuation marks
- Convert text into lowercase
- Remove unnecessary symbols

Step 3: NLP Processing
- Tokenization
- Stop-word removal
- Lemmatization

Step 4: Feature Extraction
- Apply TF-IDF or word embeddings

Step 5: Model Training
- Split dataset into training and testing sets
- Train CNN, LSTM, or BERT models

Step 6: Prediction
- Input a news article
- Predict whether the news is fake or real

XI. EXPERIMENTAL SETUP

Component	Specification

© MAY 2026 | IRE Journals | Volume 9 Issue 11 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I11-1717820

Programming Language	Python
NLP Libraries	NLTK, SpaCy
Deep Learning Framework	TensorFlow / Keras
Dataset	Kaggle Fake News Dataset
IDE	Jupyter Notebook
Hardware	Intel i5, 8GB RAM

XII. EVALUATION METRICS

The proposed model performance is evaluated using multiple metrics.

12.1 Accuracy
Accuracy = Correct Predictions / Total Predictions

12.2 Precision
Precision measures correctly predicted fake news instances.

12.3 Recall
Recall measures the model's ability to detect actual fake news.

12.4 F1-Score
F1-score represents the harmonic mean of precision and recall.

12.5 Confusion Matrix
A confusion matrix is used to analyse classification performance.

XIII. RESULTS AND STATISTICAL ANALYSIS

Model	Accuracy
Naive Bayes	84%
Logistic Regression	87%
SVM	89%
CNN	93%
LSTM	95%
Bi-LSTM	96%
BERT	98%

Observations
- Deep learning models perform better than traditional machine learning algorithms.
- LSTM and Bi-LSTM effectively capture sequential information.
- BERT achieves the highest accuracy due to contextual embeddings.
- Transformer models provide better semantic understanding of language.

XIV. ADVANTAGES OF PROPOSED SYSTEM

- Automated fake news verification
- High classification accuracy
- Real-time prediction support
- Scalable architecture
- Reduction in misinformation spread
- Efficient handling of large datasets

XV. APPLICATIONS

Social Media Monitoring
Helps detect misleading posts and fake information.
News Verification Platforms
Supports journalists and fact-checking organizations.
Healthcare
Identifies medical misinformation.
Politics
Detects fake political campaigns and propaganda.
Cybersecurity
Reduces misinformation-based cyber threats.
Education
Encourages media literacy and information awareness.

XVI. CHALLENGES

The proposed system still faces several limitations:
- Multilingual fake news detection
- Understanding sarcasm and satire
- Dataset imbalance
- High computational requirements
- Adversarial attacks
- Lack of explainable predictions

XVII. FUTURE SCOPE

Future improvements may include:
1. Multimodal fake news detection using text, images, and videos.
2. Integration of Explainable AI techniques.
3. Browser extensions for real-time verification.
4. Blockchain-based content verification systems.
5. Graph Neural Networks for propagation analysis.
6. Cross-language fake news detection.

© MAY 2026 | IRE Journals | Volume 9 Issue 11 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I11-1717820

XVIII. CONCLUSION

Fake news detection using NLP and Deep Learning has become an essential research area for reducing misinformation in the digital world. NLP techniques help preprocess and analyse textual information, while deep learning models such as CNN, LSTM, Bi-LSTM, and BERT provide highly accurate classification results.

Experimental analysis demonstrates that transformer-based architectures outperform traditional machine learning methods because of their contextual language understanding capabilities. The proposed framework can assist governments, educational institutions, media organizations, and social networking platforms in improving the reliability of digital information.

REFERENCES

[1] Shu K., Sliva A., Wang S., Tang J. & Liu H. Fake News Detection on Social Media: A Data Mining Perspective.
[2] Wang W. Y. “Liar, Liar Pants on Fire”: A Benchmark Dataset for Fake News Detection.
[3] Kaliyar R. K., Goswami A. & Narang P., FakeBERT: Fake News Detection in Social Media with a BERT-Based Approach.
[4] Devlin J., Chang M. W., Lee K. & Toutanova K, BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.
[5] Zhou X. & Zafarani R. A Survey of Fake News: Theories, Detection Methods, and Opportunities.
[6] TensorFlow Documentation.
[7] Kaggle Fake News Dataset.
[8] NLTK Documentation.
[9] SpaCy Documentation.
[10] Vaswani A., et al. Attention Is All You Need.