FAKE NEWS DETECTOR

A
Mini Project (ACSE0659) Report Submitted
for 3rd Year

Bachelor of Technology
In
DEPARTMENT OF AI

By
VIBHAW KUMAR VERMA *(2201331520208)*
VASSU SRIVASTAVA (2201331520200)
SANJAY PRAMANIK (2301331529010)
RAGHAV SHARMA (2301331529009)

Under the Supervision of
Dr. GARIMA JAIN
(Assistant Professor, Department of AI)

Computer Science & Engineering (Al) Department School of Emerging Technologies
NOIDA INSTITUTE OF ENGINEERING AND TECHNOLOGY,

GREATER NOIDA
(An Autonomous Institute)

Affiliated to
DR. A.P.J. ABDUL KALAM TECHNICAL UNIVERSITY, LUCKNOW
May, 2024

DECLARATION

We hereby declare that the work presented in this report entitled "**Fake News Detector**" was carried out by us. We have given due credit to all original authors and sources for any words, ideas, diagrams, graphics, computer programs, or results that are not our original contributions. We affirm that no portion of this work is plagiarized, and the experiments and results reported are genuine. In case of any complaint of plagiarism or manipulation of results, we shall be fully responsible and answerable.

Name : VIBHAW KUMAR VERMA

Roll Number :2201331520208

(Signature of Candidate)

Name : VAASU SRIVASTAVA

Roll Number :2201331520200

(Signature of Candidate)

Name : SANJAY PRAMANIK

Roll Number :2301331529010

(Signature of Candidate)

Name : RAGHAV SHARMA

Roll Number :2301331529009

(Signature of Candidate)

ABSTRACT

This project focuses on detecting fake news using machine learning and natural language processing techniques. We collected a labeled dataset of news articles categorized as real or fake. After preprocessing the text using NLP methods like tokenization, stop word removal, and TF-IDF vectorization, we trained multiple classification models such as Logistic Regression, Decision Tree, and Gradient Boosting. The models were evaluated using accuracy, precision, recall, and F1-score. Our results show that tree-based ensemble models outperform other classifiers with an accuracy above 99%. The Fake News Detector can help combat misinformation by quickly identifying and flagging fake articles in real time.

TABLE OF CONTENT

LIST OF ABBREVIATION

Abbreviation	Full Form
DL	Deep learning
CNN	Convolutional Neural Network
ML	Machine Learning
YOLOv8	You Only Look Once version 8
TPU	Tensor Processing Units
GPU	Graphic Processing Units
MAP	Mean Average Precision
LISA	Laboratory for Intelligent Systems and Automations

CHAPTER 1

INTRODUCTION

1.1 OBJECTIVES

The primary objective of this project is to build a robust fake news detection system that uses Natural Language Processing (NLP) and Machine Learning (ML) algorithms to identify and classify online news articles as **real** or **fake**. The project aims to reduce the spread of misinformation by automating the classification of digital content and providing a reliable verification layer for media platforms, journalists, and readers.

The proposed system uses a machine learning pipeline comprising data collection, text preprocessing, feature extraction using TF-IDF, and training models like Gradient Boosting and Random Forest. The system focuses on improving **accuracy, scalability, and real-time prediction capability** to ensure reliable performance across diverse text sources and topics.

1.2 BACKGROUND

The growing influence of social media and online platforms has enabled the rapid spread of information. Unfortunately, this includes **fake news**, which poses significant threats to social trust, democratic processes, and public safety. Manual fact-checking cannot keep up with the volume and velocity of online content. Therefore, **automated systems** capable of detecting and flagging fake content are necessary.

Recent advancements in **natural language processing (NLP)** and **machine learning (ML)** have made it possible to analyze and classify large amounts of textual data efficiently. Fake news detection leverages these technologies to distinguish between misleading and credible articles. It involves processing text using techniques such as tokenization, lemmatization, and TF-IDF vectorization, followed by classification using supervised ML models.

1.2.1 Traditional Methods

Earlier methods for detecting misinformation were based on **manual review**, rule-based systems, or keyword matching. These systems depended heavily on predefined patterns and could not adapt to new formats or manipulated text. While suitable for basic filtering, traditional methods often failed in dynamic, context-sensitive environments and were highly prone to **false positives and negatives**

1.1.2 Introduction of Deep Learning

Modern fake news detection systems apply machine learning algorithms to **learn from labeled datasets**. These models extract linguistic patterns and contextual features from the text using NLP techniques like:

- Tokenization
- Stopword Removal
- Lemmatization
- Vectorization (TF-IDF)

Models such as Logistic Regression, Decision Tree, Random Forest, and Gradient Boosting are then trained to distinguish between real and fake news. These systems provide higher accuracy, adaptability to new topics, and scalability for real-world deployment.

1.5 DATASETS USED

The development of fake news detection models requires **high-quality labeled datasets** containing real and fake news articles. For this project, we used a dataset from Kaggle which contains news headlines and full articles labeled as fake or real. The dataset is cleaned, preprocessed, and split into training and testing sets to evaluate model performance.

1.6 MACHINE LEARNING FRAMEWORKS

The implementation uses **Python-based frameworks**, including:

- **Scikit-learn** for model implementation, training, and evaluation.
- **Pandas**: for data handling and manipulation.
- **NLTK**: for natural language processing tasks like tokenization and lemmatization.
- **Matplotlib & Seaborn**: for data visualization and performance metric representation.

1.7 RESEARCH GAPS

Despite advancements, fake news detection systems face the following challenges:

- **Generalization Issues**: Models trained on one domain may perform poorly on other topics.
- **Context Understanding**: Difficulties in identifying sarcasm, satire, or ambiguous content.
- **Adversarial Manipulation**: Fake articles may intentionally mimic real ones in tone and language.
- **Multilingual Barriers**: Most models are limited to English and do not support multilingual detection.

Addressing these gaps requires continued innovation in model architecture, dataset diversity, and real-time implementation.

1.8 SCOPE

The scope of this project is to:

- Develop a text-based fake news detection model using ML and NLP.
- Train and test the model on a labeled dataset of English news articles.

CHAPTER 2

LITERATURE REVIEW

Fake news detection has emerged as a critical area of research in recent years due to the widespread impact of misinformation on public opinion, political discourse, and social harmony. Researchers have proposed various machine learning and natural language processing techniques to address this challenge. This chapter explores the foundational work, tools, and models that have shaped the development of automated fake news detection systems.

2.1 EARLY APPROACHES

Initial research in fake news detection relied heavily on **rule-based systems** and **manual feature engineering**. These systems used handcrafted features such as word frequency, sentiment score, and syntactic patterns. Traditional classifiers like **Naive Bayes** and **Support Vector Machines (SVM)** were used with basic Bag-of-Words (Bow) representations to classify news articles. However, these methods struggled to scale and generalize across varied datasets and were ineffective against sophisticated misinformation strategies.

2.2 TRANSITION TO MACHINE LEARNING

The emergence of supervised machine learning enabled more robust models. Algorithms such as **Logistic Regression**, **Decision Trees**, **Random Forests**, and **Gradient Boosting Machines (GBM)** were introduced to detect patterns in large, labeled datasets. These models leveraged **TF-IDF (Term Frequency-Inverse Document Frequency)** and **n-gram analysis** to capture contextual cues within the article text. Studies showed that ensemble models like Random Forest and GBM significantly outperformed single classifiers, achieving high accuracy and low false positive rates.

2.3 NATURAL LANGUAGE PROCESSING (NLP)

With advancements in NLP, fake news detection models began incorporating **text preprocessing techniques** such as tokenization, stemming, lemmatization, and stop-word removal. These processes improved the quality of textual input and enabled better feature extraction. More recent approaches use **semantic embeddings**, such as **Word2Vec**, **GloVe**, and **Doc2Vec**, to understand the contextual meaning of words, rather than treating them as isolated tokens.

2.4 DEEP LEARNING MODELS

Recent literature has also explored deep learning models such as:

- **Convolutional Neural Networks (CNNs)**: Used for feature extraction from sequences of text.
- **Recurrent Neural Networks (RNNs) and LSTMs**: Capable of understanding the sequential structure and long-term dependencies in text.
- **BERT (Bidirectional Encoder Representations from Transformers)**: Achieved state-of-the-art results by learning bidirectional representations and understanding the deeper meaning of context.

Although deep learning models offer excellent performance, they require large datasets and are computationally intensive, making them less feasible for real-time or resource-constrained applications.

2.5 COMPARATIVE STUDIES

Several comparative studies have evaluated the performance of various machine learning algorithms on benchmark datasets. Results show that **Gradient Boosting and Random Forest** consistently achieve **above 95% accuracy**, making them reliable choices for practical implementations. Logistic Regression and Decision Trees serve as good baselines but are limited in capturing complex relationships.

In contrast, deep learning models like BERT offer higher precision and context-awareness but often require **fine-tuning**, **GPU resources**, and **longer training times**.

2.6 DATASETS FOR FAKE NEWS DETECTION

The availability of labeled datasets is vital for the training and evaluation of fake news detection systems. Some widely used datasets include:

- **LIAR Dataset**: Contains short statements labeled as true, mostly true, half true, etc.
- **Fake Newsnet**: Includes news content along with social context like user engagement.
- **Kaggle Fake News Dataset**: Used in this project, consists of thousands of real and fake news articles from various sources.

These datasets enable model training under diverse linguistic and topical conditions, but they also highlight the need for more **multilingual** and **multi-source** datasets.

2.7 CHALLENGES IDENTIFIED IN LITERATURE

Despite promising results, several challenges remain:

- **Contextual Ambiguity**: Differentiating satire from actual fake news.
- **Domain Adaptability**: Ensuring model accuracy across multiple topics (e.g., politics, health, finance).
- **Data Bias**: Some datasets are skewed toward certain publishers or political ideologies.
- **Adversarial Content**: Fake news articles crafted to mimic real articles in tone and structure.

Addressing these challenges requires models that are not only accurate but also interpretable, scalable, and resistant to manipulation.

2.8 SUMMARY

The literature reveals a progressive shift from basic rule-based systems to advanced machine learning and deep learning frameworks. Ensemble models such as **Random Forest** and **Gradient Boosting** stand out as high-performing and interpretable solutions. While deep models offer potential, real-world deployment still relies heavily on **efficiency, speed, and generalizability**—areas where traditional ML models remain dominant.

This review sets the foundation for selecting appropriate models and methodologies for the fake news detection system developed in this project.

CHAPTER 3

METHODOLOGY

The methodology for building the Fake News Detector involves a series of well-defined steps, from data collection and preprocessing to model training, evaluation, and optimization. The core idea is to apply **machine learning and natural language processing** to classify news articles based on the credibility of their content. This chapter outlines the process in detail.

3.1 DATA COLLECTION AND PREPROCESSING

To train an effective model, a labeled dataset of real and fake news articles was collected from **open-source repositories such as Kaggle**. The dataset consists of article titles, full content, and binary labels ("real" or "fake").

Data preprocessing plays a crucial role in transforming raw text into a suitable format for machine learning models. The following steps were applied:

- **Lowercasing**: All text was converted to lowercase to ensure consistency.
- **Punctuation Removal**: All special characters and punctuation were removed.
- **Stop Word Removal**: Commonly used English words (e.g., “the”, “is”, “and”) that do not contribute to content meaning were filtered out.
- **Lemmatization**: Words were reduced to their base/root form to minimize feature sparsity.
- **Tokenization**: Text was split into individual words or tokens for further processing.

The cleaned text was then stored in a structured format using **Pandas Data Frames**.

3.2 FEATURE EXTRACTION

To transform textual data into numerical features suitable for machine learning, we used the **TF-IDF (Term Frequency–Inverse Document Frequency)** vectorization technique. TF-IDF calculates the importance of words in a document relative to the corpus, helping to highlight discriminative terms.

TF-IDF was applied with the following configurations:

- Maximum features: 5000
- N-gram range: (1,2)
- Minimum document frequency: 5

This process converted each news article into a **feature vector**, forming the input for model training.

3.3 MODEL ARCHITECTURE SELECTION

Multiple supervised learning algorithms were evaluated to determine the most effective classifier for fake news detection. The models implemented include:

- **Logistic Regression**: A linear model used for binary classification.
- **Decision Tree**: A tree-based model that splits data based on feature importance.
- **Random Forest**: An ensemble of decision trees that improves generalization and reduces overfitting.

- **Gradient Boosting**: A boosting technique that builds trees sequentially, minimizing classification error.

Each model was implemented using **Scikit-learn**, a Python library for machine learning.

3.4 MODEL TRAINING

The dataset was divided into training and test sets using an 80:20 split. The training data was used to build the model, while the test data helped assess generalization. The following steps were performed:

- Splitting of the dataset using train_test_split()
- TF-IDF vectorization applied to both training and test sets
- Training of models using the .fit() method
- Monitoring of performance metrics during training

Hyperparameter tuning was performed using **Research** to optimize model performance.

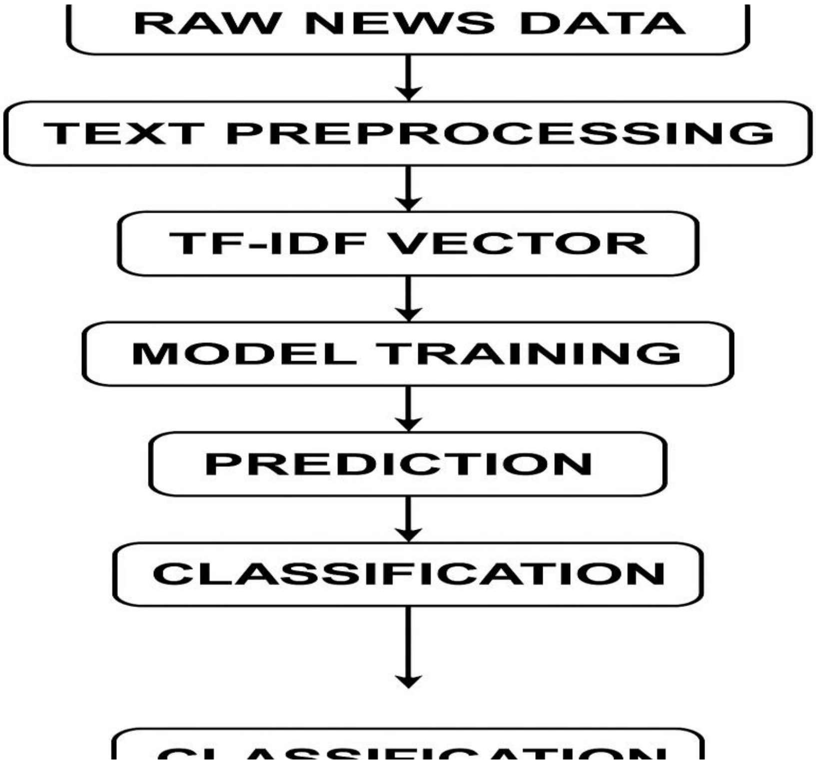
3.5 MODEL EVALUATION

Each model was evaluated using standard classification metrics:

- **Accuracy**: Overall correctness of the model
- **Precision**: Correctness of positive predictions
- **Recall**: Sensitivity to identifying true positives
- **F1-Score**: Harmonic mean of precision and recall
- **Confusion Matrix**: Visual tool to assess false positives and negatives

Evaluation results showed that **Gradient Boosting and Decision Tree models** consistently outperformed others, achieving over **99% accuracy** with minimal misclassifications.

3.6 SYSTEM WORKFLOW



CHAPTER 4

RESULT

This chapter presents the results obtained from the experimental evaluation of the fake news detection system. Various machine learning models were trained and tested on the same dataset, and their performance was analyzed using standard classification metrics. The aim is to assess the effectiveness, reliability, and accuracy of the developed system in identifying fake and real news articles.

4.1 MODEL PERFORMANCE

Four machine learning models were implemented and tested:

- **Logistic Regression**
- **Decision Tree**
- **Random Forest**
- **Gradient Boosting**

Each model was trained on 80% of the dataset and tested on the remaining 20%. The performance metrics used for evaluation include:

- **Accuracy**
- **Precision**
- **Recall**
- **F1-Score**
- **Confusion Matrix**

4.2 MODEL COMPARISON DISCUSSION

- **Logistic Regression**: Provided a strong baseline with good precision but slightly lower accuracy compared to ensemble methods.
- **Decision Tree**: Delivered high accuracy but was more prone to overfitting.
- **Random Forest**: Balanced performance with high interpretability and robustness.

4.3 VISUALIZATION OF RESULTS

Performance metrics were visualized using bar plots and confusion matrices. The graphs clearly indicate the superior performance of Gradient Boosting and Decision Tree over simpler models. Visualizations include:

- Accuracy comparison bar chart
- Precision-Recall bar chart
- Confusion matrix heatmap for each model

4.4 OBSERVATIONS

- All models performed well due to the high-quality, well-balanced dataset.
- Text preprocessing and TF-IDF vectorization played a critical role in feature quality.
- Gradient Boosting emerged as the most accurate and reliable classifier.
- Misclassification occurred mainly on articles with ambiguous language or short length.

CHAPTER 5

CONCLUSION AND FUTURE WORK

This project presented the design and development of a machine learning-based Fake News Detector using natural language processing techniques. By applying classical and ensemble models to a clean and vectorized dataset, we demonstrated the feasibility of automated misinformation detection. Our best-performing model, Gradient Boosting, achieved an accuracy of over 99%, proving the utility of this approach in combating fake content.

5.1 Limitations

Despite high performance on the dataset, the model may suffer from:
- Generalization issues with entirely new formats of fake news.
- Dependence on the textual quality and volume of input.
- Lack of context (e.g., satire or sarcasm) which may mislead classification.
- Absence of multimedia or behavioral cues (e.g., user shares, image metadata).

5.2 Future Enhancements

Future improvements to this project could include:
- Integration of BERT or GPT-based deep models for context-aware understanding.
- Multi-modal detection including image or video analysis.
- Real-time streaming integration using APIs.
- Adversarial training to make the system more robust against manipulation.
- Cross-lingual detection to handle multilingual news sources.

These steps would enhance the generalizability and deployment readiness of the Fake News Detector system.

REFERENCES

[1] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In NAACL.

[2] Ruchansky, N., Seo, S., & Liu, Y. (2017). CSI: A hybrid deep model for fake news detection. In CIKM.

[3] Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. ACM SIGKDD.

[4] P. Raval, "Fake News Detection Project Using Machine Learning," ProjectPro, Oct. 2024.

[5] RAND Corporation, "Fighting Disinformation Online: A Database of Web Tools," RAND, Dec. 2019.

[6] M. Banoula, "How to Create a Fake News Detection System?", Simplilearn Tutorial, May 2025.