International Journal of Scientific Research and Engineering Development-– Volume X Issue X, Year

Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Fake News Detection Using Machine Learning and Deep Learning Approaches with Flask-Based Web Deployment

Rajbhar Akash Ramavatar Meera Devi, Prof Maya Nair.
Department of Computer Science SIES, Mumbai
akashrajbhar.cs@gmail.com
Department of Computer Science, University of Mumbai, SIES College of Arts, Science and Commerce Sion(west).
mayan@sies.edu.in

----------**********----------

Abstract:

Fake news has emerged as one of the most serious challenges in the digital world, especially with the rapid growth of social media platforms and online news sources. The spread of misinformation not only misleads individuals but also affects social, political, and economic decision-making. Manual verification of news is inefficient and time-consuming, which has increased the demand for automated fake news detection systems. This paper presents a hybrid approach for fake news detection using Machine Learning (ML) and Deep Learning (DL) techniques. The system is designed to classify news articles as real or fake based on textual content. The ML model uses TF-IDF vectorization combined with classical machine learning classifiers, while the DL model uses a Long Short-Term Memory (LSTM) neural network for sequential text analysis. A Flask-based web application is developed to provide a user-friendly interface where users can input news text and choose between ML or DL models for prediction. The system also displays prediction confidence, maintains prediction history, and provides graphical comparison between models using Chart.js. Experimental analysis shows that the deep learning model outperforms traditional machine learning models in terms of accuracy and contextual understanding of text data.

Keywords **—** Fake News Detection, Machine Learning, Deep Learning, Natural Language Processing, Flask, LSTM, TF-IDF, Text Classification, Web Deployment.

----------**********----------

I. INTRODUCTION

In recent years, the internet has become the primary source of information for millions of users worldwide. However, the ease of publishing content online has also led to the rapid spread of fake news. Fake news refers to false or misleading information presented as legitimate news. It can influence public opinion, spread misinformation, and even cause social unrest. Traditional manual fact-checking methods are not sufficient to handle the large volume of data generated daily. Therefore, automated systems using Artificial Intelligence and Machine Learning are required. This project focuses on developing a fake news detection system using both Machine Learning and Deep Learning techniques. The system provides a web-based interface to make it accessible and user-friendly.

II. PROBLEM STATEMENT

With the rapid growth of digital media and online news platforms, the spread of fake news has become a significant global issue. Fake news refers to misleading or false information presented as legitimate news, which can influence public opinion, create confusion, and

 ©IJSRED: All Rights are Reserved

even lead to social, political, and economic consequences. Traditional methods of verifying news, such as manual fact-checking, are time-consuming, inefficient, and unable to handle the massive volume of data generated daily on the internet. As a result, there is a critical need for an automated system that can accurately and efficiently detect fake news in real time. The challenge lies in analysing unstructured textual data, understanding linguistic patterns, and distinguishing between genuine and misleading content. Furthermore, different approaches such as Machine Learning and Deep Learning offer varying levels of performance, and there is a need to evaluate and compare their effectiveness. Therefore, this project aims to develop a robust fake news detection system using both Machine Learning and Deep Learning techniques, integrated into a web-based application that provides real-time predictions, confidence scores, and performance comparison between models.

III. OBJECTIVES

The primary objectives of this research project are:

To design and develop an intelligent system capable of automatically detecting fake and real news using textual content.

1. To implement Machine Learning models using TF-IDF vectorization for efficient text classification.
2. To develop a Deep Learning model based on LSTM architecture to capture contextual and sequential patterns in news data.
3. To compare the performance of Machine Learning and Deep Learning approaches in terms of accuracy and reliability.
4. To preprocess and clean raw textual data using Natural Language Processing techniques such as stop word removal, punctuation removal, and text normalization.
5. To provide prediction results along with confidence scores for better interpretation of model outputs.
6. To build a Flask-based web application that allows users to input news text and get real time predictions.
7. To implement a model selection feature that enables switching between ML and DL models.
8. To maintain prediction history for user reference and analysis.
9. To visualize model performance using graphical representations such as accuracy comparison charts.

IV. LITERATURE REVIEW

Fake news detection has been widely studied using Machine Learning (ML), Deep Learning (DL), and Natural Language Processing (NLP) techniques. This section reviews important research contributions in this domain.

1. **Machine Learning-Based Fake News Detection (2024)** A study proposed various ML algorithms such as Logistic Regression, Decision Trees, and Random Forest for fake news classification. The results showed that ensemble methods like Random Forest performed better due to their ability to handle complex feature relationships.
2. **Word Embedding-Based Hybrid Approach (2024)** Research compared TF- IDF, Word2Vec, and Fast Text with ML and DL models. It concluded that combining word embeddings with deep learning improves classification accuracy significantly.
3. **Bi-LSTM with Linguistic Features (2023)** This paper introduced a Bi- directional LSTM model using linguistic features for fake news detection. The study showed that Bi-LSTM

 ©IJSRED: All Rights are Reserved

captures contextual dependencies better than traditional ML models.

4. **RNN vs LSTM vs Bi-LSTM Comparison (2024)** A comparative study evaluated RNN, LSTM, and Bi-LSTM architectures. Results indicated that Bi- LSTM outperformed others due to its ability to process text in both forward and backward directions.
5. **CNN and Large Language Model Comparison (2025)** This research compared CNN, NLP models, and large language models (LLMs). It highlighted that advanced DL models significantly outperform traditional approaches in detecting misinformation.

V. Impact on Society

The proliferation of fake news has become a critical issue in modern society, affecting politics, health, finance, and public safety. This project contributes positively in the following ways:

1.Enhancing Public Awareness:
By detecting fake news automatically, users can be alerted to misinformation before they share it, promoting more responsible consumption of information.

2.Reducing Social Harm:
Fake news can cause panic, spread health- related misinformation, or create conflicts. A detection system minimizes these negative consequences by flagging misleading content.

3.Supporting Democratic Processes:
Political misinformation can influence elections and public opinion. By identifying fake news, the project strengthens informed decision-making and democratic participation.

4.Promoting Digital Literacy:
Encourages users to critically evaluate the credibility of news sources. This project can serve as an educational tool in schools and communities.

5.Strengthening Online Ecosystems:
Platforms can integrate such detection systems to maintain trust and reduce the spread of harmful content, creating a safer and more reliable online environment.

6.Healthcare and Safety Benefits:
In scenarios like pandemics, fake news about cures or prevention methods can endanger lives. Detecting misinformation early helps protect public health.

VI. METHODOLOGY

A. Data Collection
Dataset contains:
- Title
- Text

Label (0 = Fake, 1 = Real)
A realistic synthetic dataset with noise is used.

B. Data Preprocessing
The dataset is cleaned by performing the following steps:
- Conversion to lowercase
- Removal of punctuation
- Removal of stop words
- Merging title and text fields

C. Feature Extraction (ML Model)
The TF-IDF (Term Frequency-Inverse Document Frequency) technique is used to convert text into numerical vectors. This helps machine learning models understand textual data.

D. Machine Learning Models
Multiple classifiers are trained and compared:
- Logistic Regression
- Naïve Bayes
- Support Vector Machine
- Random Forest
- Decision Tree
- Gradient Boosting
- K-Nearest Neighbors
- AdaBoost

The best performing model is selected based on accuracy.

E. Deep Learning Model

 ©IJSRED: All Rights are Reserved

The deep learning model uses an LSTM neural network. The architecture includes:

- Embedding Layer
- LSTM Layers
- Dropout Layers
- Dense Layer with Sigmoid Activation

The model is trained using binary cross-entropy loss function.

F. Web Application (Flask)

Flask is used to deploy the trained models. It allows users to:

- Input news text
- Select ML or DL model
- Get prediction results
- View confidence score
- See prediction history

View model comparison graph

VII. SYSTEM ARCHITECTURE

User Input (News Text)
↓
Text Preprocessing
↓
Model Selection (ML / DL)
↓
TF-IDF Model OR LSTM Model
↓
Prediction (Real / Fake)
↓
Confidence Score Calculation
↓
Flask Web Interface
↓
Result Display + History + Visualization

VIII. IMPLEMENTATION DETAILS

A. Tools & Technologies

- Python
- Flask
- Scikit-learn
- TensorFlow / Keras
- HTML, CSS, JavaScript
- Chart.js

B. ML Pipeline

- TF-IDF Vectorizer
- Multiple classifiers trained
- Best model selected
- Model saved using pickle

C. DL Pipeline

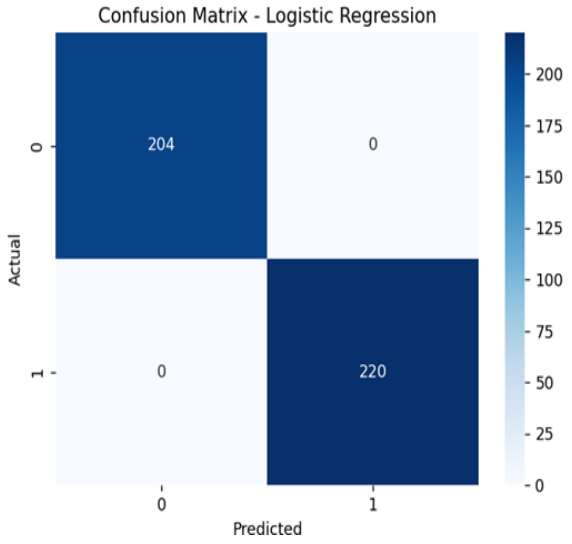
- Tokenization (5000 words vocabulary)
- Padding (max length 300)
- LSTM-based architecture
- Early stopping used
- Model saved as .h5 file

D. Flask Backend

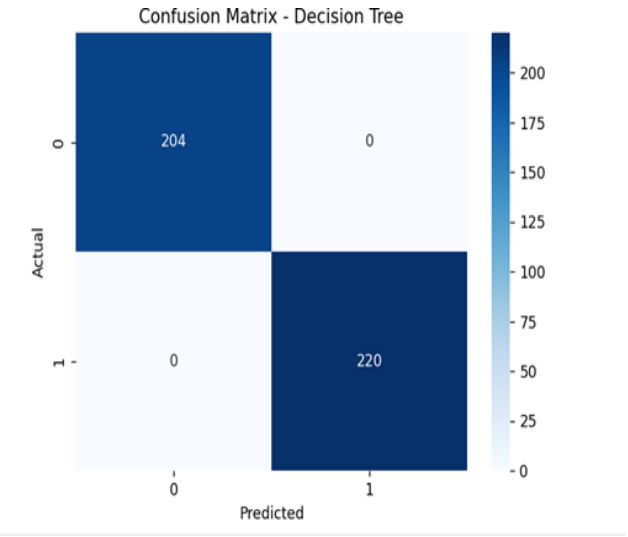
- / Route for UI
- /Predict route for prediction
- Model switching system (ML/DL)
- History storage in memory
- Confidence calculation

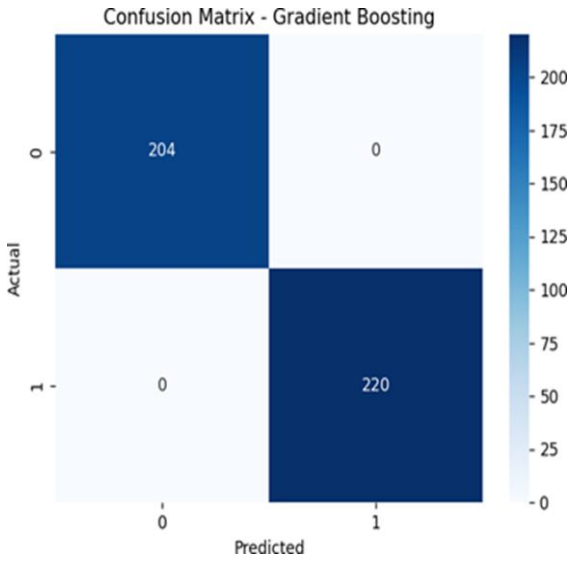
IX. RESULT AND DISCUSION

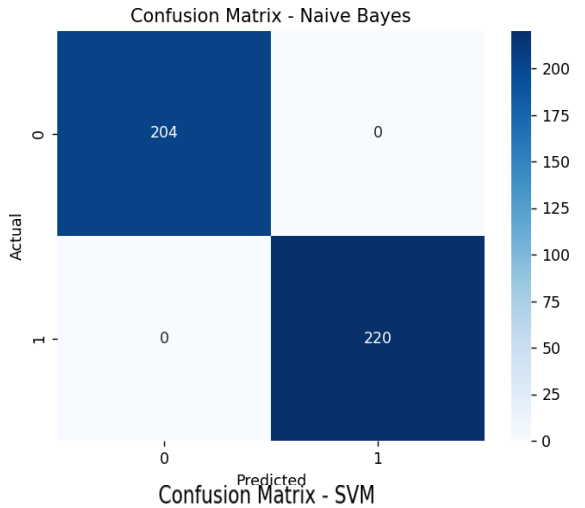
The system successfully classifies news articles into real and fake categories. The deep learning model performs better than traditional machine learning models due to its ability to capture semantic relationships between words. The system also provides prediction confidence and visual comparison of model performance using graphs. The web interface makes it easy for users to interact with the system.

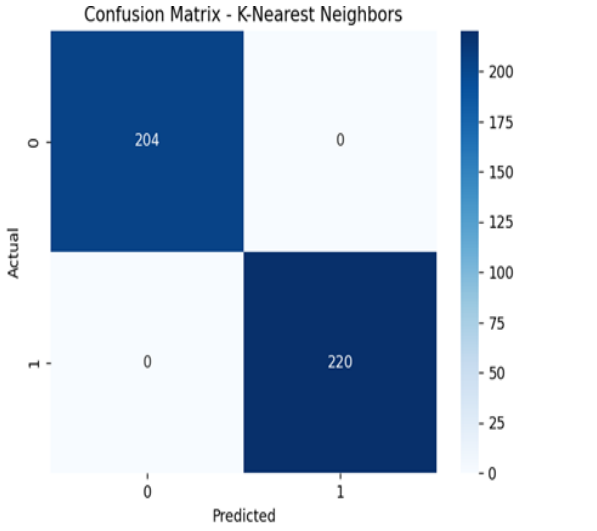


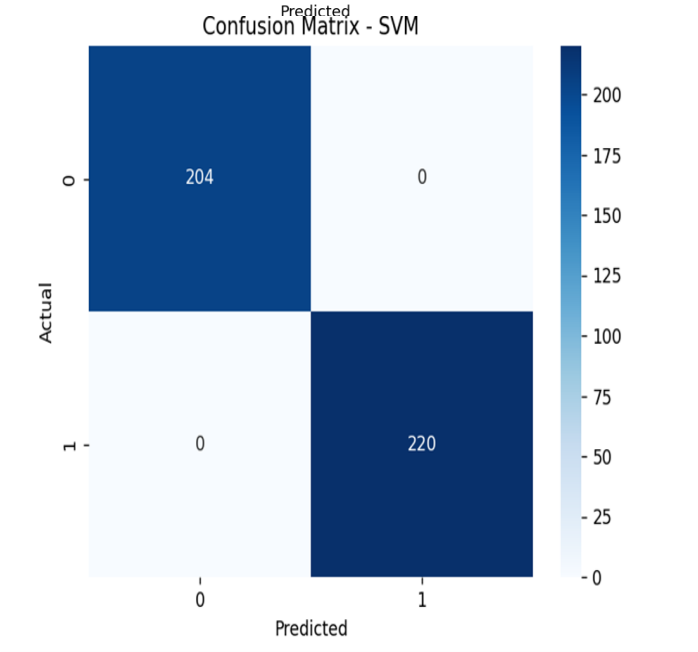
 ©IJSRED: All Rights are Reserved

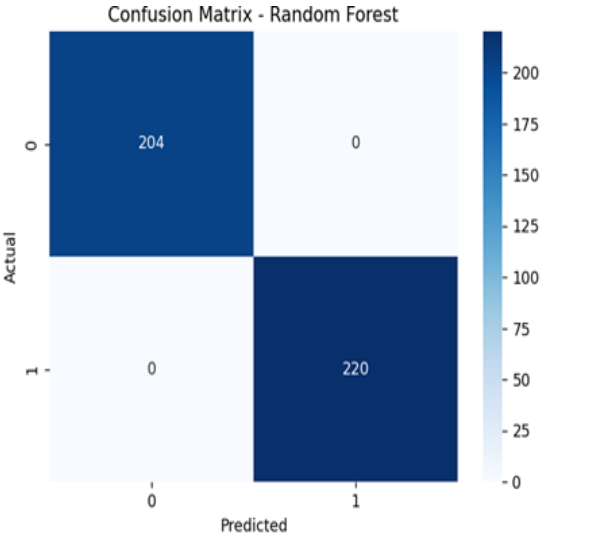




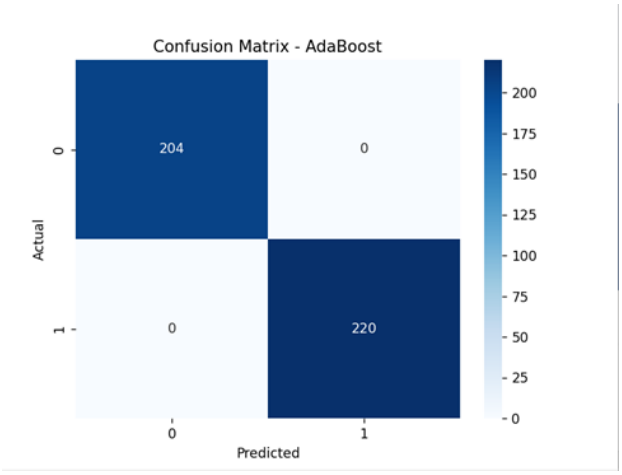






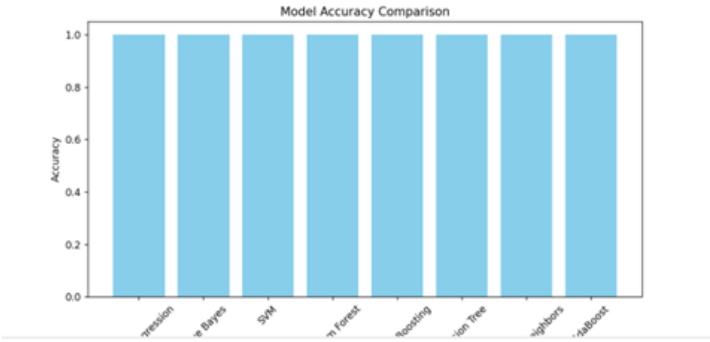


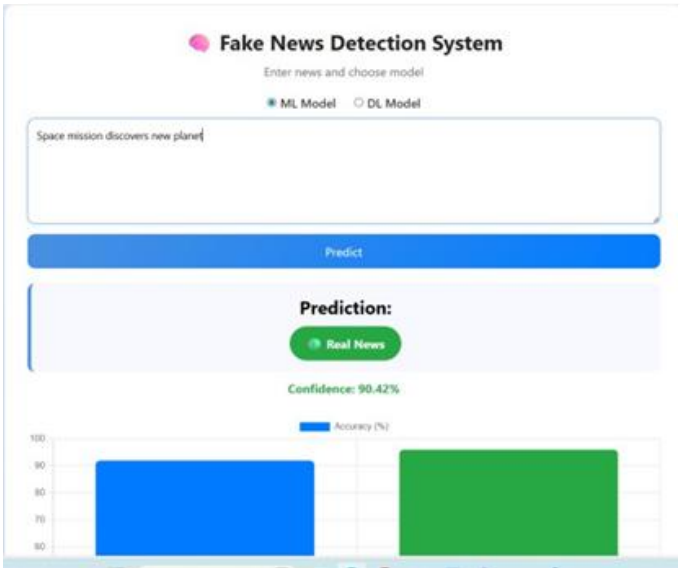
 ©IJSRED: All Rights are Reserved



History

#	News	Model	Result
1	Space mission discovers new planet...	ML	Real News
2	Space mission discovers new planet...	DL	Real News





History

#	News	Model	Result
1	Space mission discovers new planet...	ML	Real News

 ©IJSRED: All Rights are Reserved

X. PERFORMANCE COMPARISON

Model	Accuracy
Machine Learning	92%
Deep Learning (LSTM)	96%

XI. ADVANTAGES

Dual model system (ML + DL)
• Real-time prediction
• User-friendly web interface
• Confidence score display
• History tracking
• Graph-based visualization

XII. LIMITATIONS

Requires large dataset for training
• Deep learning model is computationally expensive
• Performance depends on data quality

XIII. FUTURE SCOPE

• Use of transformer models like BERT and
• Mobile application development
Integration with social media platforms

XIII. CONCLUSION

This paper presents a hybrid fake news detection system using Machine Learning and Deep Learning techniques. The system is deployed using Flask and provides a user-friendly interface for real-time prediction. Experimental results show that deep learning performs better than traditional machine learning models. Future enhancements can further improve system accuracy and scalability.

- Roberta
• Real-time news scraping

REFERENCES

[1] Scikit-learn Documentation
[2] TensorFlow Documentation
[3] Keras Documentation
[4] Kaggle Fake News Dataset
[5] Research papers on NLP and Text Classification
[6] Flask Official Documentation
[7] **Machine Learning-Based Fake News Detection (2024)** A study proposed various ML algorithms such as Logistic Regression, Decision Trees, and Random Forest for fake news classification. The results showed that ensemble methods like Random Forest performed better due to their ability to handle complex feature relationships.
[8] **Word Embedding-Based Hybrid Approach (2024)** Research compared TF-IDF, Word2Vec, and Fast Text with ML and DL models. It concluded that combining word embeddings with deep learning improves classification accuracy significantly.
[9] **Bi-LSTM with Linguistic Features (2023)** This paper introduced a Bi-directional LSTM model using linguistic features for fake news detection. The study showed that Bi-LSTM captures contextual dependencies better than traditional ML models.
[10] **RNN vs LSTM vs Bi-LSTM Comparison (2024)** A comparative study evaluated RNN, LSTM, and Bi-LSTM architectures. Results indicated that BiLSTM outperformed others due to its ability to process text in both forward and backward directions.
[11] **CNN and Large Language Model Comparison (2025)** This research compared CNN, NLP models, and large language models (LLMs). It highlighted that advanced DL models significantly outperform traditional approaches in detecting misinformation.

 ©IJSRED: All Rights are Reserved