

Fake News Detection Using Deep Learning and Text Classification

Bhakti Ashok Chatur
Department Of Data Science
DR.D.Y.Patil ACS Pimpri
Pune,India

Kareena Nathuram Jadhav
Department Of Data Science
DR.D.Y.Patil ACS Pimpri
Pune ,India

Abstract—The rapid growth of digital media and online news platforms has significantly increased the spread of fake and misleading news. This paper presents an automated news detection system using deep learning and text classification techniques. Natural Language Processing methods are applied for text preprocessing. Models such as CNN, LSTM, Transformer, and Hybrid CNN-LSTM are implemented and evaluated. Experimental results show that the Hybrid CNN-LSTM model achieves superior accuracy, making the proposed system reliable and scalable for fake news detection.

Keywords — Fake News Detection; Deep Learning; Text Classification; NLP; CNN; LSTM

I. INTRODUCTION

The rapid expansion of the internet and digital media has transformed the way people consume news and information. Social media platforms, online news portals, and blogs provide instant access to information across the world. However, this convenience has also led to the widespread dissemination of fake and misleading news. Fake news refers to false or manipulated information presented as legitimate news with the intention of misleading readers.

The spread of fake news can influence public opinion, create social unrest, damage reputations, and affect critical events such as elections, public health awareness, and economic decisions. Due to the massive volume of news generated daily, manual verification of news authenticity is impractical and time-consuming. Therefore, there is a strong need for automated and intelligent systems that can accurately detect fake news.

In recent years, Artificial Intelligence (AI) and Deep Learning (DL) techniques have shown promising results in text analysis and classification tasks. Deep learning models are capable of learning complex linguistic patterns, semantic relationships, and contextual information from large-scale textual data. Natural Language Processing (NLP) plays a crucial role in preprocessing and transforming raw text into meaningful representations that can be processed by machine learning models.

This project focuses on detecting fake news using deep learning-based text classification techniques. Multiple models

such as Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM) networks, Transformer-based models, and a Hybrid CNN-LSTM architecture are implemented and evaluated. The objective is to analyze their performance and identify the most effective model for accurate and reliable news detection.

II. EASE OF USE :

The proposed News Detection Using Deep Learning and Text Classification system is designed with simplicity, usability, and efficiency in mind. The system architecture ensures that users can interact with it easily without requiring in-depth technical knowledge of machine learning or natural language processing.

The system follows a structured and well-defined workflow, making it convenient to use and understand. Initially, the user is only required to provide the input news text. Once the input is given, the system automatically handles all internal processes, including text cleaning, preprocessing, feature extraction, and classification. This eliminates the need for manual intervention at intermediate stages.

All preprocessing operations such as tokenization, stopword removal, lemmatization, and text vectorization are performed automatically in the background. The user does not need to configure parameters or adjust settings manually, which reduces complexity and minimizes the chances of error. This automation significantly improves ease of use and saves time.

The classification models, including CNN, LSTM, Transformer, and Hybrid CNN-LSTM, are integrated into a unified framework. Model execution and prediction are handled internally by the system, ensuring smooth operation. The output generated by the system is clear and easy to interpret, indicating whether the given news article is classified as real or fake.

The modular design of the system further enhances usability. Individual components such as data preprocessing, model training, and evaluation are organized in a structured manner. This allows easy modification, maintenance, and future enhancement without affecting the overall functionality of the system.

Overall, the proposed system maintains simplicity while delivering high performance. Its automated processing, minimal user input requirement, and clear output presentation make it suitable for students, researchers, and general users. The ease of use ensures that the system can be effectively utilized for real-world fake news detection applications.

III. EQUATIONS

To evaluate the performance of the proposed news detection system, standard classification metrics are used. These metrics help in analyzing the effectiveness of deep learning models in distinguishing between real and fake news.

Accuracy measures the overall correctness of the model and is defined as the ratio of correctly classified news articles to the total number of articles.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

(1)

Precision indicates how many of the news articles predicted as fake are actually fake. It helps in reducing false positive predictions.

$$Precision = \frac{TP}{TP + FP}$$

(2)

Recall represents the proportion of actual fake news articles that are correctly identified by the model. It focuses on minimizing false negatives.

$$Recall = \frac{TP}{TP + FN}$$

(3)

F1-score provides a balanced evaluation by combining precision and recall. It is especially useful when the dataset is imbalanced.

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

(4)

Where:

TP represents True Positives,
TN represents True Negatives,
FP represents False Positives, and
FN represents False Negatives.

These evaluation metrics are used to compare the performance of CNN, LSTM, Transformer, and Hybrid CNN-LSTM models implemented in the proposed system.

IV. RESULTS AND DISCUSSION :

The performance of the proposed news detection system was evaluated using standard classification metrics such as accuracy, precision, recall, and F1-score. Multiple deep learning models including CNN, LSTM, Transformer, and Hybrid CNN-LSTM were trained and tested on the dataset. The experimental results show that deep learning models outperform traditional machine learning techniques in

detecting fake news. Among all the implemented models, the Hybrid CNN-LSTM model achieved the highest accuracy. This improvement is due to its ability to capture both local textual features through convolutional layers and long-term dependencies through recurrent layers.

CNN models demonstrated strong performance in identifying important word patterns, while LSTM models effectively captured sequential relationships in text data. Transformer-based models provided better contextual understanding but required higher computational resources. The Hybrid CNN-LSTM model balanced accuracy and efficiency, making it the most effective approach for the proposed system.

The results confirm that combining multiple deep learning techniques improves robustness and reliability in fake news classification.

V. CONCLUSION :

Fake news detection has become a critical challenge in the digital era due to the rapid spread of misinformation across online platforms. This study presented an automated news detection system using deep learning and text classification techniques. Various models such as CNN, LSTM, Transformer, and Hybrid CNN-LSTM were implemented and evaluated.

The experimental analysis indicates that deep learning models are highly effective in identifying fake news from textual data. Among the evaluated models, the Hybrid CNN-LSTM model delivered the best performance in terms of accuracy and overall reliability. The system requires minimal user input and performs all processing automatically, making it suitable for real-world applications.

The proposed approach provides a scalable and efficient solution for detecting fake news and can support media verification systems, researchers, and social media monitoring platforms.

VI. FUTURE SCOPE :

Although the proposed system achieves high accuracy, there are several opportunities for future enhancement. The system can be extended to support multilingual news detection to handle content in different languages. Incorporating multimodal data such as images, videos, and social media metadata can further improve detection accuracy.

Future work can also focus on developing lightweight models for real-time deployment on mobile and edge devices. Additionally, explainable artificial intelligence techniques can be integrated to improve transparency and user trust by explaining model predictions. Continuous learning mechanisms can help the system adapt to evolving fake news patterns.

VII. ACKNOWLEDGMENT :

The authors would like to express their sincere gratitude to the Department of Computer Science, Dr. D. Y. Patil Arts, Commerce and Science College, Pimpri, for providing guidance and support throughout the completion of this project. The authors also thank faculty members and peers for their valuable suggestions and encouragement.

REFERENCES :

- [1] Shu, K., Sliva, A., Wang, S., Tang, J., and Liu, H., "Fake News Detection on Social Media: A Data Mining Perspective," ACM SIGKDD Explorations, 2017.
- [2] Ruchansky, N., Seo, S., and Liu, Y., "CSI: A Hybrid Deep Model for Fake News Detection," Proceedings of CIKM, 2017.
- [3] Zhou, X., and Zafarani, R., "A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities," ACM Computing Surveys, 2018.