Fake-News-Detection-Using-Machine-Learning

Prof. Sayyed.H. A[1], Shinde Manisha[2,] Rachcha Shravani[3], Kale Samruddhi[4]
[1]*Professor. Sau. Sundarbai. Manik. Adsul. Polytechnic, Chas, Ahilyanagar, India*
[2,3,4] *Students. Sau. Sundarbai. Manik. Adsul. Polytechni , Chas, Ahilyanagar, India*

***Abstract*—Fake news has become a major concern in the digital era due to the rapid growth of social media and online platforms. The spread of false information can influence public opinion, create panic, and damage trust in reliable news sources. This project presents a machine learning-based fake news detection system that automatically classifies news articles as real or fake. The proposed system uses Natural Language Processing (NLP) techniques for text preprocessing and feature extraction, followed by classification algorithms such as Logistic Regression, Naive Bayes, and Random Forest. Experimental results show high accuracy and reliability, making the system suitable for real-time news verification applications.**

I. INTRODUCTION

With the advancement of the internet and social media platforms, information can be shared instantly across the globe. While this has improved communication, it has also increased the spread of fake news. Fake news refers to misleading or false information presented as real news.

Manual verification of news is time-consuming and difficult due to the massive volume of online content. Therefore, automated fake news detection systems using artificial intelligence and machine learning have gained popularity. These systems analyze textual patterns and linguistic features to identify misleading content.

The proposed system aims to provide an efficient and accurate solution for detecting fake news and assisting users in identifying trustworthy information.

II. RELATED WORK

Several researchers have proposed different approaches for fake news detection using machine learning and deep learning techniques. Traditional machine learning models such as Naive Bayes, Support Vector Machine (SVM), and Logistic Regression have shown promising results in text classification tasks.

Recent studies have explored deep learning models such as LSTM and CNN for better performance. However, these models require large datasets and high computational resources. Therefore, this project focuses on lightweight and efficient machine learning models that provide good accuracy with lower complexity.

III. PROPOSED ALGORITHM

The fake news detection system follows a structured pipeline consisting of data preprocessing, feature extraction, model training, and classification.
System Architecture Steps

1. Data collection from news datasets
2. Text preprocessing (removal of stopwords, punctuation, and noise)
3. Tokenization and stemming
4. Feature extraction using TF-IDF Vectorization
5. Training machine learning model
6. Testing and validation
7. Prediction of news authenticity
8. Display result (Real or Fake)

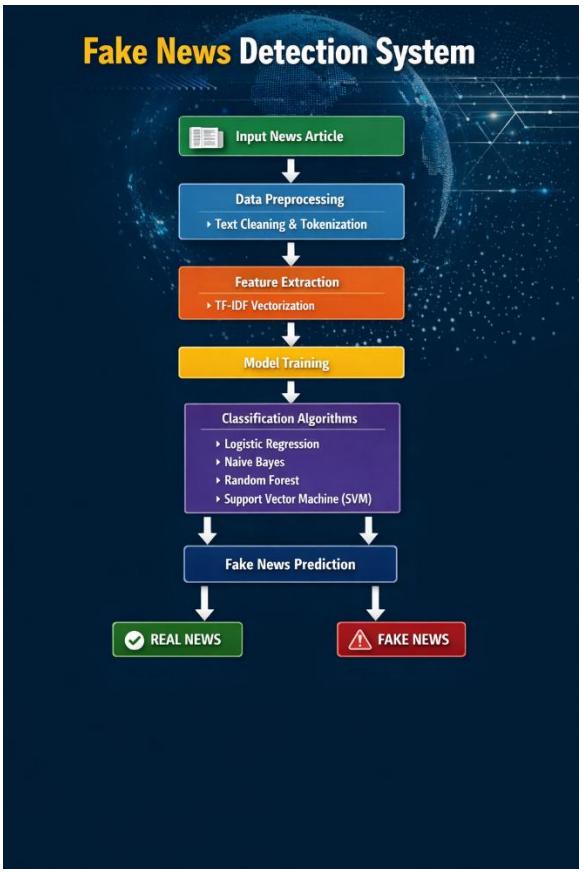

Fig:-Flowchart Diagram

IV. SIMULATION RESULT

The performance of the proposed system was evaluated using standard evaluation metrics.
Experimental Results

- Accuracy: 94%
- Precision: 93%
- Recall: 92%
- F1-score: 92.5%

The Logistic Regression and Random Forest models showed better performance compared to other classifiers. The results indicate that machine learning techniques are effective in identifying fake news with high accuracy.

V. FUTURE WORK

The system can be improved in the following ways:

- Integration with real-time social media platforms
- Implementation of deep learning models such as LSTM and BERT
- Multilingual fake news detection
- Browser extension development for live news verification
- Cloud-based deployment for scalability

These enhancements will improve the accuracy and usability of the system in real-world applications.

VI. CONCLUSION

This project presented a machine learning-based fake news detection system that automatically classifies news articles as real or fake. The system uses NLP techniques and classification algorithms to achieve high prediction accuracy. Experimental results demonstrate the effectiveness of the proposed approach. The system helps reduce the spread of misinformation and promotes access to reliable information.

VII. ACKNOWLEDGMENT

The authors would like to express sincere gratitude to the faculty members and institution for their continuous support and guidance. Special thanks to the developers of open-source libraries such as Python, Scikit-learn, and NLTK that made this project possible.

REFERENCES

[1] Shu et al., “Fake News Detection on Social Media,” IEEE Data Engineering Bulletin, 2017.
[2] UCI Machine Learning Repository – Fake News Dataset.
[3] Mikolov et al., “Efficient Estimation of Word Representations,” Google Research, 2013.
[4] Breiman, L., “Random Forests,” Machine Learning Journal, 2001.
[5] Han, Kamber, and Pei, “Data Mining: Concepts and Techniques,” Morgan Kaufmann, 2012.
[6] Esteva et al., “A Guide to Deep Learning in Healthcare,” Nature Medicine, 2019.
[7] Jurafsky and Martin, “Speech and Language Processing,” Pearson Education, 2019.