International Student Conference on Next-Gen Computing: Application of AI, Big Data, Quantum Computing, Signal Processing and Cloud Innovations (ICNGC-2026)

International Journal of Science, Engineering and Technology ISSN: 2348-4098, P-ISSN: 2395-4752

AI-Based Fake News Detection Using Machine Learning and Explainable AI

Ms. Sayana Garudik[1], Ms. Pari Purohit[2], Mrs. Neeku Sahu[3], Mrs. Shruti Mehta[4]

[1,2,3]B.tech ;CSE CSIT, Durg

[4]Assistant Professor, CSE CSIT, Durg

Abstract- Fake news has become a serious issue in the digital world, spreading misinformation rapidly through social media and online platforms. Detecting fake news manually is difficult due to the large volume of data. In this research, we propose an AI-based fake news detection system using machine learning (ML) models. The dataset is preprocessed and multiple ML algorithms such as Logistic Regression (LR), Random Forest (RF), Support Vector Machine (SVM), Naive Bayes (NB), K-Nearest Neighbors (KNN), and XGBoost (XGB) are applied. Feature selection techniques are used to improve accuracy, and Explainable AI tools like SHAP and LIME are used to interpret model predictions. The best-performing model is deployed for real-time fake news detection through web-based applications.

Keywords: Fake News, Machine Learning, NLP, XAI, SHAP, LIME, Social Media, Classification

I.INTRODUCTION

In recent years, fake news has become a major challenge in digital communication. Social media platforms allow information to spread quickly, but they also enable the rapid spread of misinformation. Fake news can influence public opinion, politics, and social stability.

According to recent studies, a large percentage of online news is either misleading or completely false. Traditional methods of detecting fake news are slow and inefficient. Therefore, automated systems using Artificial Intelligence (AI) and Machine Learning (ML) are needed.

Machine learning models can analyze large datasets and identify patterns that distinguish fake news from real news. In this research, we develop an AI-based system that detects fake news using multiple ML models and improves performance through feature selection and hyperparameter tuning.

Key Contributions

- Developed an AI-based model for fake news detection
- Applied multiple ML algorithms for comparison
- Used feature selection techniques to improve performance
- Applied SHAP and LIME for model explainability
- Proposed deployment using web/mobile application

II. LITERATURE REVIEW

Several researchers have worked on fake news detection using machine learning and deep learning techniques.

Some studies used Logistic Regression and Naive Bayes for text classification but achieved moderate accuracy. Others applied deep learning models like LSTM and CNN, which improved accuracy but required high computational power.

International Journal of Science, Engineering and Technology ISSN: 2348-4098, P-ISSN: 2395-4752

Many works focused only on model accuracy and ignored interpretability. Some studies did not apply feature selection or hyperparameter tuning, leading to overfitting.

In contrast, our work combines:

- Feature reduction
- Hyperparameter tuning
- Explainable AI (SHAP & LIME)
- Real- time deployment

This makes the system more accurate, efficient, and transparent.
Recent advancements in Artificial Intelligence and Natural Language Processing (NLP) have significantly improved fake news detection systems. NLP techniques help machines understand textual content by analyzing sentence structure, keywords, semantics, and writing patterns. Researchers are increasingly combining NLP with machine learning models to improve classification accuracy and detect misleading information more efficiently.

Several studies have applied deep learning models such as Long Short-Term Memory (LSTM), Recurrent Neural Networks (RNN), and Convolutional Neural Networks (CNN) for fake news detection. These models automatically extract hidden features from news content and provide better performance compared to traditional machine learning approaches. However, deep learning models often require large datasets and high computational resources.

Social media platforms such as Facebook, Instagram, WhatsApp, and Twitter have become major sources for spreading fake news. Due to rapid sharing and reposting mechanisms, false information reaches millions of users within a short time. Researchers emphasize the importance of real-time fake news detection systems to reduce misinformation and maintain digital trust.

Feature extraction techniques such as TF-IDF, Bag of Words (BoW), and Word Embeddings are commonly used in fake news detection systems. These methods convert textual information into numerical representations that can be processed by machine learning models. Proper feature extraction improves the overall performance and efficiency of the detection system.
Explainable Artificial Intelligence (XAI) techniques such as SHAP and LIME are gaining importance in modern AI systems. These techniques help users understand why a particular news article is classified as fake or real. Explainability improves transparency, increases trust in AI systems, and helps researchers identify model biases.

Despite significant improvements in fake news manipulated media. Future systems should focus on multilingual detection, real-time monitoring, and integration with social media platforms for better performance and wider applicability.

III. METHODOLOGY

This research follows a structured pipeline including data collection, preprocessing, model training, evaluation, and deployment.

A. Dataset Description

The dataset is collected from Kaggle and contains news articles labeled as real or fake. It includes features such as:

International Journal of Science, Engineering and Technology ISSN: 2348-4098, P-ISSN: 2395-4752

- Title
- Text content
- Author
- Publication date

The dataset consists of thousands of news samples used for training and testing.

B. Data Preprocessing

Data preprocessing is essential for improving model performance.

- Text Clearing : Removal of punctuation, Stopwords,and special characters
- Tokenization: Breaking text into words
- Stemming/Lemmatization: Reducing words to root form
- Vectorization: Using TF-IDF to convert text into numerical form

C. Machine Learning Models

We applied multiple supervised ML models:

Model	Description
LR	Predicts probability of fake or real news
SVM	Finds optimal boundary for classification
RF	Ensemble model using decision trees
KNN	Classifies based on nearest neighbors
NB	Probabilistic classifier
XGB	Boosting model for high accuracy

D. Hyperparameter Tuning

Hyperparameters are optimized using GridSearchCV to improve performance.

Examples :

- Learning rate
- Number of estimators
- Maximum depth

Model	Accuracy	Precision	Recalll	F1
LR	0.92	0.91	0.93	0.92
SVM	0.95	0.94	0.96	0.95
RF	0.97	0.96	0.97	0.97
NB	0.89	0.88	0.90	0.89
XGB	0.98	0.98	0.98	0.98

E. Feature Selection

Feature selection reduces unnecessary data and improves efficiently.

- TF -IDF score used to select important words
- Low -impact features removed

Hyperparameter Tuning Table

Model	Parameter	Value
SVM	Kernel	RBF

International Journal of Science, Engineering and Technology ISSN: 2348-4098, P-ISSN: 2395-4752

RF	Estimators	100
XGB	Learning Rate	0.1

F. Explainable AI (XAI)

To understand model predictions:

- **SHAP** : Shows importance of each word in prediction
- **LIME :** Explains individual predictions locally

These techniques improve transparency and trust.

IV. RESULT AND ANALYSIS

A. Evaluation Metrics

The performance of machine learning models is evaluated using different statistical metrics. These metrics help measure the effectiveness and reliability of the proposed fake news detection system.

We evaluate models using:

- Accuracy
- Precision
- Recall
- F1 -Score

B. Model Performance

XGBoost achieved the best performance

C. Effect of Feature Selection

Feature reduction improved:

- Accuracy
- Speed
- Model efficiency

D. Explainability Results

- SHAP showed important keywords influencing predictions
- LIME explained individual news classifications

E. Deployment

The model can be deployed:

- On a Web Application
- As a Mobile app

Usera can input news text and get instant results.

F. Advantages of Proposed System

- Fast detection
- High accuracy
- Reduces misinformation
- Real-time prediction
- User friendly

International Journal of Science, Engineering and Technology ISSN: 2348-4098, P-ISSN: 2395-4752

G. Future Scope

- Deep learning integration
- Real -Time Twitter Detection
- Multilingual fake news Detection
- Mobile app deployment

COMPARISON TABLE

Author	Model	Accuracy
Sharma et al.	SVM	92%
Kumar et al.	RF	95%
Proposed work	XGB	98%

V. CONCLUSION

Fake news detection is a critical problem in today's digital world. This research presents an AI-based system that effectively detects fake news using machine learning models.

By combining:

- Feature selection
- Hyperparameter tuning
- Explainable AI

The system achieves high accuracy and transparency.

Future work includes:

- Using deep learning models
- Expanding dataset size
- Real- time social media integration.

REFERENCES

1. H. Ahmed, I. Traore, and S. Saad, "Detection of Online Fake News Using N-Gram Analysis and Machine Learning Techniques," Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments, pp. 127–138, 2018.
2. K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake News Detection on Social Media: A Data Mining Perspective," ACM SIGKDD Explorations Newsletter, vol. 19, no. 1, pp. 22–36, 2017.
3. R. Kumar and A. Sharma, "Fake News Detection Using Machine Learning Algorithms," International Journal of Computer Applications, vol. 182, no. 48, pp. 20–25, 2019.
4. S. Kaliyar, A. Goswami, and P. Narang, "FakeBERT: Fake News Detection in Social Media with a BERT-Based Deep Learning Approach," Multimedia Tools and Applications, vol. 80, pp. 11765–11788, 2021.
5. Y. Goldberg, "A Primer on Neural Network Models for Natural Language Processing," Journal of Artificial Intelligence Research, vol. 57, pp. 345–420, 2016.
6. F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.

International Journal of Science, Engineering and Technology ISSN: 2348-4098, P-ISSN: 2395-4752

7. M. T. Ribeiro, S. Singh, and C. Guestrin, "Why Should I Trust You? Explaining the Predictions of Any Classifier," Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1135–1144, 2016.
8. S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," Advances in Neural Information Processing Systems, vol. 30, 2017.
9. J. Brownlee, "Machine Learning Mastery with Python," Machine Learning Mastery, 2019.
10. Kaggle, "Fake News Dataset," [Online]. Available: https://www.kaggle.com. [Accessed: May 2026].
11. T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient Estimation of Word Representations in Vector Space," arXiv preprint arXiv:1301.3781, 2013.
12. A. Graves, "Supervised Sequence Labelling with Recurrent Neural Networks," Springer, 2012.