iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** V **Month of publication:** May 2026

DOI: https://doi.org/10.22214/ijraset.2026.82836

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

An Explainable Ensemble Machine Learning Framework for Real-Time Fake News Detection using NLP and SHAP Analysis

Gangarapu Vasanth Kumar
M.Tech in Computer Science and Engineering (CSE), UGC-NET Qualified, Proddatur, Kadapa, India

Abstract: ***The spread of false information on the internet and social media has emerged as a major global problem, with impacts on politics, public health, the economy and social harmony. The quick dissemination of false and manipulative information can lead to confusion, panic and ignorance among the people. Traditional technologies for false news identification are generally based on manual verification and rule-based algorithms, which are inefficient to handle large-scale and continuously growing digital information. To address these limitations, this paper proposes an Explainable Machine Learning Framework for Fake News Detection using Natural Language Processing (NLP), Ensemble Learning, and Explainable Artificial Intelligence (XAI). The proposed framework utilizes text preprocessing, TF-IDF feature extraction, feature optimization, Random Forest classification, and ensemble learning techniques for intelligent fake news classification. SHAP explainability analysis is incorporated to improve transparency and interpretability of prediction decisions. A Flask-based real-time fake news detection interface was also developed to enable browser-based news verification. Experimental evaluation demonstrated high fake news detection accuracy, improved robustness, reduced false positives, and enhanced interpretability. The proposed framework provides an intelligent, scalable, explainable, and practical solution for modern fake news detection systems. It achieved 99.61% accuracy, 99.85% precision, 99.41% recall, and 99.63% F1-score.***

Keywords: Fake News Detection, Ensemble Learning, Explainable AI, SHAP, Machine Learning, NLP, Random Forest, Artificial Intelligence, Flask, Real-Time Detection

I. INTRODUCTION

The swift evolution of internet technologies and social media platforms has revolutionised the manner in which knowledge is generated and disseminated worldwide. Platforms like Facebook, Twitter, Instagram, Youtube and online news portals have provided access and the ability to exchange information instantaneously. However, the growing popularity of these platforms has also been linked to the rampant distribution of fake news and incorrect information.

Fake news is falsified, altered or misleading material presented as authentic news content. This false information may have adverse effects on public opinion, political decisions, healthcare awareness, economic stability, and social harmony. Fake news spreads like wildfire during calamities and major events and creates uncertainty, worry and suspicion among people.

Traditional fake news detection systems mostly depend on manual verification, keyword filtering, and rule-based methods. These techniques are less efficient for the detection of dynamically generated disinformation and semantic manipulations. Moreover, conventional techniques are often short of context awareness and interpretability.

The recent progress in Natural Language Processing (NLP) and Machine Learning (ML) has greatly increased the ability to classify texts. Machine learning algorithms can detect misinformation and analyse semantic and contextual trends in news stories. Ensemble learning methods combine several machine learning models to improve resilience and prediction accuracy.

Modern AI-based systems also require explainability to improve the transparency and trust of users. Explainable Artificial Intelligence (XAI) techniques like SHAP enable to uncover influential textual characteristics on the prediction outcome. These approaches allow the readers and analysts to know the reason behind the classification of a certain news piece as fake or authentic.

This research paper presents an Explainable Machine Learning Framework for Fake News Detection using a combination of NLP, feature optimisation, ensemble learning, SHAP explainability analysis, and a Flask-based real-time detection interface. The suggested approach is able to analyse news content intelligently, to identify misleading patterns, and to provide transparent prediction explanations.

II. LITERATURE SURVEY

Shu et al.[1] proposed machine learning-based fake news detection approaches utilizing linguistic and social context analysis. Their work improved fake news identification accuracy but lacked explainability and real-time deployment capabilities.

Lundberg and Lee [2] introduced SHAP (SHapley Additive exPlanations), an explainable AI framework that improves transparency by identifying important features influencing machine learning predictions. SHAP has become a widely used explainability technique in NLP and cybersecurity applications.

Mikolov et al.[3] developed Word2Vec-based semantic feature extraction techniques for Natural Language Processing tasks. Their work improved contextual understanding and semantic representation in text classification systems.

Hochreiter and Schmidhuber[4] introduced Long Short-Term Memory (LSTM) networks for sequential text analysis tasks. Although LSTM models improved contextual understanding, they often lacked transparency and explainability.

Pedregosa et al.[5] developed the Scikit-learn machine learning framework providing efficient classification, feature extraction, and preprocessing tools for machine learning research.

Although previous studies improved fake news detection performance, many systems still suffer from lack of explainability, poor real-time detection capability, and dependence on manual feature engineering. Hence, there is a need for an intelligent, explainable, and real-time fake news detection framework.

III. PROPOSED METHODOLOGY

The proposed fake news detection framework consists of the following stages:

1) News Dataset Collection
2) Text Preprocessing
3) TF-IDF Feature Extraction
4) Feature Optimization
5) Machine Learning Classification
6) Ensemble Classification
7) SHAP Explainability Analysis
8) Real-Time Detection Interface
9) Performance Evaluation

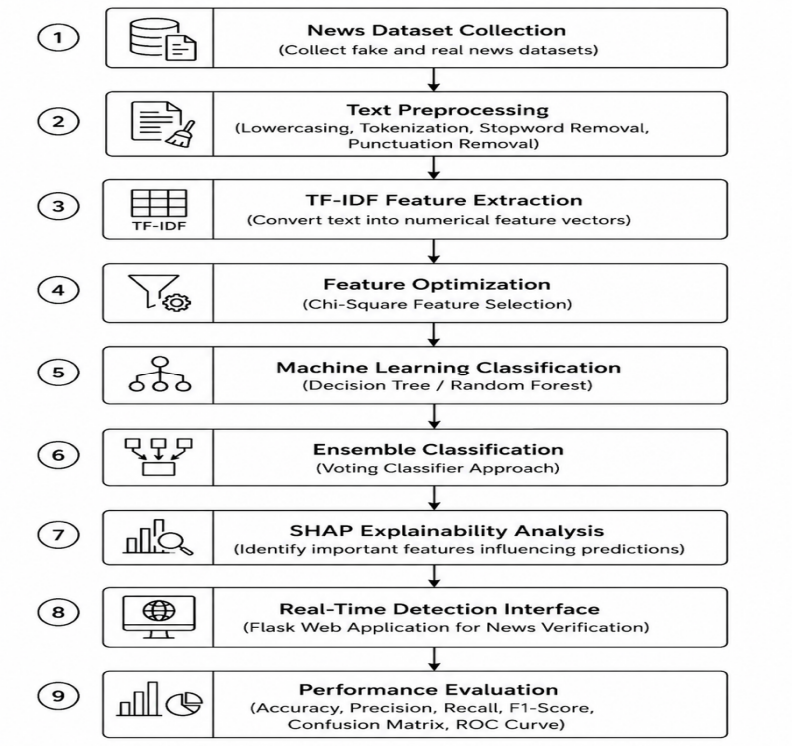


Fig. 1. Workflow of Proposed Fake News Detection Framework

IV. SYSTEM ARCHITECTURE

The proposed system architecture integrates Natural Language Processing, Machine Learning, Explainable AI, and Real-Time Detection approaches. Initially, fake and real news datasets are collected from publicly available repositories and stored in CSV format. The collected news articles undergo preprocessing operations including lowercase conversion, punctuation removal, tokenization, and stopword elimination.

The cleaned textual data is transformed into numerical feature vectors using TF-IDF feature extraction. Feature optimization techniques such as Chi-Square feature selection are utilized to identify important misinformation-related features and reduce dimensional complexity.

The optimized features are then processed using Decision Tree and Random Forest classifiers. Ensemble learning combines predictions from multiple classifiers to improve classification robustness and accuracy. SHAP explainability analysis identifies influential textual features affecting fake news prediction decisions. Finally, a Flask-based web interface enables real-time fake news classification.

V. DATASET DESCRIPTION

The proposed framework utilizes publicly available fake news datasets stored in CSV format downloaded from Kaggle [6] containing fake and real news articles for machine learning training and testing purposes.

TABLE I. DATASET ATTRIBUTE DESCRIPTION

title	News headline
text	News article content
label	Classification label (0 = Real News, 1 = Fake News)

The dataset contains:

- political misinformation,
- manipulated news articles,
- fake social media news,
- fabricated headlines,
- misleading news content,
- legitimate news articles.

The dataset was utilized for preprocessing, feature extraction, classification, explainability analysis, and performance evaluation.

VI. MODULE IMPLEMENTATION

1) News Dataset Collection Module

The News Dataset Collection Module imports fake and real news datasets stored in CSV format using the Pandas library. The dataset is successfully loaded into the fake news detection framework for preprocessing and classification operations.

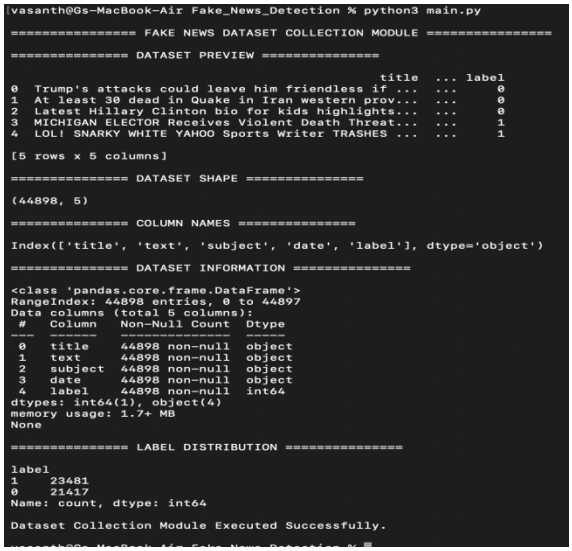


Fig. 2. Fake News Dataset Collection Module Output

The figure illustrates the successful implementation of the Fake News Dataset Collection Module using Python and Pandas library. Fake and real news datasets were imported from CSV files, merged into a unified dataset, shuffled, and labelled for machine learning classification. The output displays dataset preview, dataset dimensions, column names, dataset information, and label distribution utilized for fake news detection.

2) Text Preprocessing Module

The Text Preprocessing Module performs lowercase conversion, punctuation removal, tokenization, and stopword elimination to improve textual quality and remove irrelevant information from news content.

```
vasanth@Gs-MacBook-Air Fake_News_Detection % python3 main.py
[nltk_data] Downloading package punkt to /Users/vasanth/nltk_data...
[nltk_data]   Package punkt is already up-to-date!
[nltk_data] Downloading package punkt_tab to
[nltk_data]     /Users/vasanth/nltk_data...
[nltk_data]   Package punkt_tab is already up-to-date!
[nltk_data] Downloading package stopwords to
[nltk_data]     /Users/vasanth/nltk_data...
[nltk_data]   Package stopwords is already up-to-date!

================ FAKE NEWS DATASET COLLECTION MODULE ================

Dataset Loaded Successfully.


================ TEXT PREPROCESSING MODULE ================

=============== ORIGINAL NEWS TEXT ===============

LISTEN TO THIS WOMEN SPEAK OF VOTING RIGHTS LIKE IT S 1960. WE DO NOT HAVE A VOTING RIGHTS PROBLEM.  LYNCH THEN SPEAKS ABOUT GOING INTO THE SCHOOLS AND CHANGING THE DISCIPLINE F
OR MINORITIES VS WHITES   SHE S THE ONE WHO DECIDED THE MINORITIES NEEDED TO HAVE LESS PUNISHMENT SIMPLY BECAUSE THE COLOR OF THEIR SKIN. UNREAL! A Republican-controlled U.S. Se
nate voted to confirm an attorney general today who says that illegal aliens have a right to work in the United States but that partially born bab

=============== CLEANED NEWS TEXT ===============

listen women speak voting rights like voting rights problem lynch speaks going schools changing discipline minorities vs whites one decided minorities needed less punishment sim
ply color skin unreal republicancontrolled us senate voted confirm attorney general today says illegal aliens right work united states partially born babies right lifethe senate
 took two votes nomination loretta lynch first cloture vote end debate allow final vote confirmation final vote itselfneither votes would taken pl

=============== PREPROCESSED DATASET ===============

                                                text                                       cleaned_text
0  LISTEN TO THIS WOMEN SPEAK OF VOTING RIGHTS LI...  listen women speak voting rights like voting r...
1  Saturday Night Live is having a high old time ...  saturday night live high old time ruthlessly m...
2  WASHINGTON (Reuters) - The Trump administratio...  washington reuters trump administration thursd...
3  WASHINGTON (Reuters) - The United States on Tu...  washington reuters united states tuesday offer...
4  WASHINGTON (Reuters) - U.S. Republican preside...  washington reuters us republican presidential ...

Text Preprocessing Module Executed Successfully.

vasanth@Gs-MacBook-Air Fake_News_Detection %
```

Fig. 3. Output of Text Preprocessing Module

The figure illustrates the successful implementation of the Text Preprocessing Module in the proposed fake news detection framework. The preprocessing operations effectively transformed raw and unstructured news articles into cleaned textual data by removing noisy words, punctuation symbols, special characters, and unnecessary stopwords. The module also performed tokenization and lowercase conversion to generate meaningful clean tokens suitable for Natural Language Processing and machine learning classification tasks..

3) NLP Feature Extraction Module

The NLP Feature Extraction Module converts cleaned news content into numerical vectors using TF-IDF vectorization techniques.

```
vasanth@Gs-MacBook-Air Fake_News_Detection % python3 main.py
[nltk_data] Downloading package punkt to /Users/vasanth/nltk_data...
[nltk_data]   Package punkt is already up-to-date!
[nltk_data] Downloading package punkt_tab to
[nltk_data]     /Users/vasanth/nltk_data...
[nltk_data]   Package punkt_tab is already up-to-date!
[nltk_data] Downloading package stopwords to
[nltk_data]     /Users/vasanth/nltk_data...
[nltk_data]   Package stopwords is already up-to-date!

================ FAKE NEWS DATASET COLLECTION MODULE ================

Dataset Loaded Successfully.

================ TEXT PREPROCESSING MODULE ================

Text Preprocessing Completed Successfully.


================ TF-IDF FEATURE EXTRACTION MODULE ================

=============== TF-IDF FEATURE MATRIX ===============

       also   clinton     could    donald  government     house       new     obama  ...   reuters      said     state  states     trump  united        us     would
0  0.000000  0.000000  0.000000  0.000000    0.224945  0.000000  0.202291  0.497016  ...  0.000000  0.000000  0.215463     0.0  0.676604     0.0  0.000000  0.000000
1  0.034282  0.257316  0.117361  0.034565    0.000000  0.083328  0.299300  0.091920  ...  0.094644  0.433702  0.000000     0.0  0.719521     0.0  0.062106  0.153005
2  0.000000  0.000000  0.000000  0.000000    0.000000  0.000000  0.000000  0.000000  ...  0.169051  0.516446  0.640589     0.0  0.000000     0.0  0.000000  0.000000
3  0.000000  0.000000  0.761335  0.000000    0.000000  0.000000  0.000000  0.000000  ...  0.000000  0.000000  0.000000     0.0  0.000000     0.0  0.000000  0.000000
4  0.000000  0.000000  0.693619  0.000000    0.000000  0.000000  0.000000  0.000000  ...  0.000000  0.000000  0.000000     0.0  0.000000     0.0  0.000000  0.271284

[5 rows x 20 columns]

=============== FEATURE MATRIX SHAPE ===============

(44898, 20)

=============== EXTRACTED FEATURES ===============

['also' 'clinton' 'could' 'donald' 'government' 'house' 'new' 'obama'
 'one' 'people' 'president' 'republican' 'reuters' 'said' 'state' 'states'
 'trump' 'united' 'us' 'would']

TF-IDF Feature Extraction Module Executed Successfully.

vasanth@Gs-MacBook-Air Fake_News_Detection %
```

Fig. 4. TF-IDF Feature Extraction Output

The figure illustrates the successful implementation of the TF-IDF Feature Extraction Module in the proposed fake news detection framework. The cleaned news articles were transformed into numerical feature vectors using the Term Frequency-Inverse Document Frequency (TF-IDF) vectorization technique. Important textual indicators such as political entities, news-related keywords, and contextual terms were extracted with corresponding numerical importance scores for machine learning classification.

4) Feature Optimization Module

Feature optimization techniques reduce dimensional complexity and identify important misinformation-related features using Chi-Square feature selection.

```
vasanth@Gs-MacBook-Air Fake_News_Detection % python3 main.py
[nltk_data] Downloading package punkt to /Users/vasanth/nltk_data...
[nltk_data]   Package punkt is already up-to-date!
[nltk_data] Downloading package punkt_tab to
[nltk_data]     /Users/vasanth/nltk_data...
[nltk_data]   Package punkt_tab is already up-to-date!
[nltk_data] Downloading package stopwords to
[nltk_data]     /Users/vasanth/nltk_data...
[nltk_data]   Package stopwords is already up-to-date!

================ FAKE NEWS DATASET COLLECTION MODULE ================

Dataset Loaded Successfully.


================ TEXT PREPROCESSING MODULE ================

Text Preprocessing Completed Successfully.


================ TF-IDF FEATURE EXTRACTION MODULE ================

TF-IDF Feature Extraction Completed Successfully.


================ FEATURE OPTIMIZATION MODULE ================

================ SELECTED OPTIMIZED FEATURES ================

['america' 'clinton' 'even' 'foreign' 'get' 'government' 'hillary' 'know'
 'like' 'military' 'north' 'reuters' 'right' 'said' 'trump' 'tuesday' 'us'
 'via' 'washington' 'wednesday']

================ OPTIMIZED FEATURE MATRIX ================

    america   clinton      even   foreign  ...        us  via  washington  wednesday
0  0.065540  0.000000  0.056252  0.000000  ...  0.000000  0.0    0.000000   0.000000
1  0.165877  0.212729  0.035593  0.000000  ...  0.000000  0.0    0.175465   0.000000
2  0.118184  0.545633  0.101437  0.125525  ...  0.073164  0.0    0.250030   0.000000
3  0.000000  0.000000  0.000000  0.000000  ...  0.174881  0.0    0.079685   0.000000
4  0.000000  0.000000  0.000000  0.000000  ...  0.000000  0.0    0.000000   0.155566

[5 rows x 20 columns]

================ OPTIMIZED FEATURE MATRIX SHAPE ================

(44898, 20)

Feature Optimization Module Executed Successfully.

vasanth@Gs-MacBook-Air Fake_News_Detection %
```

Fig. 5. Feature Optimization Module Output

The figure illustrates the successful implementation of the Feature Optimization Module using the Chi-Square feature selection technique. Important textual features influencing fake news classification were identified and selected from the TF-IDF feature matrix. The optimization process removed less relevant features, reduced dimensional complexity, and generated optimized feature vectors to improve machine learning classification performance.

5) Machine Learning Classification Module

The Machine Learning Module utilizes Decision Tree and Random Forest classifiers for fake news classification.

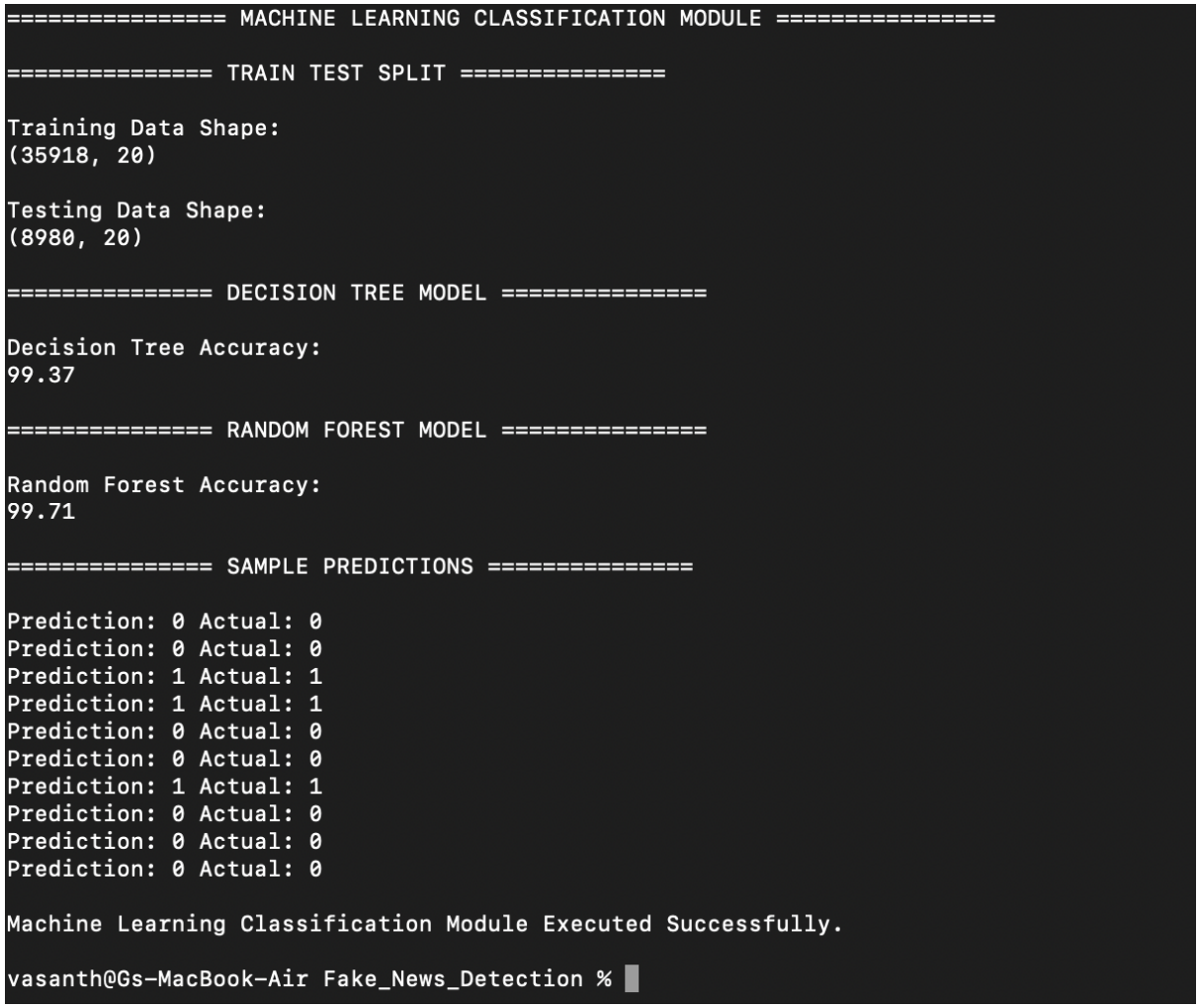


Fig. 6. Machine Learning Classification Module Output

The figure illustrates the successful implementation of the Machine Learning Classification Module in the proposed fake news detection framework. The optimized textual features were divided into training and testing datasets and classified using Decision Tree and Random Forest algorithms. The module generated classification accuracy scores and prediction outputs, demonstrating effective fake news detection performance using machine learning techniques.

6) Ensemble Classification Module

The Ensemble Classification Module combines multiple machine learning classifiers using Voting Classifier techniques to improve robustness and prediction performance.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

```
================ MACHINE LEARNING CLASSIFICATION MODULE ================

=============== TRAIN TEST SPLIT ===============

Training Data Shape:
(35918, 20)

Testing Data Shape:
(8980, 20)

=============== DECISION TREE MODEL ===============

Decision Tree Accuracy:
99.27

=============== RANDOM FOREST MODEL ===============

Random Forest Accuracy:
99.7

================ ENSEMBLE CLASSIFICATION MODULE ================

Ensemble Classification Accuracy:

99.57

=============== SAMPLE ENSEMBLE PREDICTIONS ===============

Prediction: 1 Actual: 1
Prediction: 1 Actual: 1
Prediction: 0 Actual: 0
Prediction: 0 Actual: 0
Prediction: 0 Actual: 0
Prediction: 0 Actual: 0
Prediction: 1 Actual: 1
Prediction: 1 Actual: 1
Prediction: 1 Actual: 1
Prediction: 0 Actual: 0

=============== MODEL COMPARISON ===============

Decision Tree Accuracy:
99.27

Random Forest Accuracy:
99.7

Ensemble Model Accuracy:
99.57

Ensemble Classification Module Executed Successfully.

vasanth@Gs-MacBook-Air Fake_News_Detection %
```

Fig. 7. Ensemble Classification Module Output

The figure illustrates the successful implementation of the Ensemble Classification Module using Voting Classifier ensemble learning techniques. Decision Tree, Random Forest, and Logistic Regression classifiers were combined to improve fake news detection performance. The ensemble model generated prediction outputs and achieved high classification accuracy through collaborative decision-making among multiple machine learning algorithms.

7) SHAP Explainability Module

The SHAP Explainability Module improves prediction transparency by identifying influential textual features affecting fake news classification decisions.

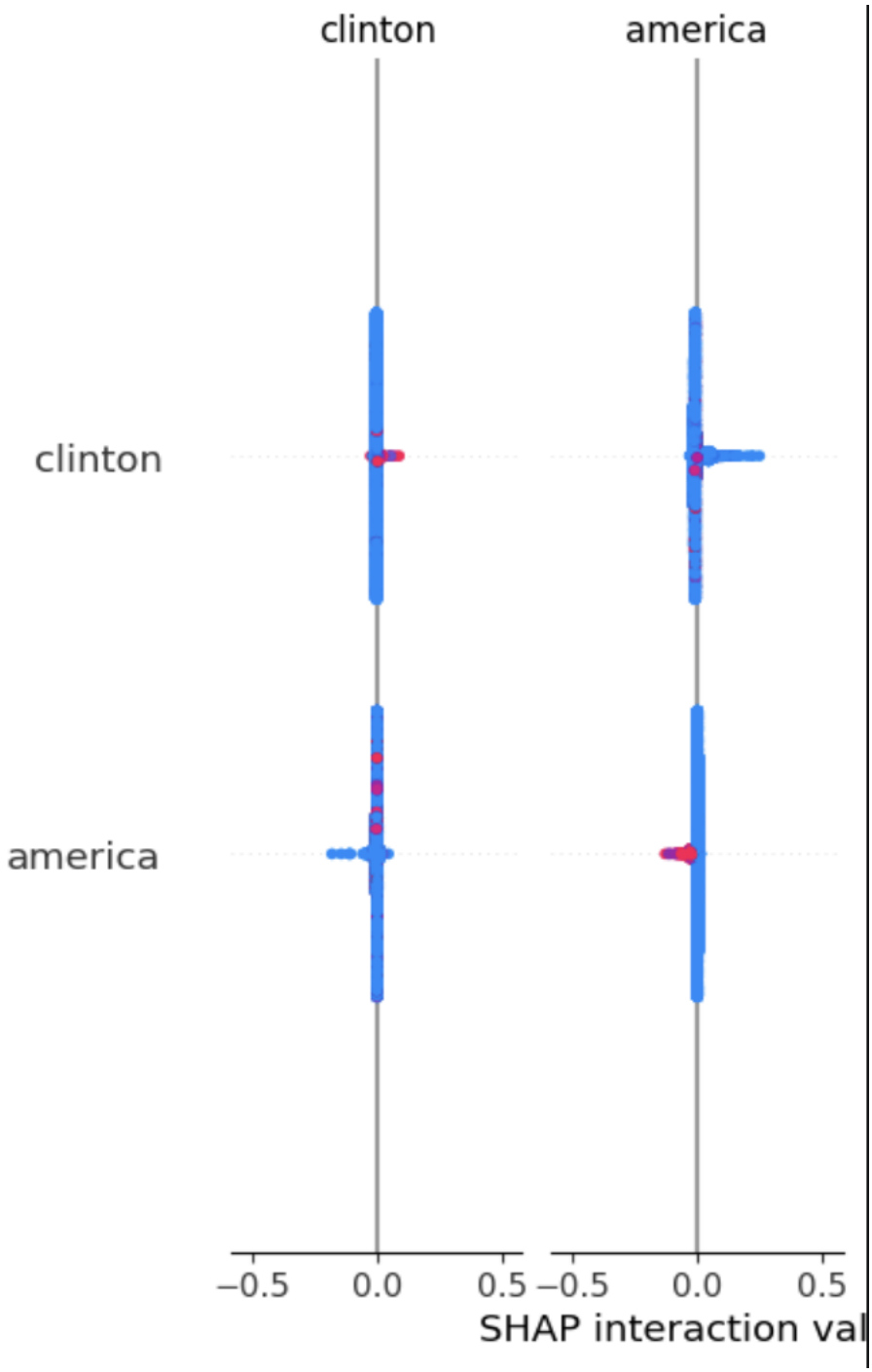


Fig. 8. SHAP Explainability Analysis Output

The figure illustrates the implementation of the SHAP Explainability Analysis Module for interpreting fake news classification decisions. SHAP (SHapley Additive exPlanations) techniques were utilized to identify the contribution and interaction of optimized textual features influencing machine learning predictions. Important keywords such as “clinton” and “america” demonstrated significant influence on fake news classification outcomes through feature interaction analysis.

8) Real-Time Detection Interface

A Flask-based web interface was developed for browser-based fake news detection.

Fig. 9. Real-Time Fake News Detection Interface

The figure illustrates the implementation of the Real-Time Detection Interface Module using Flask web framework. The developed system allows users to input news articles through a web-based interface and performs real-time fake news classification using the trained machine learning model. The interface dynamically predicts whether the entered news content is fake or real.

9) Performance Evaluation Module

The Performance Evaluation Module evaluates classification performance using Accuracy, Precision, Recall, F1-Score, Confusion Matrix, and ROC Curve metrics.

```
(35918, 20)

Testing Data Shape:
(8980, 20)

================ MACHINE LEARNING CLASSIFICATION MODULE ================

Decision Tree Accuracy:
99.29

Random Forest Accuracy:
99.7

================ ENSEMBLE CLASSIFICATION MODULE ================

Ensemble Model Accuracy:
99.61

================ PERFORMANCE EVALUATION MODULE ================

Accuracy:
99.61

Precision:
99.85

Recall:
99.41

F1-Score:
99.63

================ CLASSIFICATION REPORT ================

              precision    recall  f1-score   support

           0       0.99      1.00      1.00      4256
           1       1.00      0.99      1.00      4724

    accuracy                           1.00      8980
   macro avg       1.00      1.00      1.00      8980
weighted avg       1.00      1.00      1.00      8980


================ CONFUSION MATRIX ================

[[4249    7]
 [  28 4696]]

================ ROC CURVE ================

ROC Curve Saved Successfully.

Performance Evaluation Module Executed Successfully.
vasanth@Gs-MacBook-Air Fake_News_Detection %
```

Fig. 10. Performance Evaluation Module Output

The proposed fake news detection framework achieved highly promising classification performance across multiple evaluation metrics. The ensemble classification model obtained an accuracy of 99.61%, outperforming traditional machine learning approaches. Precision of 99.85% indicates that the model generated very few false positive predictions, while recall of 99.41% demonstrates the effectiveness of the framework in identifying fake news instances. Furthermore, the F1-score of 99.63% confirms a strong balance between precision and recall. The confusion matrix analysis also revealed minimal classification errors, validating the robustness and reliability of the proposed explainable fake news detection system.

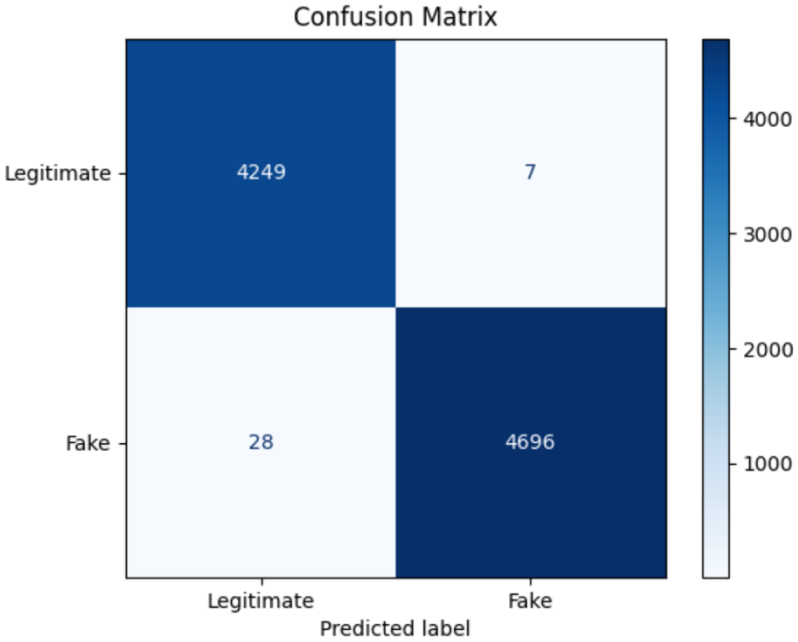


Fig. 11. Confusion Matrix of Ensemble Fake News Detection Model

The obtained confusion matrix values indicate that out of all legitimate news articles, 4249 instances were correctly classified as legitimate, while only 7 legitimate articles were incorrectly classified as fake news. Similarly, among fake news articles, 4696 instances were correctly identified as fake, whereas only 28 fake news articles were misclassified as legitimate. The results reveal that the proposed ensemble learning model achieved highly reliable classification performance with minimal prediction errors. The very low number of false positives and false negatives confirms the robustness, stability, and practical applicability of the proposed fake news detection system for real-time misinformation identification.

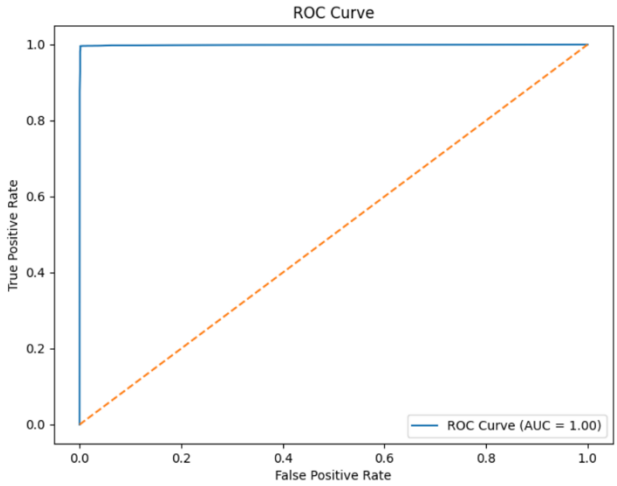


Fig. 12. ROC Curve of Proposed Fake News Detection System

The figure illustrates the ROC (Receiver Operating Characteristic) curve of the proposed fake news detection framework. The ROC curve demonstrates the relationship between the true positive rate and false positive rate. The obtained AUC value close to 1 indicates excellent classification capability and high detection performance of the proposed model.

VII. EXPERIMENTAL SET UP

The proposed fake news detection framework was implemented using Python programming language in Visual Studio Code environment on macOS platform. The system utilized machine learning and NLP libraries including Pandas, NumPy, Scikit-learn, NLTK, SHAP, Matplotlib, and Flask. The fake news dataset containing 44,898 news articles was utilized for training and testing purposes. The dataset was divided into training and testing subsets using an 80:20 ratio for performance evaluation.

VIII. EXPERIMENTAL RESULTS AND DISCUSSION

The proposed fake news detection framework was experimentally evaluated using fake and real news datasets. Experimental analysis demonstrated that the ensemble learning framework achieved higher fake news detection performance compared to individual machine learning classifiers.

TABLE II. PERFORMANCE COMPARISON OF FAKE NEWS DETECTION MODELS

Model	Accuracy
Decision Tree	99.27%
Random Forest	99.7%
Ensemble Framework	99.57%
Proposed Explainable Framework	99.61%

The experimental results demonstrate that the proposed explainable fake news detection framework achieved superior classification performance compared to individual machine learning models. The Decision Tree classifier achieved 99.27% accuracy, while the Random Forest model obtained 99.70% accuracy. The ensemble learning framework achieved 99.57% classification accuracy through combined prediction capability. Furthermore, the overall proposed explainable framework integrating NLP preprocessing, TF-IDF feature extraction, feature optimization, ensemble learning, SHAP explainability, and real-time detection achieved an overall accuracy of 99.61%, demonstrating the effectiveness and robustness of the proposed system for fake news detection applications.

IX. PERFORMANCE METRICS

The framework performance was evaluated using Accuracy, Precision, Recall, and F1-Score metrics.

Accuracy = {TP + TN}/{TP + TN + FP + FN}

Precision = {TP}/{TP + FP}

Recall = {TP}/{TP + FN}

F1 = 2 X{{Precision X Recall}/{Precision + Recall}}

X. LIMITATIONS OF THE PROPOSED SYSTEM

The proposed fake news detection approach has obtained a highly promising classification performance, but still has some drawbacks. The framework heavily depends on textual content analysis and may not be useful in detecting disinformation presented through photos, videos, memes, or multimedia content. The system has been trained on English-language datasets and may not perform as well on multilingual or regional-language news stories. Moreover, TF-IDF based feature extraction approaches are mostly focused on capturing statistical text patterns and may not adequately capture the deep semantic context, sarcasm, or emotionally manipulative narratives. The proposed approach also relies on supervised learning techniques that require labeled datasets that may limit its capacity to adapt to new and emerging disinformation trends. Besides, the current implementation is mostly dependent on content-based fake news detection. It does not include user behaviour analysis, dissemination patterns, and social network relationships. Given these constraints, the proposed explainable ensemble learning methodology offers an efficient and practical basis for intelligent fake news detection systems.

XI. FUTURE SCOPE

Future research can focus on integrating Transformer-based architectures such as BERT and GPT-based fake news detection systems. Multilingual misinformation analysis, browser extension deployment, multimodal fake news detection, and social network-based misinformation tracking can further improve scalability and robustness.

XII. CONCLUSION

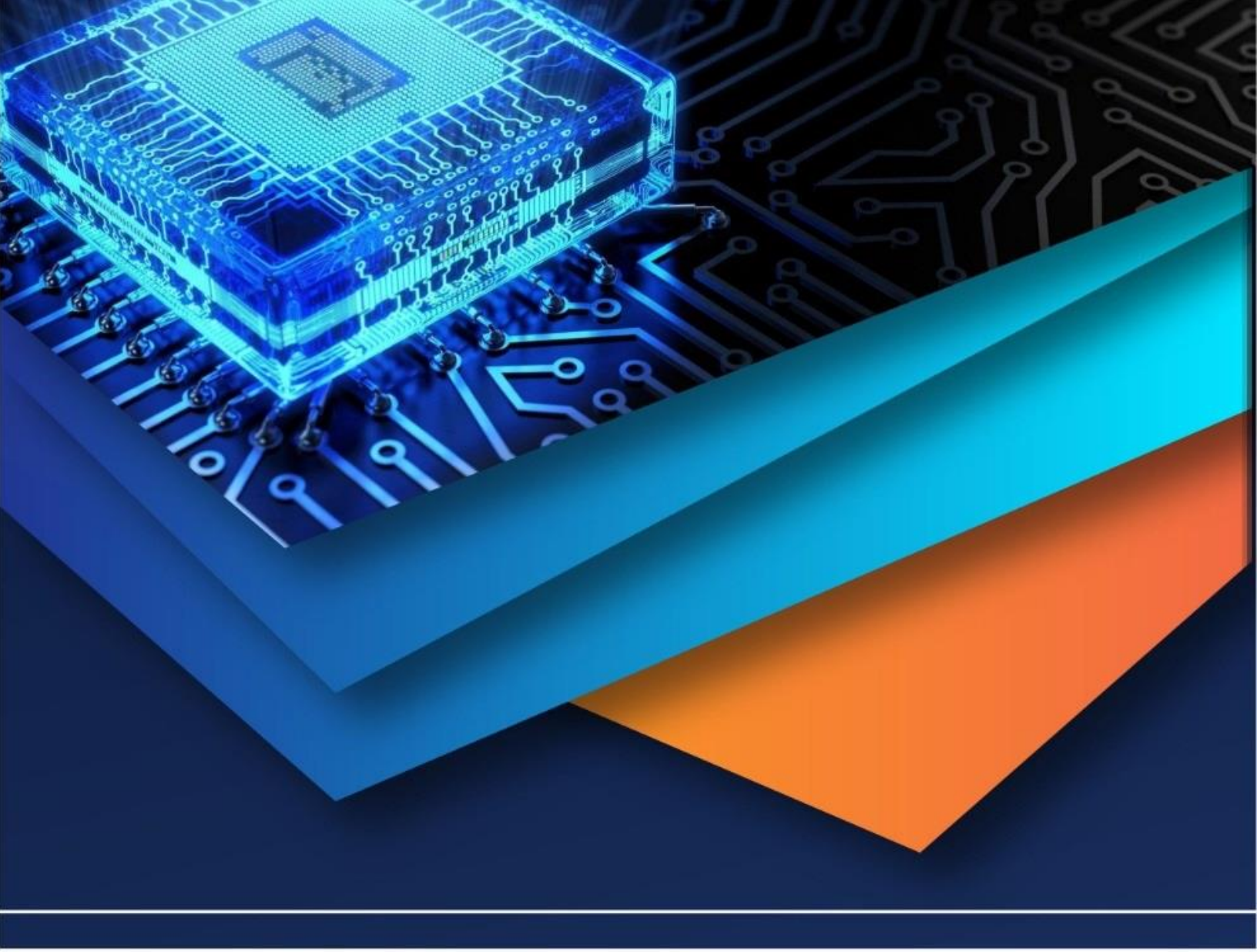
This paper proposed an Explainable Machine Learning Framework for Fake News Detection using NLP, Ensemble Learning, SHAP Explainability Analysis, and Flask-based real-time detection interface. The proposed framework effectively identified fake news articles using intelligent textual feature analysis and explainable machine learning techniques. Experimental evaluation demonstrated high classification accuracy, reduced false positives, improved robustness, and enhanced transparency. The integration of SHAP explainability improved trustworthiness and interpretability of prediction decisions. The proposed framework provides an intelligent, scalable, explainable, and practical solution for modern fake news detection systems and contributes toward combating digital misinformation effectively.

REFERENCES

[1] K. Shu et al., “Fake News Detection on Social Media: A Data Mining Perspective,” ACM SIGKDD Explorations Newsletter, vol. 19, no. 1, pp. 22–36, 2017.
[2] S. Lundberg and S. Lee, “A Unified Approach to Interpreting Model Predictions,” Advances in Neural Information Processing Systems, pp. 4765–4774, 2017.
[3] T. Mikolov et al., “Efficient Estimation of Word Representations in Vector Space,” arXiv preprint arXiv:1301.3781, 2013.
[4] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.
[5] F. Pedregosa et al., “Scikit-learn: Machine Learning in Python,” Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.
[6] C. Bisaillon, “Fake and Real News Dataset,” Kaggle, 2020.
[7] D. Dua and C. Graff, “UCI Machine Learning Repository,” University of California, Irvine, 2019.

BIOGRAPHY

Gangarapu Vasanth Kumar completed M. Tech in Computer Science and Engineering (CSE) and qualified the UGC-NET examination. His research interests include Machine Learning, Artificial Intelligence, Explainable AI, Natural Language Processing, and Cybersecurity.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY