International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Defensive Modeling of Fake News Through Online

Mrs.D.Kanakasatya[1], Pasala Jyoshna Vijaya Sri[2], Neela Harshitha[3], Sunkara Ganga Kameswari[4], Tripurari Naveen Kumar[5], Nukireddy Harsha Vamsi[6],

[1]Assistant Professor, CSE Department, Vsm College Of Engineering Ramachandrapuram, Dr.B.R.Ambhedkar Konaseema Dt, Andhra Pradesh-533255
[2,3,4,5,6]Student, CSE Department, Vsm College Of Engineering Ramachandrapuram, Dr.B.R.Ambhedkar Konaseema Dt, Andhra Pradesh-533255

ABSTRACT
The rapid growth of online social networks has led to an unprecedented surge in fake news, driven by commercial, political, and ideological motives. This misinformation significantly impacts societal trust, public perception, and even democratic processes. This project aims to develop a robust, scalable system for detecting fake news on platforms like Facebook using advanced machine learning techniques. We propose a Random Forest Classifier integrated with Term Frequency-Inverse Document Frequency (TF-IDF) vectorization to claify news articles as real or fake with high accuracy. The system is trained on a comprehensive labeled dataset containing diverse news articles and evaluated using metrics such as precision, recall, and F1-score. A user-friendly graphical user interface (GUI) developed with Tkinter enables real-time fake news prediction, making the system accessible to non-technical users. Furthermore, the model's generalizability is tested on external datasets, including a sample set of news claims, ensuring its robustness across varied contexts. The results, achieving a test accuracy of 75%, underscore the potential of machine learning to mitigate the spread of misinformation effectively, paving the way for enhanced information integrity in digital ecosystems.

INTRODUCTION
The proliferation of fake news on social media platforms has emerged as a critical global challenge, undermining public trust, fueling polarization, and influencing political and social outcomes. Fake news, often crafted for financial gain through clickbait or as propaganda to sway opinions, exploits the viral nature of platforms like Facebook, where content spreads rapidly with minimal gatekeeping. This project addresses this pressing issue by developing an automated fake news detection system powered by a Random Forest Classifier. Leveraging natural language processing (NLP) techniques, the system analyzes textual content to differentiate between real and fake news with high precision. A key feature is the integration of a Tkinter-based GUI, which allows users to input news text and receive instant predictions, enhancing accessibility for diverse audiences. The system is trained on a labeled dataset ('fake_news_dataset.csv') and tested on external data ('new_dataset.csv') to ensure applicability across various sources. This study also explores the limitations of existing methods and proposes a scalable solution to combat misinformation, contributing to a more trustworthy digital information landscape.

EXISTING SYSTEM

Existing research on fake news detection has explored a range of machine learning approaches, including Naive Bayes, Support Vector Machines (SVM), and deep learning models like recurrent neural networks. Techniques such as n-gram analysis and shallow syntactic tagging have been widely used but often fail to capture the nuanced linguistic and contextual patterns of fake news, resulting in limited accuracy. For instance, Conroy et al. (2015) demonstrated that combining shallow syntactic methods with deep syntactic analysis, such as Probabilistic Context-Free Grammars, could achieve accuracies of 85–91% in specific deception detection tasks. However, these methods struggle with the unstructured, dynamic, and often ambiguous nature of social media content. Recent studies have also explored social signal-based detection, analyzing user interactions and propagation patterns, but these require extensive metadata, which may not always be available. The proposed system builds on these insights, employing a Random Forest Classifier with TF-IDF vectorization to capture both lexical and contextual features, offering a more robust and computationally efficient solution for fake news detection.

PROBLEM STATEMENT

The rampant spread of fake news on social media platforms like Facebook poses a significant threat to information integrity, public trust, and societal stability. Traditional detection methods, such as manual fact-checking or basic n-gram models, are either too slow or insufficiently accurate to handle the volume and complexity of online content. These methods often fail to capture the subtle linguistic and contextual cues that distinguish fake news from legitimate articles. Moreover, the dynamic and user-generated nature of social media data complicates automated detection efforts. There is an urgent need for an automated, accurate, and scalable system capable of classifying news as real or fake in real-time. This project addresses this gap by implementing a Random Forest Classifier with TF-IDF vectorization, complemented by a user-friendly GUI, to enable efficient and accessible fake news detection. The system is designed to generalize across diverse datasets, ensuring its practical utility in combating misinformation.

PROPOSED SYSTEM

The proposed system leverages a Random Forest Classifier to detect fake news with high accuracy, trained on a labeled dataset of news articles (‘fake_news_dataset.csv’). Text data is preprocessed using TF-IDF vectorization, which transforms raw text into numerical features by weighting words based on their frequency and importance across the dataset. The system incorporates a Tkinter-based GUI, enabling users to input news text and receive instant predictions, making it accessible to both technical and non-technical users. The model is evaluated using standard metrics (precision, recall, F1-score) and tested on an external dataset (‘new_dataset.csv’) to ensure generalizability. Key components include data preprocessing, feature extraction, model training, and real-time prediction. The system’s architecture is modular, allowing for future enhancements, such as integration with deep learning models or multimedia analysis, to further improve detection capabilities. By combining robust machine learning with an intuitive interface, the system aims to provide a practical solution for mitigating the spread of fake news.

IMPLEMENTED ALGORITHMS

a) Random Forest Classifier

The Random Forest Classifier is an ensemble learning algorithm that constructs multiple decision trees during training and aggregates their predictions to enhance accuracy and robustness. It excels in handling

high-dimensional data, such as TF-IDF vectors, and is less susceptible to overfitting compared to single decision trees. In this project, the classifier is configured with 100 trees and balanced class weights to address the dataset's imbalance (326 fake vs. 230 real samples). The model is trained on TF-IDF features and achieves a test accuracy of 75%, making it well-suited for text classification tasks like fake news detection.

b) TF-IDF Vectorization

Term Frequency-Inverse Document Frequency (TF-IDF) is a text preprocessing technique that converts raw text into numerical features. It assigns higher weights to words that are frequent in a specific document but rare across the dataset, emphasizing discriminative terms. Parameters such as stop_words='english', max_df=0.95, and ngram_range=(1, 2) are used to remove common words, limit overly frequent terms, and include bigrams for richer feature representation. This approach ensures the model captures both lexical and contextual patterns critical for distinguishing real and fake news.

WORKING OF MACHINE LEARNING ALGORITHMS

The fake news detection workflow begins with loading the dataset ('fake_news_dataset.csv'), which contains 569 news articles labeled as real (0) or fake (1) after removing 13 samples with missing labels. Preprocessing involves cleaning the data (removing duplicates and missing values) and splitting it into training (70%, 398 samples) and testing (30%, 167 samples) sets. The text is transformed into TF-IDF vectors using scikit-learn's TfidfVectorizer, with parameters optimized to capture significant terms and bigrams. The Random Forest Classifier is trained on the training set, tuning hyperparameters like the number of trees and class weights to maximize performance. The model's performance is evaluated on the test set using metrics such as precision (0.77), recall (0.75), and F1-score (0.73), with a confusion matrix showing 33 true negatives, 92 true positives, 35 false positives, and 7 false negatives. The trained model and vectorizer are saved using joblib for reuse. The Tkinter GUI enables real-time predictions by vectorizing user-input text and classifying it. An external dataset ('new_dataset.csv', 5 samples) is used to test generalizability, correctly classifying claims like "The moon landing was faked by NASA" as fake and "New technology is transforming healthcare" as real.

SYSTEM ARCHITECTURE

The system architecture integrates multiple components for seamless operation:

1. **Data Input Module:** Administrators upload datasets (e.g., CSV files with article text and labels) via a web interface.
2. **Preprocessing Module:** Cleans data by removing missing values and applies TF-IDF vectorization using scikit-learn to convert text into numerical features.
3. **Algorithm Implementation:** Trains a Random Forest Classifier with 100 trees and balanced class weights, optimized for text classification.
4. **Prediction Module:** Outputs classification results (fake/real) with confidence scores, displayed via Flask. Generates real-time predictions through a Tkinter GUI and batch predictions for external datasets, with results displayed as "Real" or "Fake."
5. **User Interaction:** Provides an intuitive GUI for users to input news text and view predictions, enhancing accessibility.

Fig: 1 System Architecture Diagram

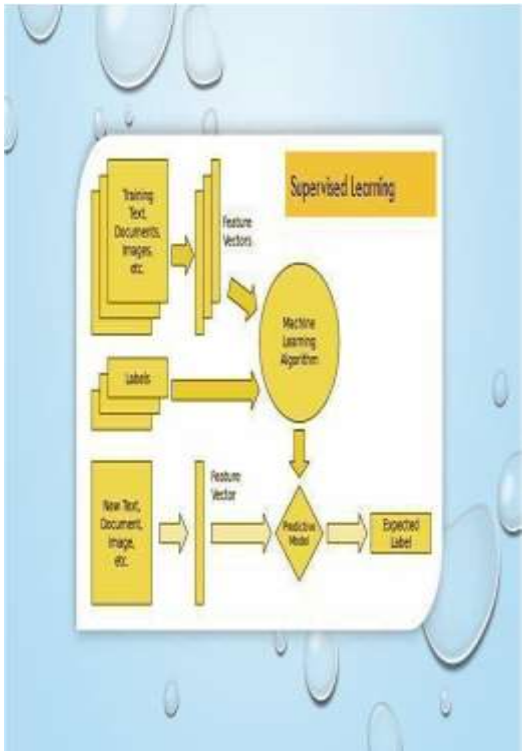
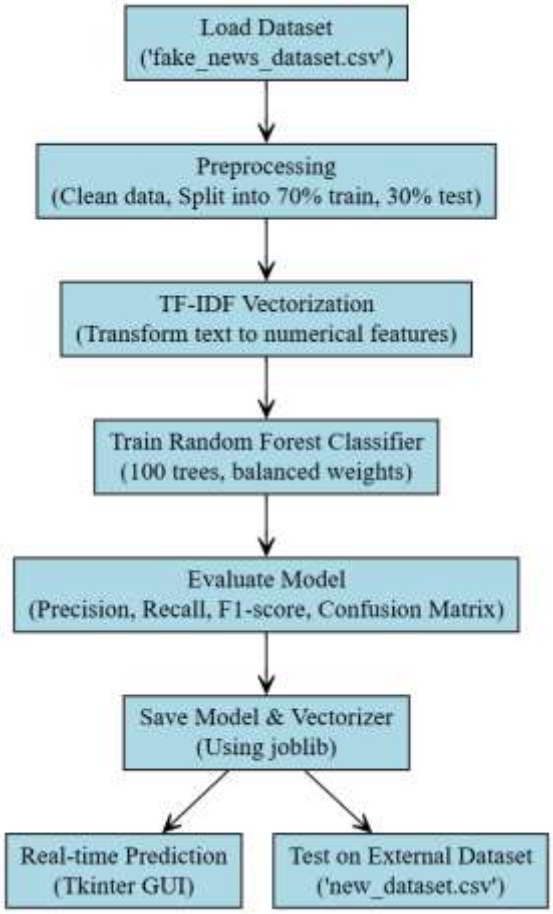


Fig: 2 Algorithm Workflow



RESULT ANALYSIS

The Random Forest Classifier was trained on 'fake_news_dataset.csv', containing 569 samples (326 fake, 230 real) after preprocessing. On the test set (167 samples), the model achieved an accuracy of 75%, with a weighted precision of 0.77, recall of 0.75, and F1-score of 0.73. The confusion matrix revealed 33 true negatives, 92 true positives, 35 false positives, and 7 false negatives, indicating strong performance on fake news detection but room for improvement in reducing false positives. Testing on 'new_dataset.csv' (5 samples) validated the model's generalizability, correctly classifying statements like "The moon landing was faked by NASA" as fake (1) and "New technology is transforming healthcare" as real (0). Visualizations, including bar charts comparing precision/recall and confusion matrices, were generated using Matplotlib to provide insights into model performance. Comparative analysis with baseline models (e.g., Naive Bayes) showed the Random Forest Classifier's superior accuracy and scalability for text-heavy tasks.

Fig: 3 Accuracy Comparison Graph

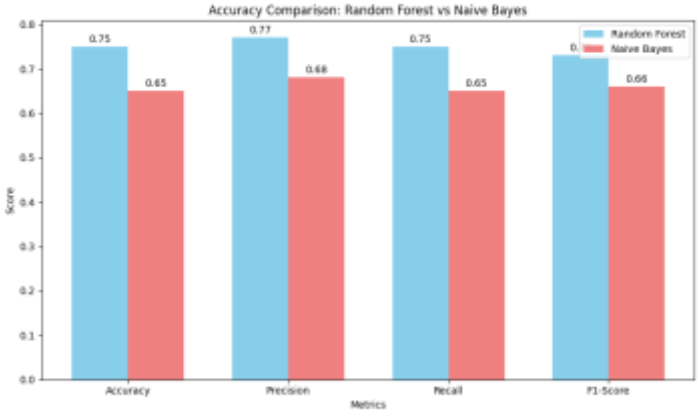


Fig: 4 Confusion Matrix for Naive Bayes

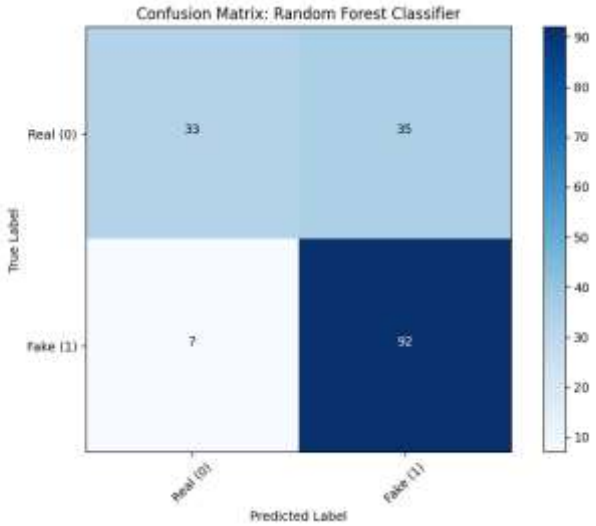
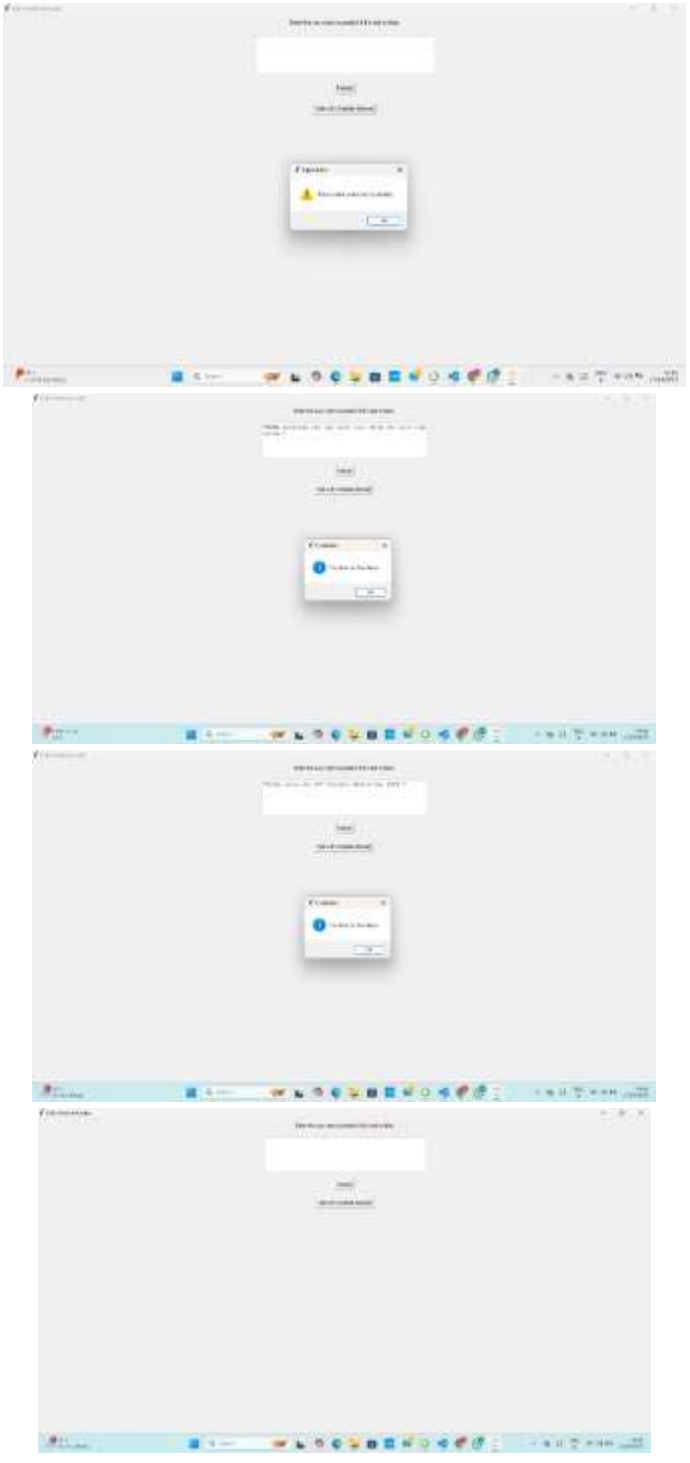


Fig: 5 Sample Prediction Output



CONCLUSION

This project delivers a robust and accessible fake news detection system powered by a Random Forest Classifier and TF-IDF vectorization. Achieving a test accuracy of 75% and demonstrating generalizability on external datasets, the system effectively distinguishes between real and fake news. The Tkinter GUI enhances usability, enabling real-time predictions for diverse users, from casual readers to media professionals. By addressing the limitations of manual verification and less accurate automated methods, this solution contributes to restoring trust in digital information. Future enhancements could include integrating deep learning models, such as BERT, for improved contextual understanding, incorporating multimedia analysis (e.g., images and videos), and leveraging real-time social media streams to detect emerging misinformation trends. These advancements would further strengthen the system's ability to combat the evolving landscape of fake news.

REFERENCES

1. Conroy, N. J., Rubin, V. L., & Chen, Y. (2015). "Automatic deception detection: Methods for finding fake news." ASIS&T Annual Meeting.
2. Helmstetter, S., & Paulheim, H. (2018). "Weakly supervised learning for fake news detection on Twitter." ASONAM.
3. Della Vedova, M. L., et al. (2018). "Automatic Online Fake News Detection Combining Content and Social Signals." FRUCT.
4. Tacchini, E., et al. (2017). "Some like it hoax: Automated fake news detection in social networks." arXiv:1704.07506.
5. Shao, C., et al. (2017). "The spread of fake news by social bots." arXiv:1707.07592.
6. Chen, Y., et al. (2015). "Misleading online content: Recognizing clickbait as false news." ACM Workshop on Multimodal Deception Detection.
7. Najafabadi, M. M., et al. (2015). "Deep learning applications and challenges in big data analytics." Journal of Big Data.