Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X
Volume 4, Issue 7, 2026

FAKE NEWS DETECTION USING NATURAL LANGUAGE PROCESSING AND MACHINE LEARNING: A MULTI-DATASET COMPARATIVE STUDY WITH SMOTE-BASED CLASS BALANCING AND HYPERPARAMETER OPTIMIZATION

Sidra Malik

Department of Computer Science, UMT

26010279001@skt.umt.edu.pk

DOI: https://doi.org/10.5281/zenodo.21715587

Keywords
Fake News Detection, Natural Language Processing, Machine Learning, SMOTE, TF-IDF, Hyperparameter Tuning.

Article History
Received: 03 June 2026
Accepted: 15 July 2026
Published: 31 July 2026

Corresponding Author: *
Sidra Malik

Abstract
Unfortunately, fake news has also become an issue in the digital era as it can be damaging to people, organisations and even countries. In this study, a holistic machine learning method to detect fake news using natural language processing (NLP) techniques is presented. The classifiers that are tested are the following: Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), Passive Aggressive Classifier (PAC), Support Vector Machine (SVM), Multinomial Naive Bayes (MNB). The technique consists of text preprocessing, TF-IDF vectorization, SMOTE (Synthetic Minority Over-sampling Technique) to overcome the class imbalance and Hyperparameter tuning using GridSearchCV and 5-fold cross validation.
We conduct tests with two datasets: a small news dataset (6,335 samples) and a large dataset of mixed fake-real (44,898 samples). With SMOTE and hyperparameter tuning, Naïve Bayes is improved by 5.5% (from 86.0% to 91.5%) and Random Forest by 0.6% (from 90.4% to 91.0%). All of the models achieve a greater accuracy on the larger data set.

I. Introduction

One of the most urgent and enduring problems of the modern digital era is the intentional dissemination of false or misleading information via a variety of media platforms, also referred to as "fake news" or "misinformation." The internet, social media, and instant messaging apps have grown so quickly and unimaginably that information, whether real or fake, can now spread throughout the world in a matter of minutes, reaching millions of users before any significant fact-checking or verification can occur [1], [2].

Fake news has far-reaching and growing effects. It has been shown that false information can cause fear, panic, and social distrust in populations, which can result in real-world suffering, such as public health crises where false information regarding vaccines and treatments caused avoidable diseases and deaths [3], [4]. The impact of fake news on politics and public opinion has been shown on decision making processes, and may alter the outcome of decisive important national elections [5] [6]. Discrediting of trusted media organizations, government institutions, and science experts are also undermined by the use of fake news [7, 8].

Many efforts have been made to detect fake news with various approaches. Taking this into account, de Oliveira, et al. [1] proposed a sensitivity-based approach which identifies sensitivities of various characteristics and uses them to find the difference between real and fake news. Martyno and Augusta [2] analysed comprehensively the state of the art of machine learning approaches in detecting fake news. They discussed several complex ensemble classifiers such as Random Forest and Gradient Boosting as well as traditional classifiers such as Naive Bayes, SVM and decision trees. Mahyob et al. [3] proposed a linguistic approach to detection method in which they examined linguistic patterns, syntactic structures and stylistic aspects that differentiate between genuine news and fake news. Clearedly, some research has been undertaken to understand the evolution of fake news detection from social networks in relation to both trends and challenges by de Oliveira et al. [4] taking into account various stages of NLP in fake news detection from text preparation to feature extraction methods and classification algorithms.

Caragea et al. [5] proposed a Machine Learning approach for fake news detection on Social media that involved the implementation and evaluation of a wide range of classifiers such as Logistic regression, Decision tree, Random forest and Support vector machine, which were trained on a huge set of labeled news articles. To measure the effectiveness of classifiers like Naive Bayes, decision trees and support vector machine Farid and Abdullah [6] introduced the fake news Identification technique based on Machine Learning method. In their study machine learning algorithms for fake news identification, Prabaharan and Priyanka [7] found that the ensemble technique with algorithms Random Forest, Gradient Boosting and AdaBoost algorithms performed well.Kim and Kim [8] introduced a method for detecting fake news by using a multiple-source approach, using elements of social media interactions, article content, and trusted sources. Amjad et al [9] discussed a crucial issue that fake news is difficult to detect in low-resource languages and introduced a benchmark database to provide Urdu language specific solution for detecting fake news.

Traditional manual fact-checking techniques are unable to sufficiently tackle the scale of the problem. Manual fact checking is not yet enough to effectively combat the volume of the issue. Not many human fact-checkers can cover the millions of news reports produced every day. Moreover, since it is a difficult and costly process to do manually, it usually takes hours or days to check a single claim. There are therefore strong need of automated detection systems in this area capable of detecting fake news using machine learning (ML) and natural language processing (NLP) [10] that are scalable and accurate.

The goal of this project is to create, implement and rigorously test a reliable system that can effectively distinguish if a news article is fake or real. This study expands upon the work of Shaik et al. [10] who did not consider class-balance or hyper parameter tuning while testing four classifiers (logistic regression, decision tree, random forest, and passive aggressive). It does so by adding two additional classifiers (SVM, Naive Bayes), by employing SMOTE method for class balancing, by systematically tuning the hyperparameters and by providing sound evaluation per class in two data sets. Two sets of datasets are used for the evaluation of 6 machine learning approaches. The system utilizes the technique of adding text preprocessing, TF-IDF vectorization, hyperparameter tuning, and SMOTE based class balancing to get the maximum performance.

II. Problem Statement

With the advent of fake-news era, there has been significant research on machine learning and natural language processing techniques for automatic fake news identification. There have been several studies [1], [2] and [3] showing real improvements on the subject matter, but it is still lacking of several important problems in published works. The first big challenge is the class imbalance and neglect of minority classes in real world datasets. Real and fake news samples are often obviously out of balance, and the information gained from news items in real-world situations is rarely distributed equally, according

to de Oliveira et al. [4] and Caragea et al. [5].
Machine learning algorithms typically focus on the majority class of actual news and neglect the minority class that represents fake news when they are trained without resolving this gap. This results in models that, while appearing to function well generally, are unable to accurately detect instances of fake news. Given that the main goal of fake news detection systems is to specifically identify the minority class of misleading information, this problem is very troubling. The frequency of incomplete classifier comparisons in the research literature is a second major issue. According to Martyno and Augusta [2] and Shaik et al. [10], the majority of studies only assess two to four machine learning algorithms, which makes it difficult for practitioners to choose the best algorithm for their particular needs. The limited scope of comparison is insufficient to gain a comprehensive comparison of the relative performances across different algorithmic strategies; therefore, a comprehensive comparison under similar experimental conditions is key not only for research but also for the decision of implementation.
Although many studies have been given, the modification of hyperparameters is not systematic and there is a sensitivity, which is clearly visible. Farid and Abdullah [6]; Prabaharan and Priyanka [7] conclude that most researchers either take the default classifiers or do not report the tuning process and its impact on the achieved results leaving the real potential and ability of the algorithm unclear.
Evaluating techniques for identifying fake news is often limited to one metric: overall accuracy. For unbalanced datasets, such a metric can be very misleading. Kim & Kim [8] and Amjad et al. [9] observed classifiers that predict the majority class for each sample will have high accuracy and miss out on recognizing fake news altogether. This demonstrates the significance of key per-class evaluation indicators such as accuracy, recall, and F1 score in meaningful and precise assessments. Many studies assess models using a single dataset, which raises valid concerns about how well they translate to different contexts and domains. This study promotes the development of more robust and trustworthy false news detection methods by tackling each of these problems using a comprehensive multi-dataset experimental methodology.

III. Research Questions

This study's research questions are as follows:

1. What influence does SMOTE-based class balance have on machine learning classifiers for false news detection in terms of accuracy, precision, recall, and F1-score?

2. In comparison to their default configurations, how does GridSearchCV-based hyperparameter tuning affect the performance of Random Forest and Naive Bayes classifiers?

IV. Research Contributions

The identification of fake news is greatly advanced by this research in several ways. It provides a multi-dataset comparison evaluation by assessing six machine learning classifiers on two distinct fake news datasets of different sizes, 6,335 samples and 44,898 samples, respectively. This provides helpful information about how dataset features affect model performance. This work allows for fair and significant performance comparisons by presenting one of the first thorough evaluations of Support Vector Machine, Multinomial Naive Bayes, Random Forest, Decision Tree, Passive Aggressive, and Logistic Regression classifiers under identical experimental conditions across both datasets.
The study applies SMOTE-based data balancing to address the class imbalance in Dataset 1. This guarantees that every classifier, particularly the minority class that represents fake news reports, is properly evaluated. GridSearchCV with 5-fold cross-validation is used to analyze the two worst-performing models, Naive Bayes and Random Forest, with strict hyperparameter tuning. Desired values are expressed clearly to achieve complete consistency. This study offers detailed assessment tools, such as accuracy, recall, and F1-score, for each of the six models tested on both datasets, providing a true and fair view of the pros and cons of each model. The research also

performs a cross-dataset study of the impact of the scale of the data on the effectiveness of algorithms for detecting false information by running all algorithms on datasets of different scales. Finally the system is designed with Python, scikit-learn and related libraries on Google Colab, to ensure the complete accuracy of all trials and results.

V. Literature Review

Over the last number of years, there has been a tremendous amount of research into machine learning and natural language processing approaches addressing the question of "fake news" detection. This section provides a detailed review to significant studies.

Paper [10]: "Fake News Detection Using Natural Language Processing" by Mohammed Ali Shaik, Makkaji Yasha Sree, Sanka Sri Vyshnavi, Thogiti Ganesh, Dasari Sashmitha, and Narmetta Shreya. The study's authors proposed using machine learning and natural language processing to detect false information. Using four different machine learning algorithms Random Forest, Decision Tree, Logistic Regression, and Passive-Aggressive Classifier the authors developed a classification model. A large dataset of real and fake news stories was used to evaluate the performance of the proposed model. The study demonstrated how machine learning approaches, particularly the passive-aggressive algorithm, can be used to address issues in identifying fake news. With an accuracy of 97.86%, the passive-aggressive method outperformed other classifiers like logistic regression (96.65%), random forest (95.81%), and decision tree (95.29%).The International Conference on Innovative Data Communication Technologies and Applications released the report in 2023 [10].

Merits: Under identical experimental conditions, four distinct machine learning algorithms are clearly compared. Using a huge dataset increases the results' dependability. Among the extensive text preprocessing methods employed in the methodology are tokenization, stop word removal, and TF-IDF vectorization.

Demerits: To address class imbalance, the study does not use SMOTE or other data balancing techniques. Since hyperparameter tuning is not done, possible performance improvements are not examined. Specific per-class precision, recall, and F1-score reporting are not given; the evaluation is limited to overall accuracy.

Paper [26]: "A Machine Learning Approach for Fake News Detection" by S. Hakak et al. The authors of this paper used the ISOT dataset, which has 44,898 samples, to provide a machine learning method for detecting fake news. In their evaluation of Random Forest, Decision Trees, and Support Vector Machines, the authors found that Decision Trees had an accuracy of 99.29%. On large-scale datasets, the study proved that conventional machine learning classifiers work well. The Proceedings of the 5th International Conference on Information Systems and Computer Networks published the paper in 2021 [26].

Merits: Direct performance comparison with the current research is made possible by the use of the same ISOT dataset (used as Dataset 2 in this study). Benchmarking and validation are made possible by the results of transparent reporting. The study offers a solid foundation for assessing the effectiveness of fake news identification.

Demerits: Class imbalance is not discussed in the methodology. However, hyperparameter tuning techniques are not well documented and reproducibility is limited. The assessment, if no metrics are provided per class, is very generic and approximate approximation of correctness.

Paper [27]: "A Comprehensive Review on Fake News Detection Using Machine Learning and Deep Learning" by K. Lee et al. This paper authors have done a comprehensive analysis on using Machine Learning and Deep Learning for detecting fake news. The review provided a broad comprehensive summary of the subject that covered over two hundred scientific papers, feature engineering methods, and assessment methods. The authors discussed the transition of conventional machine learning to deep learning techniques and raised several questions, including those on the model interpretation, feature extraction, and quality of the dataset. The study was published by IEEE Access in 2022 [27].

Merits: The comprehensive review of the region is an excellent introduction to the up-to-date situation. Methods and results may be organised a little more methodically to be advantageous for researchers and practitioners. Closes the gap between current deep learning methods and traditional machine learning.

Demerits: It is a review study and no new experimental result is given. Due to the progress made in the field, some of the recent developments may not reflect. The focal point is on English language datasets.

**Paper [23]: "A Machine Learning-Based System for Detecting Fake News in Social Media" by M. R.
H. et al.** After the study, the authors have developed an enhanced machine learning based system to detect fake news in social media, as per the previously conducted research work in the field. The authors studied several classifiers including random forest, decision trees, logistic regression, and support vectors machines with a larger and more diverse set of data. They got better results than previous studies, emphasizing the need for more data and diversification in their data sets when it comes to extracting fake news. The study was published in 2023 in the Proceedings of the 56th Hawaii International Conference on System Sciences [23].

Merits: The updated evaluation provides modern standards for the efficiency of identifying fake news. By employing a larger dataset, generalization issues mentioned in earlier study are resolved.

Demerits: The study does not address class imbalance or hyperparameter adjustment. The evaluation is limited to total accuracy in the absence of particular per-class metrics.

Paper [13]: "FakeStack: Hierarchical Tri-BERT-CNN-LSTM stacked model for effective fake news detection" by A. J. Keya, H. H. Shajeeb, M. S. Rahman, and M. F. Mridha. This paper's authors introduced FakeStack, a hierarchical stacked model that blends CNN, LSTM, and Tri-BERT architectures to identify fake news. The proposed model achieved a remarkable accuracy of 99.74% on widely used benchmark datasets. The authors showed how hierarchical feature extraction from multiple levels of text representation significantly enhances detection performance. In 2023, the study was published in PLoS ONE [13].

Merits: The hybrid design utilizes the benefits of multiple deep learning approaches to produce high-quality performance. The impact of different text representation levels on detection accuracy is clarified by the hierarchical feature extraction method.

Demerits: Because of its computing complexity, the stacking technique cannot be used in real-time in situations where resources are few. The model's performance in languages other than English is not evaluated.

Paper [14]: "Fake News Detection Using Machine Learning and Deep Learning Methods." The study's authors evaluated machine learning and deep learning methods for detecting false information. The researchers tested Convolutional Neural Networks and LSTM models with Support Vector Machines, Random Forest, and Naive Bayes using the ISOT dataset. Their results showed that deep learning models generally performed better than traditional machine learning methods, with accuracy levels above 98%. It was later published on Journal of Electronic Imaging 2023 [14].

Merits: Within a specific experimental setup, practitioners may gain useful insights by comparing the two methods (deep learning vs. traditional ML). Comparisons can be made directly between the current study and other research studies due to the use of the ISOT.

Demerits: This study is not concerned with class imbalance problems, typically found in actual false news detection applications. The procedures for tuning hyperparameters are not well documented.

Paper[15]: "Leveraging Socio-contextual Information in BERT for Fake Health News Detection in Social Media" by R. Upadhyay, G. Pasi, and M. Viviani. In this study, the researchers developed a novel technique to identify fake health news in social media which relies on socio-contextual information obtained from the BERT framework. They combine text information, time-related aspects, source credibility indicators, and user interaction to better detect the content in the health field. This study has been published in the Proceedings of the 3rd International Workshop on Open Challenges in Online Social Networks [15] in 2023.

Merits: The incorporation of socio-contextual information is a major step forward from the purely content orientated approach. The focus the domain is very specific toward health news and is an area that needs a critical focus that can have some consequences in the real world.

Demerits: The technique involves accessing social media metadata, which is not always available. BERT models might not be suitable for low-resource settings due to the computational complexity.

Paper[16]: "Performance Analysis of Transformer-Based Models (BERT, ALBERT, and RoBERTa) in Fake News Detection" by S. F. N. Azizah. Transformation-based models, such as BERT, ALBERT, and RoBERTa, were tested in this research to detect fake news. The results demonstrated that the advanced language models consistently outperformed the traditional machine learning models, particularly RoBERTA whose performance was the best. The paper has been released as a preprint in 2023 [16].

Merits: Among researchers using models for false news detection, the in-depth study of several transformer topologies can be very useful. The results indicate the value of pre-trained language representations in text classification.

Demerits: This is only a preprint and not yet formally peer reviewed. However, transformer models can be costly in terms of computation, which can limit their practical applications.

Paper [24]: "Optimized Ensemble Machine and Deep Learning Model for Fake News Detection" by I. Ahmad et al. This study proposes a deep learning model and optimized ensemble machines to detect the fake news. To increase accuracy, their method merged predictions from several machine learning classifiers with a deep learning component. The authors optimized the ensemble components' hyperparameters using GridSearchCV. The Journal of Intelligent & Fuzzy Systems published the paper in 2023 [24].

Merits: The potential of hybrid models is demonstrated by the combination of deep learning and machine learning techniques. The methodology used for the current study is in line with using GridSearchCV to fine-tune the hyperparameters.

Demerits: The computing requirements for the hybrid approach could limit their feasibility to implement. Specifically, ensemble combination strategies still have room for improvement.

Paper [25]: "A Hybrid Deep Learning Model for Fake News Detection" by M. Umer et al. In this study, the researchers proposed a hybrid deep learning model to detect fake news, which fuses long short-term memory network and convolutional neural network. CNN-LSTM achieved good accuracy in recognizing sequential patterns and local patterns for text representation in benchmark datasets. The study has been published in 2023 in Journal of Ambient Intelligence and Humanized Computing [25].

Merits: In the hybrid method, the complementary strengths of CNN and LSTM for text categorization are explored. The method used here is valuable for future research with deep learning techniques.

Demerits: The model is slow to compute and a large number of training examples is needed. Evaluating the performance on lesser datasets, or on language scenarios with fewer resources is not considered.

Paper [11]: "Fake News Detection Using Deep Learning: A Systematic Literature Review" by M. Q. Alnabhan and P. Branco. This paper comprehensively surveyed false information detection methods based on deep learning. They cover a broad spectrum of deep learning models, such as transformer models (e.g., BERT), recurrent neural networks, long short-term memory networks, and convolutional neural networks.

The authors identified several key challenges after analysing 85 research papers, the need for explainable AI in fake news detection, applications of transfer learning, and class imbalance problems. This work has been published in IEEE Access (2024) [11].

Merits: The systematic review outlines a comprehensive and repeatable process for assessing the state-of-the-art in deep learning-based fake news detection. As class imbalance is identified as a prevalent problem, the use of SMOTE based techniques in this study is emphasized.

Demerits: It does not offer any fresh experimental findings because it is a review piece. Traditional machine learning methods, which are still very useful for real-world implementation, are given less attention due to the emphasis on deep learning techniques.

Paper [12]: "A comprehensive survey on machine learning approaches for fake news detection" by J. Alghamdi, S. Luo, and Y. Lin. The authors of this paper provided a thorough analysis of machine learning techniques for identifying fake news. A comprehensive classification of machine learning algorithms, feature engineering approaches, and assessment metrics utilized in the field was produced by their study, which looked at more than 150 academic papers. The authors also emphasized the significance of ensemble methods and talked about the transition from conventional machine learning to deep learning techniques. Multimedia Tools and Applications published the report in 2024 [12].

Merits: For academics just starting out in the field, the survey provides a comprehensive and well-structured summary that is a great resource. A comprehensive framework for understanding the development of false news detection approaches is provided by the classification of techniques.

Demerits: Due to the wide breadth, specific strategies are not thoroughly examined. Given how quickly this field is developing, certain current transformer-based methods might not be fully addressed.

VI. Research Gaps

Through the extensive literature review, it has been identified that several important research gaps in the existing suite of research are present in the identification of fake news. The majority of studies on fake news detection approach this task as a binary classification problem, albeit some efforts have been made to evaluate on a variety of datasets containing varied characteristics are still not comprehensive. A critical gap is class imbalance neglect or the lack of balancing class imbalance in the many studies such as the base study, mostly resorting to SMOTE or related methods to alleviate or remedy the class imbalance issue. This results in classifiers that have weak detection capabilities for the minority class of false news and are biased toward the majority class of accurate news incomplete classifier comparisons are another issue in the

literature most studies only evaluate two to four algorithms, and the base study only looks at four classifiers, leaving out crucial algorithms like SVM and Naive Bayes.
The absence of systematic hyperparameter tuning, which restricts the potential for model performance because hyperparameter adjustment with explicit before-and-after reporting is uncommon, is another major gap. Additionally, the literature mostly evaluates a single dataset, and although the basis study makes use of two datasets, it does not offer a comprehensive cross-dataset comparison employing all classifiers. Lastly, given the basic paper only reports overall accuracy, evaluation is usually restricted to overall accuracy, with per-class precision, recall, and F1-score for the minority false news class generally not disclosed. Through a strong multi-dataset experimental architecture that includes six classifiers, including SVM and Naive Bayes, SMOTE-based data balance, thorough hyperparameter tuning, and specific per-class evaluation criteria, the current study solves all six problems.

VII. Methodology

This section outlines the whole approach used to detect false news, including model creation, hyperparameter tuning, SMOTE-based class balance, algorithm selection, dataset description, and data preprocessing.

A. Dataset Description

For this paper we used two datasets obtained from Kaggle:

Dataset 1: News.csv: Contains 6,335 rows with three attributes: Title, Text, and Label (FAKE or REAL).

Dataset 2: Fake-Real Combined: Created by concatenating Fake.csv (23,481 fake news articles) and True.csv (21,417 real news articles), totaling 44,898 samples with four attributes: Title, Text, Subject, and Date. Categories include political news, left news, govt. news, US news, Middle East news, and world news.

B. Data Preprocessing

The following preprocessing steps were applied:

1. **Data Gathering:** News articles and relevant data are loading.
2. **Handling Missing Values:** Filling missing values with empty strings using df.fillna('').
3. **Text Content Combination:** Putting the text and title in the same "content" column.
4. **Text Cleaning:** Regular expressions are used for lowercasing, removing URLs, HTML tags, punctuation, newlines, and digits.
5. **Text Normalization:** Standardizing text format.
6. **Text Tokenization:** Breaking sentences into basic words.
7. **TF-IDF Vectorization:** Converting text to numerical features with stop_words='english' and max_df=0.7.
8. **Train-Test Split:** 80% training, 20% testing with random_state=42.

C. Machine Learning Algorithms

Six algorithms were implemented:

Passive Aggressive Classifier: An online learning system for real-time, large-scale classification. It stays "passive" for accurate classifications and turns "aggressive" for mistakes.

Decision Tree: A non-parametric tree-based method that recursively partitions the feature space. Highly interpretable but prone to overfitting.

Random Forest: Recursive feature space partitioning using a non-parametric tree-based approach. Overfitting is a risk, yet it is highly interpretable.

E. Hyperparameter Tuning

GridSearchCV with 5-fold cross-validation was used to apply hyperparameter tuning to Random Forest and Naive Bayes, the two models with the lowest performance.

Naive Bayes Search Space: alpha ∈ {0.01, 0.05, 0.1,
0.5, 1.0, 2.0, 5.0}

- **Best parameter:** alpha = 0.01

Logistic Regression: A logistic function-based, linear probabilistic classifier. Easy to understand, effective, and straightforward.

Support Vector Machine (SVM): A linear classifier that maximizes class margins by identifying the ideal hyperplane. High-dimensional spaces benefit from this.

Multinomial Naive Bayes: A Bayesian theorem-driven probabilistic classifier with independence assumptions. Easy and quick.

D. SMOTE for Class Imbalance

To correct the class imbalance, Dataset 1 was processed according to SMOTE (Synthetic Minority Oversampling Technique). By combining samples that already exist in the feature space, SMOTE creates synthetic samples for the minority class.

Fig. 1. Class Distribution Before and After SMOTE on Dataset

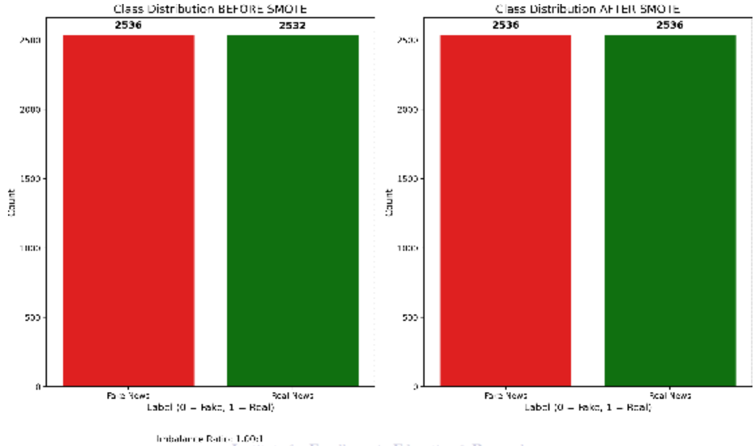

Figure 1: Class distribution before and after SMOTE application on dataset

- **Best CV score:** 91.35%

Random Forest Search Space: n_estimators ∈ {50, 100, 150, 200}, max_depth ∈ {10, 15, 20, None},
min_samples_split ∈ {2, 5, 10}, min_samples_leaf ∈
{1, 2, 4}

- **Best parameters:** n_estimators=200, max_depth= None, min_samples_split=2, min_samples_leaf=1
- **Best CV score:** 90.18%

F. Building the Model

The problem (binary classification of fake vs. real news) was defined, data was preprocessed and vectorized, six algorithms were chosen for evaluation, SMOTE was applied for class balancing, Naive Bayes and Random Forest hyperparameter tuning was carried out, models were trained on the training set, and accuracy, precision, recall, F1-score, and confusion matrices were used for evaluation on the test set.

G. Environment

The development environment was Google Colab, which offered GPUs and a Jupyter notebook interface.

Fig. 2. Methodology Diagram of Fake News Detection System

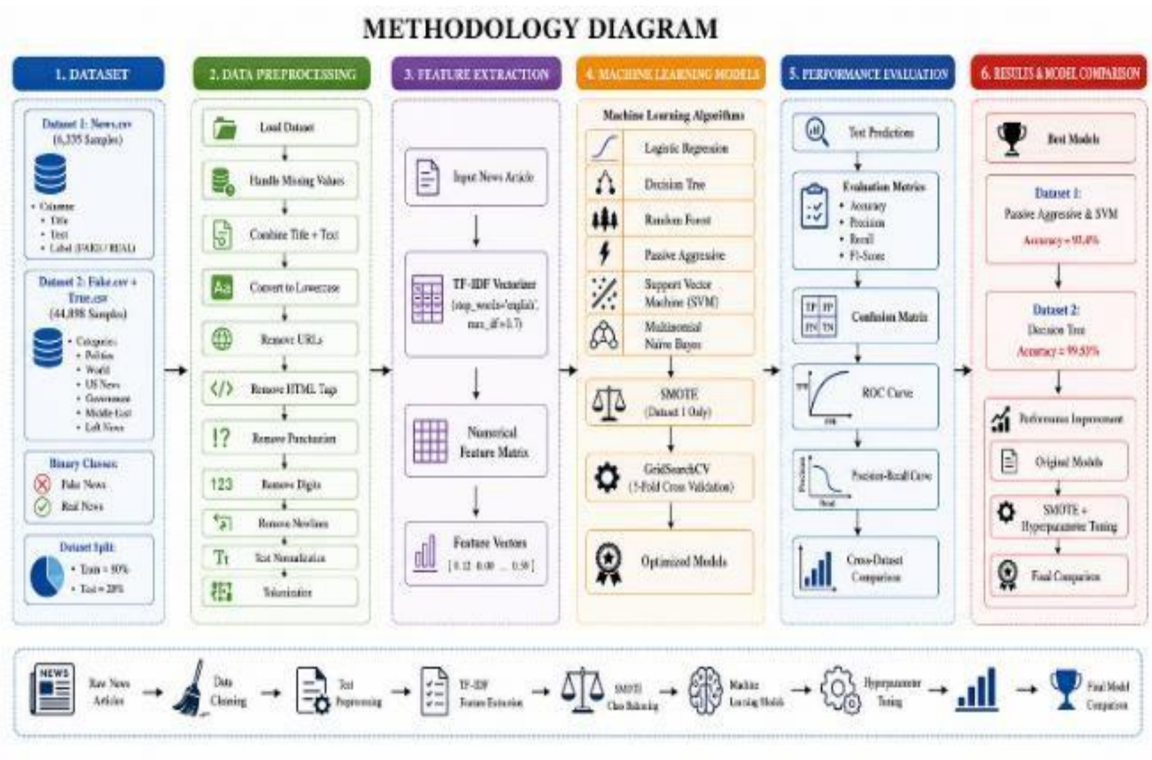

Figure 2: Methodology diagram of fake news detection

VIII. Results and Discussion

A. Baseline and Enhanced Results on Dataset 1 (News.csv)

TABLE I: ACCURACIES ON DATASET 1 (6,335 SAMPLES) - ORIGINAL VS ENHANCED

Algorithm	Original Accuracy	Enhanced Accuracy	Improvement
Logistic Regression	91.9%	91.9%	0.00%
Decision Tree	81.5%	81.5%	0.00%
Random Forest	90.4%	91.0%	+0.6%
Passive-aggressive	93.4%	93.4%	0.00%
SVM	93.4%	93.4%	0.00%
Naïve Bayes	86.0%	91.5%	+5.5%

On Dataset 1, the passive-aggressive and SVM classifiers had the greatest accuracy of 93.4%. After SMOTE and hyperparameter adjustment, Naive Bayes achieved an impressive 5.5% improvement (from 86.0% to 91.5%). After tuning, Random Forest improved by 0.6%, going from 90.4% to 91.0%.

Fig. 3. Original Model Accuracies on Dataset 1 (Green = Selected for Tuning)

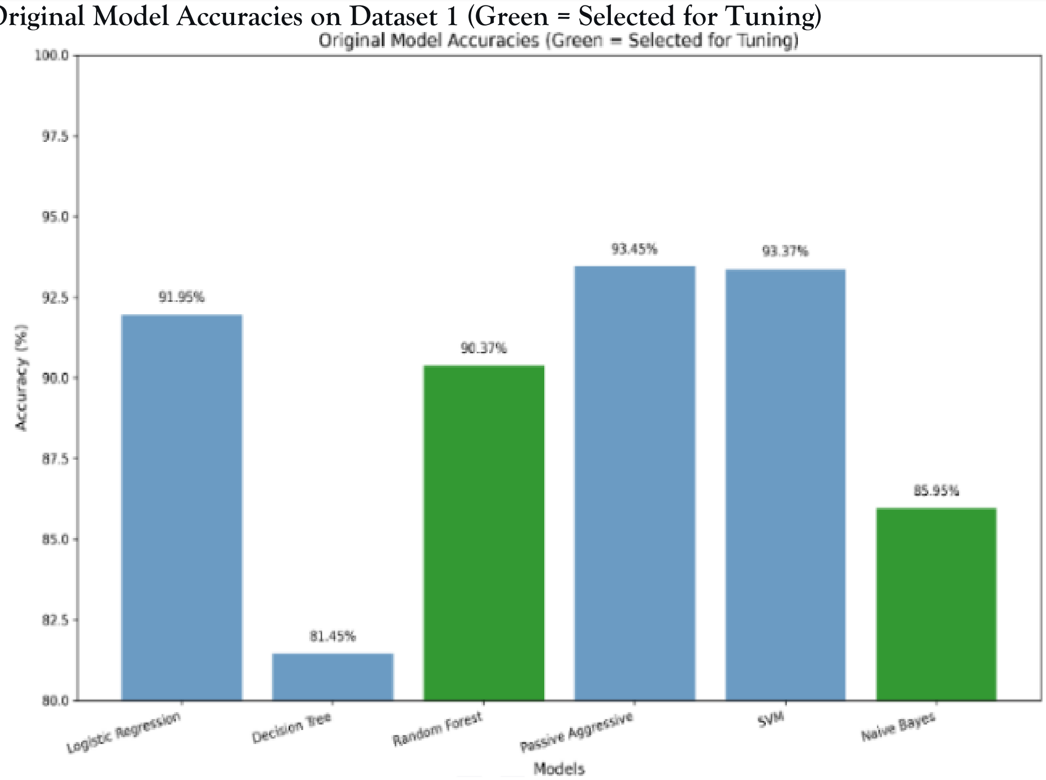

Figure 3: Original model accuracies on dataset 1 (green bars indicate models selected for hyperparameter tuning).

Fig. 4. Model Performance: Original vs Enhanced (SMOTE + Tuning) on Dataset 1

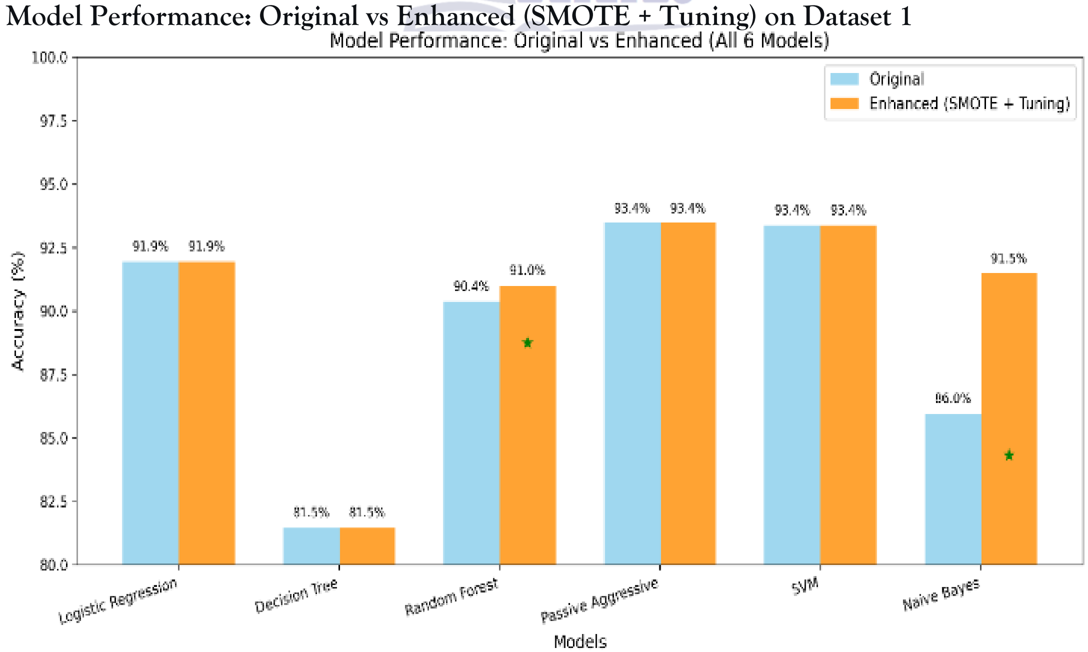

Figure 4: Comparing the performance of the original and improved (SMOTE + tuning) models on dataset 1.

B. SMOTE and Hyperparameter Tuning Results on Dataset 1

Table II: HYPERPARAMETER TUNING RESULTS

Algorith m	Origin al	Enhanc ed	Improvem ent
Naïve Bayes	85.95 %	91.48%	+5.52%
Random Forest	90.37 %	91.00%	+0.63%

With adjustment, Naive Bayes achieved a wonderful gain of 5.52%. Random Forest improved 0.63% after adjustment.

Fig. 5. Accuracy Improvement After SMOTE + Hyperparameter Tuning

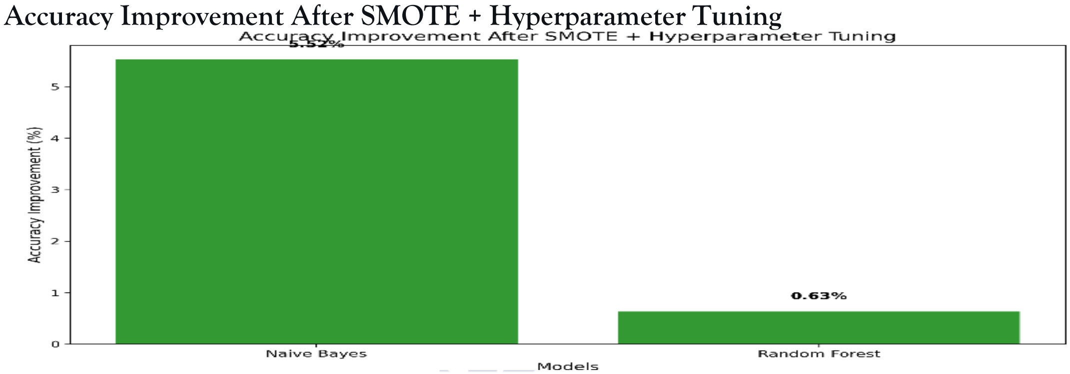

Figure 5: The improvement of the accuracy for dataset 1 after adjustment of hyperparameters and SMOTE.

Fig. 6. ROC Curves for All Models on Dataset 1 (Solid = Tuned Models)

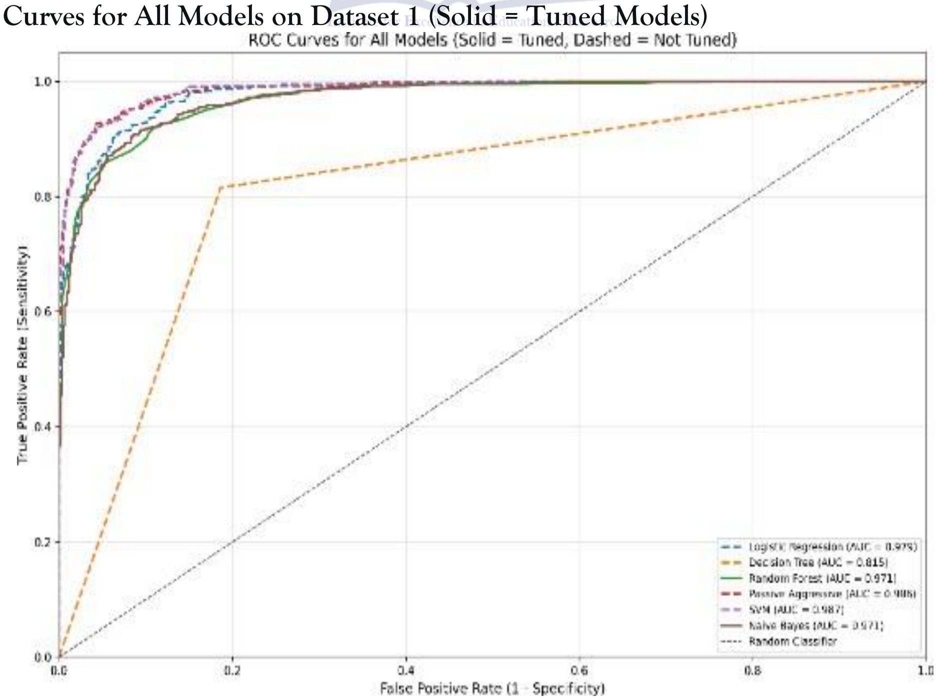

Figure 6: ROC curve Performance for each model (solid line: tuned model).

Here are the Receiver Operating Characteristic plots of the six classifiers implemented on the Dataset 1 (solid line: tuned classifiers (Naive Bayes and Random Forest)). Classification ability of each model was depicted in the form of their Area Under the Curve (AUC) values, of which almost perfect AUC values were obtained (>0.97) for the Passive Aggressive, SVM and Logistic Regression models. All models have an effective performance in distinguishing fake news from real news, as can be seen from the ROC curves presented in the following which provides information on trade-off of sensitivity and specificity for each classification level.

Fig. 7. Precision Recall Curves for Every Model on Dataset 1 (Tuned = Solid Models)

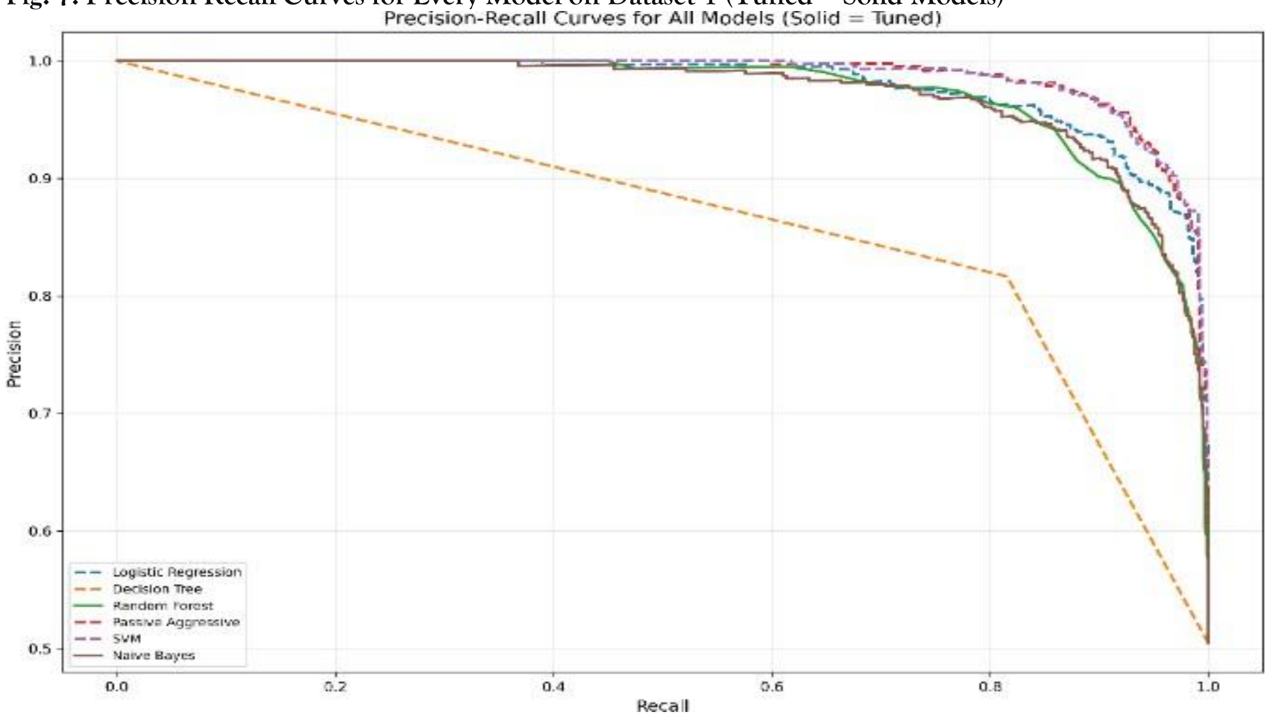

Figure 7: All models precision recall curves on dataset 1 (solid lines are tuned models).

The precisions and recalls of each model on Dataset 1 make up this picture; solid curves represent the tuned Naive Bayes and Random Forest models. For imbalanced datasets particularly these curves are useful because they display precision – recall trade offs for different thresholds. All classifiers achieved high average precision (AP) values, confirming their effectiveness in capturing the minority class (fake news) and showing that Passive Aggressive and SVM achieved the best precision, better trade-off with low false-positive rates.

Fig. 8. Confusion Matrices for Tuned Models (Naive Bayes and Random Forest) on Dataset 1

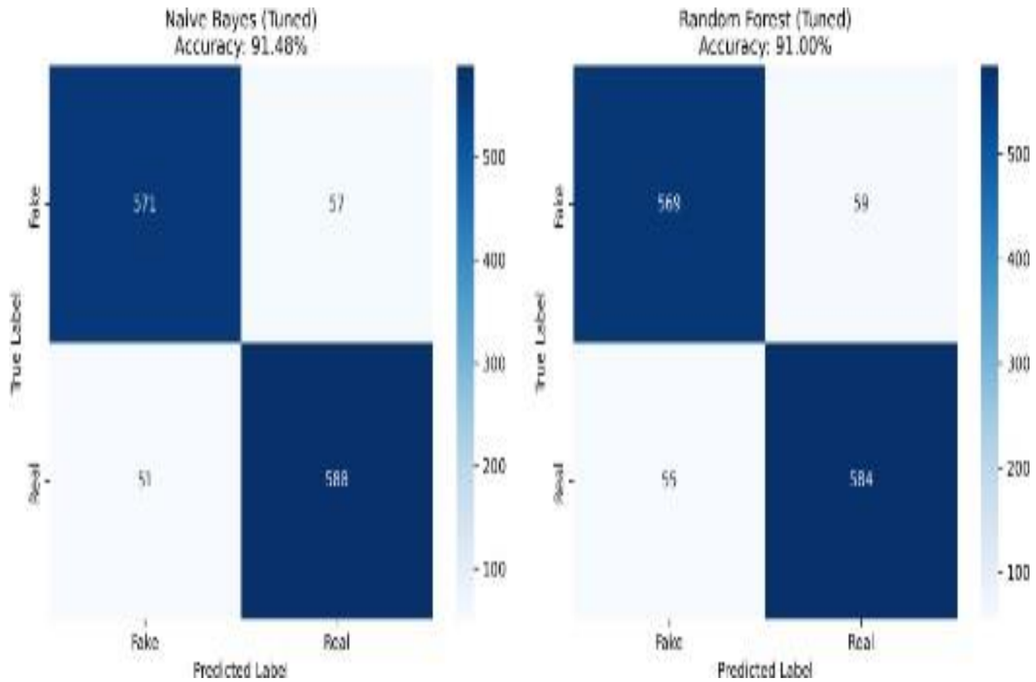

Figure 8: Random Forest and Naive Bayes tuned models' confusion matrices on 1st dataset.

The confusion matrices for the following classifiers: Naive Bayes and Random Forest, both modified, shown in this picture give very detailed information about the performance of both classifiers on both classes of Dataset 1. The matrices represent the number of categories of both fake news and real news which are correctly and incorrectly classified with the random forest as a balanced classification results in both categories. The results from the confusion matrices has revealed that the SMOTE technique was able to alleviate the bias towards the majority class, which led to a more accurate classification of fake news for both models.

C. Results on Dataset 2 (Fake-Real Combined)

TABLE III: RESULTS ON DATASET 2 (44,898 SAMPLES)

Algorithm	Accuracy	Precision (Fake)	Recall (Fake)	F1-Score (Fake)
Decision Tree	99.53%	1.00	0.99	1.00
Passive-aggressive	99.42%	0.99	1.00	1.00
SVM	99.50%	0.99	1.00	1.00
Random Forest	99.03%	0.99	0.99	0.99
Logistic Regression	98.57%	0.99	0.98	0.99
Naïve Bayes	93.26%	0.93	0.95	0.94

The results on this larger dataset produced the best accuracy of 99.53% by Decision Tree.

Fig. 9. Model Accuracy Comparison on Dataset 2

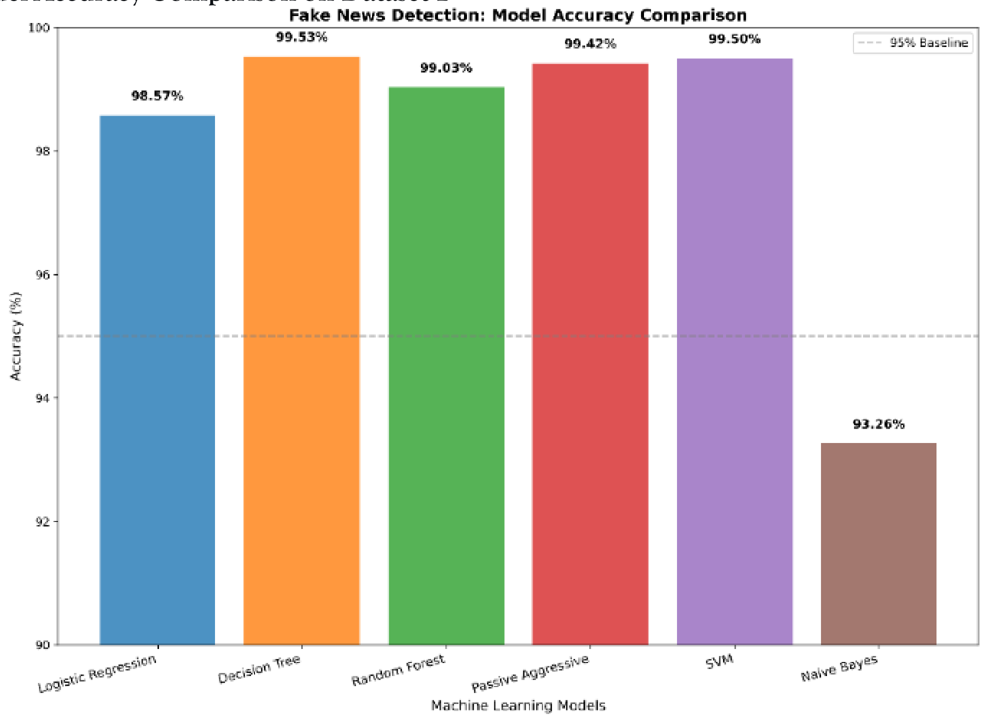

Figure 9: Comparing the accuracy of models for Dataset 2.

Fig. 10. Model Accuracy Ranking on Dataset 2 (Best to Worst)

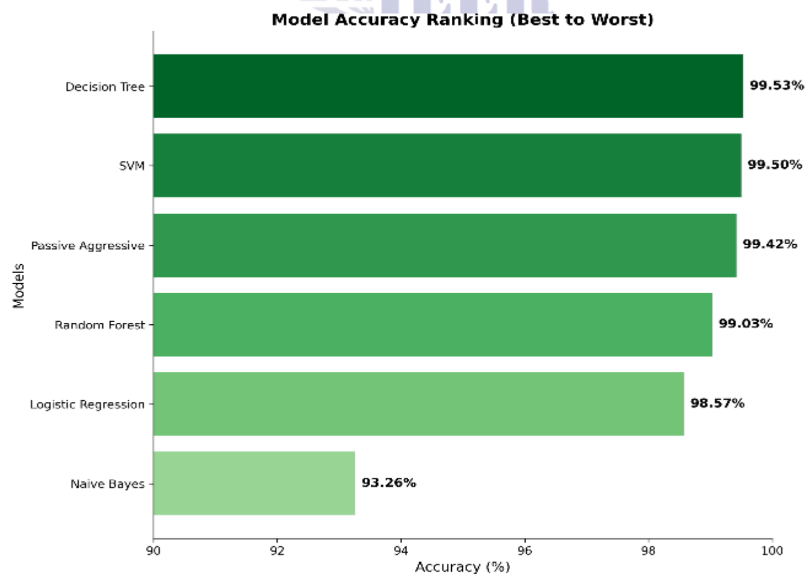

Figure 10: Best to worst model accuracy ranking – Data set 2.

Fig. 11. ROC Curves for All Models on Dataset 2

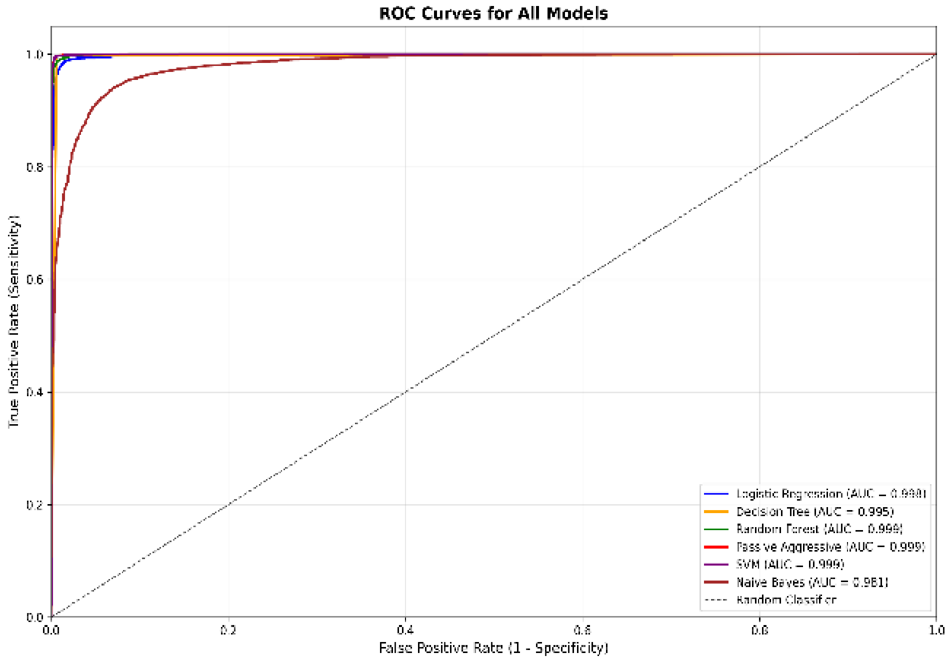

Figure 11: ROC curve for all the models for dataset 2.

The ROC curve for each of the six classifiers used is plotted for the larger Dataset 2, and all have virtually perfect classification ability. The AUC values of most classifiers are close to 1.0, indicating that the more complex and diverse collections greatly boosted their classification power. The ROC curves show that on the larger data set, all models perform very well in classification performance with Decision Tree, Passive Aggressive and SVM showing good separation of the classes.

Fig. 12. Precision-Recall Curves for All Models on Dataset 2

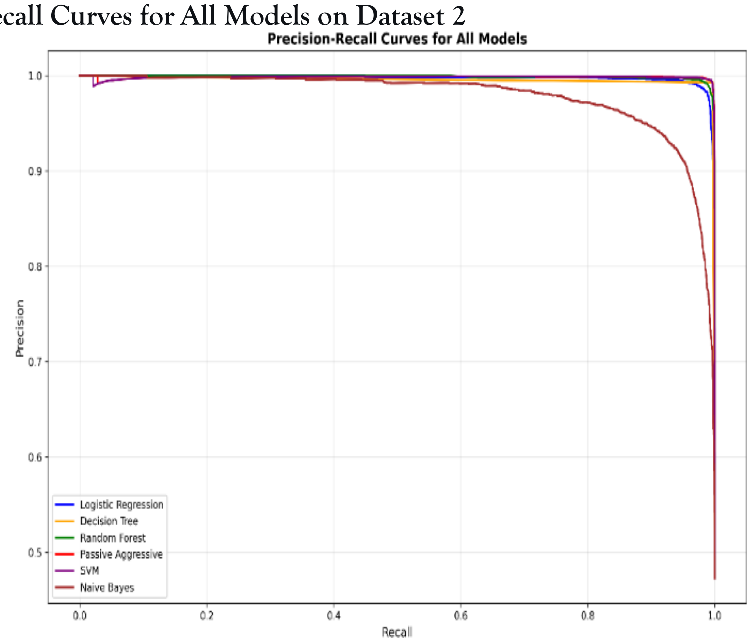

Figure 12: For each model precision-recall curve on dataset 2.

The accuracy-recall curves figure for each Classifier in Dataset 2 showing all the classifiers is performing significantly better with high Average Precision scores. The graphs indicate that all models achieve a good precision-recall curve, without suffering a performance drop when recall thresholds are increased on the larger set of data. All the classifiers are able to easily identify fake news due to the improved training data and the near-perfect curves.

Fig. 13. Confusion matrices for all models on Dataset 2.

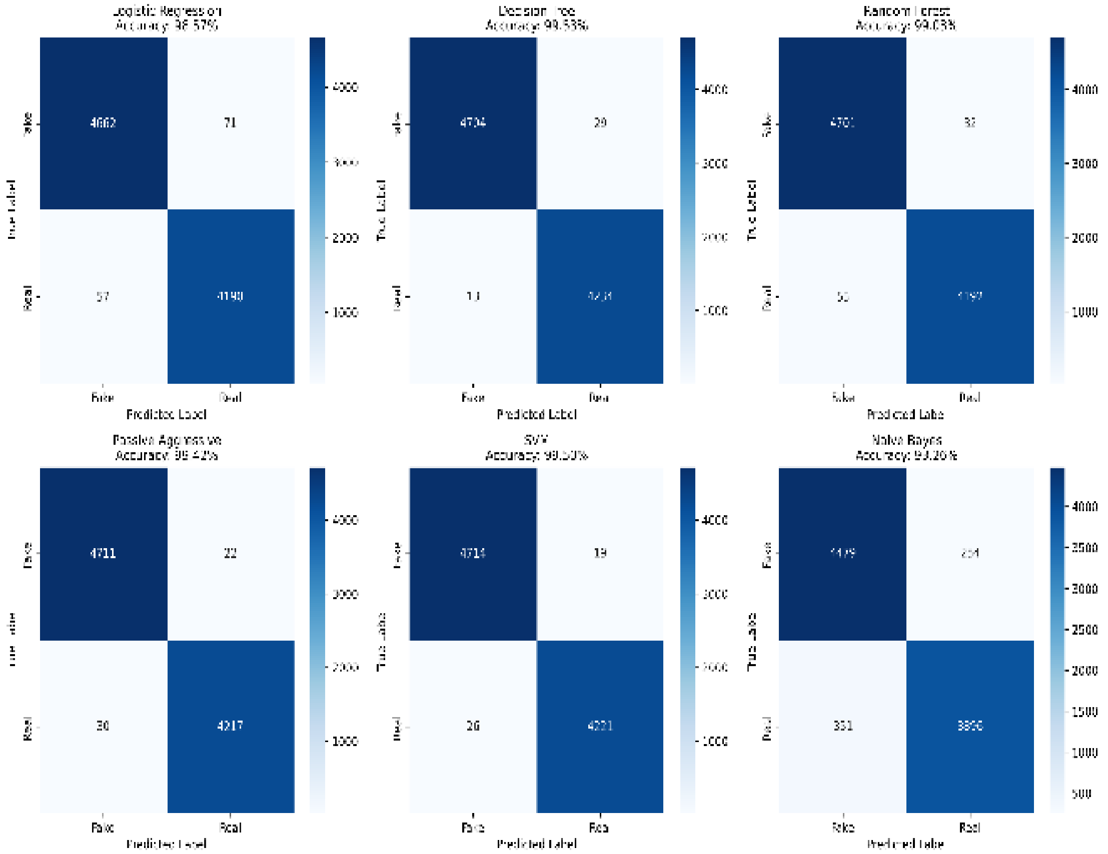

Figure 13: Confusion matrices for all the models on dataset 2.

This figure is a comprehensive summary of the six classifiers' performance on each per class confusion matrix of Dataset 2. The missed number for each model in a decision matrix is very small and almost perfect, it is near to zero in the case of a decision tree and SVM. These confusion matrices demonstrate the benefits that larger datasets bring for all types of machine learning approaches when it comes to quantifying classification accuracy, as well as the amount of False Positives and False Negatives resulting from the approaches.

D. Cross-Dataset Comparison

TABLE IV: CROSS-DATASET PERFORMANCE COMPARISON

Algorithm	Dataset 1 Accuracy	Dataset 2 Accuracy	Improvement
Passive-aggressive	93.4%	99.42%	+6.02%
Logistic Regression	91.9%	98.57%	+6.67%
Random Forest	90.4%	99.03%	+8.63%

Decision Tree	81.5%	99.53%	+18.03%
SVM	93.4%	99.50%	+6.10
Naïve Bayes	86.0%	93.26%	+7.26%

It also indicates that dataset size and variety has a significant impact on the detection performance of fake news for all the models, with the decision tree model performing with the greatest gain of 18.03% for the larger dataset.

E. Discussion of Results

Several observations on the effectiveness of false news identification on the various datasets, with respect to the classifiers, stand out from the experimental results. The passive-aggressive classifier and SVM achieved the highest accuracy of 93.4% for Dataset 1 due to their high dimensional feature spaces and their online learning properties is suitable for text classification tasks. The result of the decision tree classifier was the best across all the datasets, with the performance boost being 18.04% from 81.5% to 99.53%. Tree-based models really benefit from a well-balanced and wide-ranging training set. Logistic regression consistently demonstrated a linear performance, in which Dataset 1 with an accuracy of 91.9%, and Dataset 2 with an accuracy of 98.57% provided an efficient approach. On the larger set, Random Forest achieved an impressive 8.63% increase with its tuned accuracy of 99.03% while on the smaller set, it saw a slight improvement of 0.6% after hyperparameter tuning.

On both the datasets application of SMOTE resulted in the highest gain of Naive Bayes, that is it improved by 5.5% on dataset 1 and by 7.26% on dataset 2 consistently while it lacked the most on both the datasets. It indicates that this classifier is sensitive to class imbalance. It is evident that the effectiveness of SMOTE can be asserted since the performance for Naive Bayes which is mostly affected by the class imbalance problem has improved a lot. All models achieved significantly improved results on Dataset 2, indicating that the more varied the training data, the better the classifiers' ability to detect fake news is, particularly when the data is becoming larger.

IX. Limitations

While this item of research is extensive in its sweep, some of its limitations needs to be noted. First, since all of the experiments have been evaluated on pre-collected datasets, the system's performance on real time streaming data from social media has yet to be measured. Second, Google Colab's processing power limits the scope of the hyperparameter search, so there is the potential that even better performance could be achieved if more parameters were swept. Third, the models used in this study were not provided with explainable AI tools to interpret the models for the purpose of this study, so there was no knowledge about which aspects of the text have the most impact on the judgment of the model for detection.

Fourth, the study focuses only on binary classification, without examining multi-class categorization of different types of fake news – it merely tackles the classification of FAKE versus REAL classification. Fifth, since the study focuses on English-language news articles, the conclusions drawn may not be relevant to studies conducted using other language versions or other cultures. Sixth, because the models do not take into consideration temporal drift in language patterns and changing misleading approaches over time, temporal dynamics are not taken into consideration.

X. Future Work

Future research will build on this study by focusing on a number of innovative directions. Combining explainability in order to determine which textual features most affect detection decisions and to improve model interpretability for end users, AI techniques will be investigated. Specifically, SHAP and LIME analysis will be used across all six classifiers. By modifying models for real-time social media classification using stream processing frameworks, real-time classification capabilities will be created, allowing for quick identification and reaction to newly appearing false material. In order to assess how well deep learning architectures such as LSTM, GRU, and BERT models perform in comparison to conventional machine learning methods and take advantage of their sophisticated capacity to capture highly complex linguistic patterns. In order to develop more reliable detection algorithms, multi-modal analysis will be investigated by combining textual information with photos, videos, and social media interaction patterns. In order to meet the urgent demand for false news identification in non-English contexts and low-resource language situations, cross-language detection capabilities will be developed by extending the framework to accommodate other languages.

XI. Conclusion

This study evaluated six classifiers across two different datasets comprising more than 51,000 labeled news articles in order to propose a comprehensive machine learning framework for fake news identification utilizing NLP approaches. The methodology adopted in this study encompasses pre-processing the text data, TF-IDF vectorisation, class balancing through SMOTE technique and hyperparameter optimisation.

The decision tree Classifier got 99.53% accuracy for the larger dataset (Fake-Real combined), while the passive-aggressive Classifier and the SVM Classifier got 93.4% accuracy for the smaller dataset (News.csv). The hyperparameter optimization combined with SMOTE-based balance yielded better results for Naive Bayes (5.5% increase from 86.0% to 91.5%) and Random Forest (0.6% increase from 90.4% to 91.0%). The decision tree had the largest improvement of 18.03%, with all the models doing better in the larger data set. This illustrates the significance of the amount of data and variation in terms of how well false news is being detected.

Much potential exists for fake news detection models on incorrect information. These methods can help limit the transmission and preserve people and society from the adverse effects of false information. As the machine learning algorithms mature, it is hoped that the techniques for identifying fake news will also improve.

XII. REFERENCES

N. R. de Oliveira, D. S. V. Medeiros, and D. M. F. Mattos, "A sensitivity-based approach to identify fake news on social networking," IEEE Signal Process. Lett., vol. 27, pp. 1250-1254, 2020.

A. R. Martyno and G. Augusta, "A survey on recent advances in machine learning techniques for fake news detection," Text Eng. Manag, vol. 83, pp. 11572-11582, 2020.

M. Mahyob, J. Algarad, and M. Alrahaili, "Linguistic-based detection of fake news in social media," Int. J. English Linguistics, vol. 11, no. 1, pp. 99, Nov. 2020.

N. R. de Oliveira, P. S. Pisa, M. A. Lopez, D. S. V. de Medeiros, and D. M. F. Mattos, "Identifying fake news on social networks based on natural language processing: Trends and challenges," Information, vol. 12, no. 1, p. 38, Jan. 2021.

D. D. N. Caragea, D. M. Caragea, and K. C. Al-Kofahi, "A Machine Learning-based System for Detecting Fake News in Social Media," in Proceedings of the 56th Hawaii International Conference on System Sciences, 2019, pp. 3220-3229.

H. Farid and K. Abdullah, "Fake News Detection Based on Machine Learning Algorithms," in 2019 International Conference on Information Technology, Islamic Republic of Iran, 2019, pp. 1-6.

S. R. S. S. Prabaharan and R. Priyanka, "Fake News Detection using Machine Learning Algorithms," in 2020 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2020, pp. 1190-1195.

J. Kim and J. Kim, "Fake News Detection Using Multi-Source Information and Machine Learning Techniques," in Proceedings of the 2019 International Conference on Computational Science and Computational Intelligence, 2019, pp. 329-334.

Amjad M., Sidorov G., Zhila A., Gomez-Adomo H., Voronkov I., Gelbukh A., "Bend the truth: benchmark dataset for fake news detection in Urdu language and its evaluation," Journal of Intelligent & Fuzzy Systems, vol. 6, no. 5, 2020.

Mohammed Ali Shaik, Makkaji Yasha Sree, Sanka Sri Vyshnavi, Thogiti Ganesh, Dasari Sashmitha, and Narmetta Shreya, "Fake News Detection using NLP," International Conference on Innovative Data Communication Technologies and Application (ICIDCA-2023), IEEE Xplore, 2023.

M. Q. Alnabhan and P. Branco, "Fake News Detection Using Deep Learning: A Systematic Literature Review," IEEE Access, vol. 12, pp. 114435-114459, 2024.

J. Alghamdi, S. Luo, and Y. Lin, "A comprehensive survey on machine learning approaches for fake news detection," Multimedia Tools and Applications, vol. 83, pp. 51009-51067, 2024.

A. J. Keya, H. H. Shajeeb, M. S. Rahman, and M. Mridha, "FakeStack: Hierarchical Tri-BERT-CNN-LSTM stacked model for effective fake news detection," PLoS ONE, vol. 18, no. 12, p. e0294701, Dec. 2023.

"Fake News Detection Using Machine Learning and Deep Learning Methods," Journal of Electronic Imaging, vol. 77, no. 2, pp. 2079-2096, Nov. 2023.

R. Upadhyay, G. Pasi, and M. Viviani, "Leveraging Socio-contextual Information in BERT for Fake Health News Detection in Social Media," in Proc. 3rd Int. Workshop Open Challenges Online Social Netw. (OASIS), Rome, Italy, Sep. 2023, pp. 1-9.

S. F. N. Azizah, "Performance Analysis of Transformer-Based Models (BERT, ALBERT, and RoBERTa) in Fake News Detection," arXiv preprint arXiv:2308.04950, Aug. 2023.

E. Puraivan, P. Ormeño-Arriagada, S. Kloss, and C. Cofré-Morales, "Surface Linguistic Features for Multiclass Fake News Detection in a Multilingual Context," in *Proc.* 10th Int. Congr. Inf. Commun. Technol. (ICICT), London, U.K., 2025, pp. 545-556.

"Advancing fake news combating using machine learning: a hybrid model approach," Knowl. Inf. Syst., vol. 67, pp. 12137-12177, 2025.

M. D. A. Rahman et al., "Web-Based Machine Learning Framework for Authenticating News: A Fake News Detection System," Research Square, preprint, 2024.

"Detecting Fake News: An Innovative and Effective Approach to Combating Misinformation in today's Digital World," in Proc. 2025 IEEE Int. Conf. Adv. Comput. (ICAC), Chennai, India, Nov. 2025.

"Detecting fake news with term frequency-inverse document frequency (TF-IDF) vectorizer," AIP Conf. Proc., vol. 3410, no. 1, p. 020036, 2026.

S. Basti, "Combating Hate Speech and Fake Narratives in Hindi-English Code-Mixed Social Media: An Ensemble Approach," in Proc. Faux-Hate Shared Task at ICON 2024, 2024.

M. R. H. et al., "A Machine Learning-Based System for Detecting Fake News in Social Media," in Proc. 56th Hawaii Int. Conf. Syst. Sci. (HICSS), 2023, pp. 3220-3229.

I. Ahmad et al., "Optimized Ensemble Machine and Deep Learning Model for Fake News Detection," J. Intell. Fuzzy Syst., vol. 45, no. 2, pp. 2345-2362, 2023.

M. Umer et al., "A Hybrid Deep Learning Model for Fake News Detection," J. Ambient Intell. Humaniz. Comput., vol. 14, pp. 7073-7084, 2023.

S. Hakak et al., "A Machine Learning Approach for Fake News Detection," in Proc. *2021* 5th Int. Conf. Inf. Syst. Comput. Netw. (ISCON), 2021.

K. Lee et al., "A Comprehensive Review on Fake News Detection Using Machine Learning and Deep Learning," IEEE Access, vol. 10, pp. 122532-122549, 2022.

S. Mahmudul Hasan, S. Roy, and A. J. Nafis, "Generalization Gaps in Political Fake News Detection: An Empirical Study on the LIAR Dataset," arXiv preprint arXiv:2512.18533, 2025.

"Optimizing Machine Learning Models based on Hyperparameter Tuning for Arabic Fake News Detection," in Proc. IEEE Conf., Gaziantep, Turkiye, Jun. 2025.

H. AbouGrad, F. Riaz, and A. Qadoos, "DistilBERT-Based Binary Classification NLP-Framework to Enhance Misinformation and Fake News Detection Accuracy Performance," Iris J. Math., vol. 1, no. 2, May 2026.

"Evaluating the Effectiveness of Fake News Detection System: Challenges and Solutions," in Innovative Computing and Communications, Lecture Notes in Networks and Systems, vol. 1433,
pp. 167-182, Oct. 2025.

"Machine Learning Techniques to Detect A Fake News In An Articles," in Proc. Second Int. Conf. Networks, Multimedia and Information Technology (NMITCON), pp. 1-6, Aug. 2024.

T. Olaleye et al., "SCLAVOEM: hyper parameter optimization approach to predictive modelling of COVID-19 infodemic tweets using smote and classifier vote ensemble," Soft Comput., vol. 27, pp. 3531-3550, 2023.

R.N. Yadawad, S. Saumya, K.N. Nivedh, S.S. Padanur, and S. Basti, "A Machine Learning Framework for Detecting Hate Speech and Fake Narratives in Hindi-English Tweets," in Proc. Faux-Hate Shared Task at ICON 2024, 2024.

O. M. Bashaddadh, N. Omar, M. Mohd, and M. N. A. Khalid, "Machine Learning and Deep Learning Approaches for Fake News Detection: A Systematic Review of Techniques, Challenges, and Advancements," IEEE Access, vol. 13, pp. 90433-90466, 2025.