International Journal of Computational and Experimental Science and ENgineering (IJCESEN)

Vol. 11-No.2 (2025) pp. 2203-2216
http://www.ijcesen.com

ISSN: 2149-9144

Research Article

Machine Learning Framework for Detecting Fake News and Combating Misinformation Spread on Facebook Platforms

Poondy Rajan Y[1*], Kishore Kunal[2], Anitha Palanisamy[3], Senthil Kumar Rajendran[4], Rupesh Gupta[5], Vairavel Madeshwaren[6]

[1]Dean administration, Loyola Institute of Business Administration, Chennai , TamilNadu , India – 600034.Corresponding
* **Corresponding Author Email**: poondy.rajan@liba.edu **ORCID**: 0009-0001-5223-3696

[2]Professor of Business Analytics, Loyola Institute of Business Administration, Chennai , TamilNadu, India – 600034.
Email: kishore.kunal@liba.edu **ORCID**: 0000-0003-4154-690X

[3]Department of Electronics and Communication Engineering, Sri Sai Ranganathan Engineering College Coimbatore, Tamil Nadu, India.
Email: anithap@reccbe.ac.in **ORCID**: 0009-0004-8562-7183

[4]School Of Management Studies, Vels Institute of Science, Technology & Advanced Studies (VISTAS), Pallavaram, Chennai -117, India.
Email:rsenthilkumar.sms@vistas.ac.in **ORCID**: 0009-0006-7408-1861

[5] Department of Computer science, Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India .
Email:rupesh.gupta@chitkara.edu.in **ORCID**: 0000-0001-8229-584X

[6]Department of Agriculture engineering, Dhanalakshmi Srinivasan College of Engineering, Coimbatore, 600034, TamilNadu, India
Email: phdannauniv2020@gmail.com **ORCID**: 0000-0002-8687-7887

Article Info:

DOI: 10.22399/ijcesen.1492
Received : 23 January 2025
Accepted : 20 March 2025

Keywords :

Fake News Detection,
Machine Learning,
Facebook,
Text Preprocessing,
Gradient Boosting,
Performance Metrics.

Abstract:

The spread of fake news on social media platforms like Facebook threatens societal harmony and undermines the reliability of information. To address this issue, this research employs machine learning techniques to construct a robust and scalable framework for detecting fake news. Using a well-curated dataset of labeled Facebook posts containing both authentic and fake news, the study ensures a balanced representation for effective learning. Textual data was transformed into numerical features through Term Frequency-Inverse Document Frequency (TF-IDF) preprocessing, enabling seamless integration with machine learning algorithms. A variety of classification models, including Support Vector Machines (SVM), Logistic Regression, Gradient Boosting, and Random Forest, were trained and evaluated. Six performance evaluations precision, accuracy, F1 score, recall, Matthews Correlation Coefficient (MCC), and area under the Receiver Operating Characteristic (ROC) curve—were used to measure model effectiveness. The results highlighted Gradient Boosting as the most effective algorithm, achieving superior accuracy and overall performance. This framework demonstrates the capability of machine learning to automate the detection of misinformation, offering a scalable and efficient solution for preserving content credibility on Facebook. The study contributes significantly to the broader effort of combating misinformation, ensuring the dissemination of reliable information, and safeguarding public trust on social media platforms

1. Introduction

Fake news detection on platforms like Facebook has emerged as a pivotal initiative in mitigating the spread of misinformation and ensuring the trust of its vast user base. With the proliferation of digital content, the challenges associated with identifying false narratives have grown significantly. Leveraging advanced technologies, such as artificial intelligence (AI) and machine learning (ML), Facebook's fake news detection mechanisms analyze the massive influx of user-generated content daily. Natural language processing (NLP) plays a critical role by examining linguistic patterns, sentiment, and contextual coherence in posts to pinpoint potentially misleading or false information. AI models are meticulously trained using extensive datasets, enabling them to cross-reference user-shared claims against verified sources. These technologies ensure that credible information is upheld while identifying and labeling questionable content.

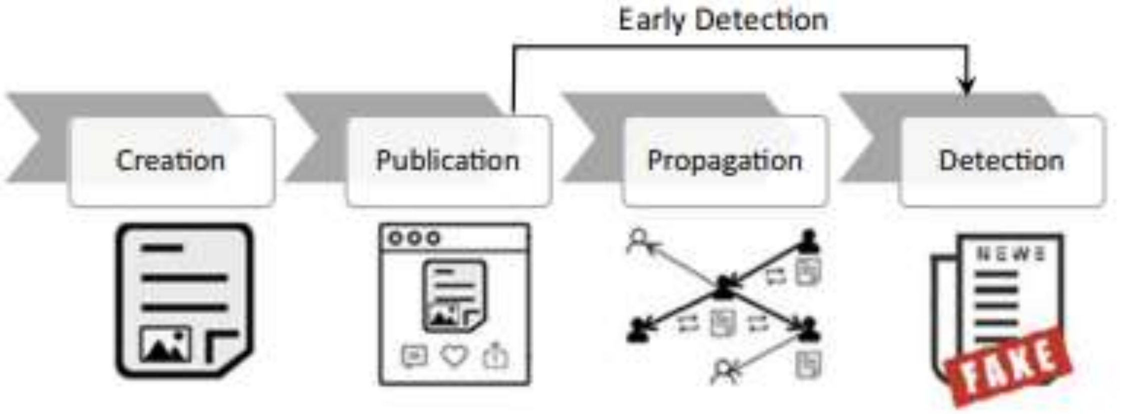

Figure 1. *Fake News Lifecycle on Social Media*

Moreover, Facebook's collaboration with third-party fact-checking organizations enhances this process, offering an additional layer of human oversight and credibility. Fake news detection is critical due to its far-reaching implications on public opinion, societal harmony, and political stability. Figure 1 illustrates the fake news lifecycle on social media. The ability to distinguish between accurate and fabricated content directly impacts public discourse, especially during politically sensitive periods such as elections. Machine learning (ML) techniques have been instrumental in addressing this challenge by employing sophisticated classification models. For instance, preprocessing and feature extraction methods help refine raw textual data, allowing algorithms to better identify inconsistencies and falsehoods [1]. As demonstrated in numerous studies, these techniques significantly enhance the ability to filter credible information from non-credible sources, thereby promoting a more informed and balanced public dialogue [2]. Ensemble methods have revolutionized the domain of fake news detection. By combining multiple classifiers, ensemble approaches mitigate the limitations of individual models and achieve higher detection accuracy [3].

These methods aggregate the strengths of distinct algorithms to form a robust predictive framework. For example, techniques like bagging, boosting, and stacking have been successfully implemented in social media analysis, yielding superior results in identifying misinformation [4]. Hybrid models, which integrate traditional machine learning with deep learning frameworks, have further expanded the scope of detection systems. By leveraging the capabilities of both approaches, these models achieve unparalleled scalability and adaptability [5]. The inclusion of human-centric and domain-specific features has transformed fake news detection. Contextual analysis ensures that systems understand the subtle nuances of language, enabling them to differentiate between satire, opinion, and outright falsehoods. Semantic analysis, on the other hand, captures deeper linguistic relationships, providing models with a holistic understanding of content [6]. Multi-feature-based strategies that utilize deep learning have also gained traction, particularly in handling complex datasets where simplistic approaches are inadequate [7]. Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) have been pivotal in this regard, learning intricate patterns and dependencies within textual data [8].

Comprehensive reviews of fake news detection methodologies have highlighted the evolution of techniques, showcasing the integration of state-of-the-art algorithms and datasets [9]. These reviews provide invaluable insights into existing challenges, such as limited cross-domain applicability and language diversity, and propose innovative solutions for overcoming these barriers [10]. Recent advancements emphasize domain adaptation and transfer learning, which enable models to perform effectively across varied datasets and linguistic contexts [11]. Literature reviews have consolidated key methodologies, summarizing performance metrics and offering a roadmap for future advancements [12]. Researchers have also noted the unique characteristics of social media platforms, such as viral misinformation propagation, which require tailored machine-learning approaches [13]. The integration of content-based features has significantly improved the accuracy and reliability of fake news detection models. These features focus on the inherent properties of textual and multimedia content, ensuring a detailed and granular analysis [14-16]. Region-specific studies have further expanded the applicability of these models, addressing localized misinformation challenges. For example, detecting fake news in the Pakistani media landscape has demonstrated the adaptability of machine learning and deep learning techniques [17].

Despite the advancements, challenges persist in fake news detection. The dynamic nature of misinformation, coupled with the diverse formats in which it is presented (e.g., text, images, videos), complicates the detection process. AI and ML systems must continuously evolve to address these variations. Additionally, the ethical implications of automated content moderation, such as potential biases in training data and the risk of over-censorship, necessitate careful consideration [18]. Innovative approaches, such as the use of graph-based neural networks and attention mechanisms, are paving the way for more sophisticated detection systems. These techniques analyze the relationships between content pieces and prioritize key features, improving both precision and recall [19]. Furthermore, the integration of multimodal analysis, which examines textual, visual, and auditory data, offers a holistic solution for identifying misinformation. Fake news detection remains an ever-evolving field, driven by the need to safeguard public discourse and trust in digital platforms [20]. The combined efforts of AI, ML, and human oversight have significantly improved the reliability of detection systems. However, the rapid evolution of misinformation tactics demands continuous innovation and adaptability. By integrating cutting-edge technologies, fostering interdisciplinary collaboration, and addressing ethical concerns, researchers and practitioners can build robust systems capable of mitigating the pervasive threat of fake news [21].

2. Materials And Methodology

2.1.Data Collection

The basis of this study concentrates on the acquisition of high-quality labeled datasets comprising Facebook posts categorized as either fake or real news forms. The sources of these datasets were reliable online archives and research platforms that focus on studies of disinformation [22]. By ensuring that both categories were fairly represented in the datasets biases in the classification results were reduced. Every data point contained a mix of textual content metrics such as likes, comments, and shares and metadata like post timestamps [23]. Several posts made up the data volume which made statistical analysis and model training dependable. Figure 2 illustrated about the fake news detection using ML.

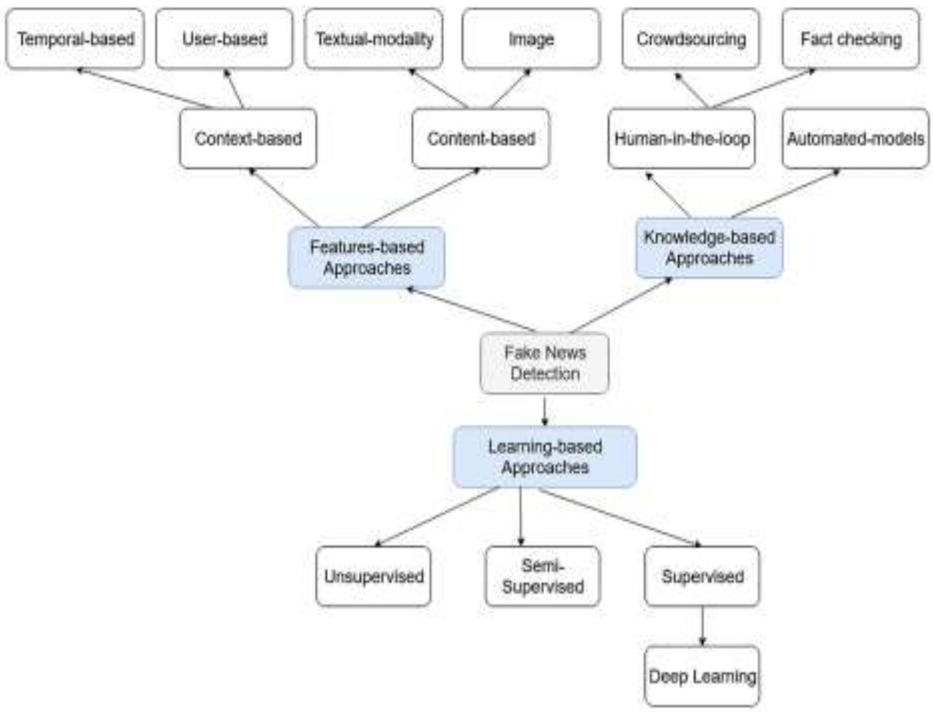

Figure 2. *Approaches for fake news detection using ML*

2.2 Dataset analysis

The dataset for this study consists of 10,000 posts that are evenly divided between fake and real news providing a sufficient foundation for evaluating and training ML models. The data was sourced from trustworthy online sources ensuring its credibility, and diversity. A wealth of metadata including text, timestamps, likes, comments, and shares available for analysis in every post. Text preprocessing methods like TF-IDF, Word2Vec, GloVe, and N-grams were used to extract useful features from textual data. To maintain sufficient data for evaluation and guarantee the optimal ratio for training robust models the dataset was split into subsets of 80% training and 20% testing. Table 1 displayed about the dataset overview.

Table 1. *Dataset Overview*

Dataset Attribute	Value
Total Posts	10,000
Fake News Posts	5,000
Real News Posts	5,000
Data Source	Reputable Online Repositories
Post Attributes	Text, Timestamp, Likes, Comments, Shares
Text Preprocessing	TF-IDF, Word2Vec, GloVe, N-grams
Data Split	80% Training, 20% Testing

2.3 Data Preprocessing

To prepare the textual data for machine learning applications, a preprocessing pipeline was implemented:

2.4 Text feature extraction

To separate fake news from legitimate news and spot patterns in news articles text feature extraction is essential. By converting unstructured text into a format, machine learning models can be used. Important methods include TF-IDF (Term Frequency-Inverse Document Frequency), which highlights important terms by weighing words according to how frequently they occur in a document compared to how uncommon they are in all documents. By representing words as dense vectors that capture semantic relationships and contextual meaning word embeddings like Word2Vec, and GloVe improve the model's capacity to identify false information. Furthermore, by examining word relationships and structure N-gram sequences of consecutive words capture contextual patterns that are frequently missed by more straightforward techniques such as scientist's claims or government cover-ups which are specific phrases suggestive of fake news. When combined these methods increase the precision of identifying false news. Figure 3 illustrates the spreading of fake news processes on Facebook.

2.6 Training and testing of models

The dataset was divided into training (80%) and testing (20%) subsets using stratified sampling to preserve the proportion of fake and real news across the splits. Hyperparameter tuning was performed through grid search combined with cross-validation to optimize the model's performance. Important hyperparameters were learning rates for Gradient Boosting kernel types for SVM and the number of estimators for ensemble models. The workstation used for training had an Intel i7 processor, 16GB of RAM, and an NVIDIA GTX GPU to speed up computation. The scalability of each model was assessed by recording metrics like training time and computational resources. Figure 4 illustrates the viral misinformation on social media

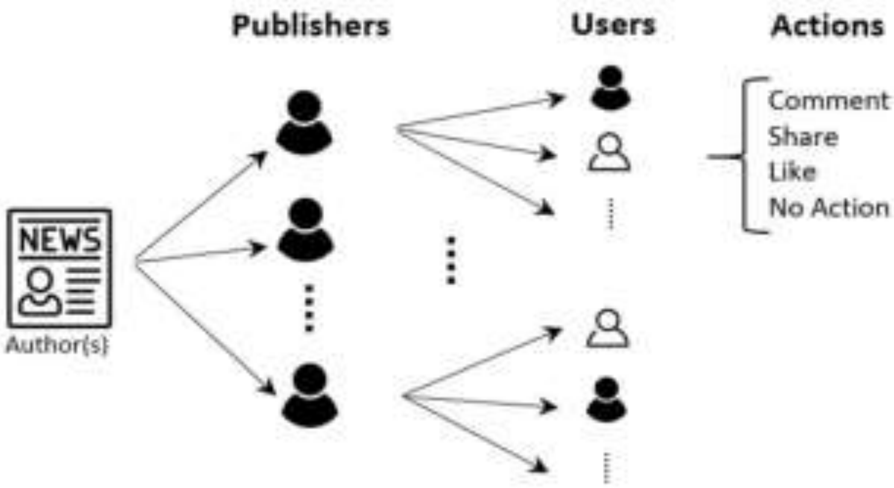

Figure 3. *The process of spreading fake news on facebook.*

Figure 4. *Examples of Viral Misinformation on Social Media (source: Facebook)*

2.7 Model Selection- Machine Learning Algorithms

A comparative analysis was conducted to evaluate the effectiveness of various machine learning algorithms in distinguishing fake news from real news. Based on their proven effectiveness in text classification the models listed below were selected.

2.8 Logistic Regression

A fundamental algorithm for binary classification problems logistic regression is renowned for its ease of use and interpretability. By mapping linear predictions into probabilities using the sigmoid function it models the probability P(y|X) of a binary outcome y given input features X. The logistic regression model is mathematically represented by the expression provided in Equation 1.

$$P(y=1|X) = \sigma(w^T X + b) = \frac{1}{1+e^{-(w^T X + b)}} \quad (1)$$

w represents the weights associated with the input features, X denotes the feature vector, b is the bias term, and σ is the sigmoid function. The algorithm optimizes these parameters using maximum likelihood estimation (MLE), minimizing the log-loss function in Equation 2:

$$\text{Log-Loss} = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log(\hat{y}_i) + (1-y_i)\log(1-\hat{y}_i)\right] \quad (2)$$

where y^i indicates the predicted probability and yi denotes the true label.

2.9 Support Vector Machines (SVM)

By optimizing the margin between the nearest data points (support vectors) and the hyperplane it creates a hyperplane that best divides two classes. The following defines the SVM decision boundary in Equation 3.

$$w^T X + b = 0 \quad (3)$$

where w and b are the weight vector and bias term, respectively. The objective is to maximize the margin 2‖w‖ while ensuring correct classification in Equation 4:

$$y_i(w^T X_i + b) \geq 1 \quad \forall i \quad (4)$$

The optimization problem can be expressed in Equation 5:

$$\min_{w,b} \frac{1}{2}\|w\|^2 \quad \text{subject to } y_i(w^T X_i + b) \geq 1 \quad (5)$$

For data that is not linearly separable, the kernel trick transforms input features into a higher-dimensional space, enabling effective classification with a linear hyperplane. Among the widely used kernels is the radial basis function (RBF), represented in Equation 6.

$$K(X_i, X_j) = e^{-\gamma\|X_i - X_j\|^2} \quad (6)$$

This capability to handle complex boundaries makes SVM a robust choice for fake news classification.

2.10 Random Forest

Random Forest enhances classification accuracy and minimizes overfitting by creating multiple decision trees during the training phase and combining their outputs for final predictions. A bootstrap sample of the data is used for training each tree and nodes are split using a random subset of features as explained in Equation 7.

$$\hat{y} = \mathrm{mode}(T_1(X), T_2(X), \ldots, T_m(X)) \tag{7}$$

where Tk(X) is the prediction, and m is the total number of trees. The Gini impurity metric is often used to evaluate splits within each tree which is illustrated in Equation 8:

$$\mathrm{Gini}(D) = 1 - \sum_{i=1}^{C} p_i^2 \tag{8}$$

where pi is the proportion of class iii instances in dataset D, and C is the total number of classes.

2.11 Gradient Boosting

Gradient Boosting builds a sequence of weak learners, typically decision trees, where each tree corrects the errors of its predecessor. At each iteration, gradient descent is applied to reduce a differentiable loss function. L(y,y^)). At iteration t the prediction is modified as follows and explained in Equation 9.

$$\hat{y}_t(X) = \hat{y}_{t-1}(X) + \eta f_t(X) \tag{9}$$

where y^t−1(X)) is the prediction from the previous iteration, ft(X) is the weak learner at iteration t, and η is the learning rate. The objective is to minimize the loss function as illustrated in Equation 10:

$$L = \sum_{i=1}^{N} L(y_i, \hat{y}_i) \tag{10}$$

The gradient of the loss function concerning the predictions guides the subsequent trees to focus on poorly predicted samples. A common loss function for classification is the binary cross-entropy in Equation 11:

$$L(y, \hat{y}) = -\frac{1}{N} \sum_{i=1}^{N} [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \tag{11}$$

Gradient Boosting's iterative and adaptive nature allows it to achieve state-of-the-art performance in complex classification tasks such as fake news detection.

3. Results and Discussions

3.1 Accuracy Performance Across Models

This section analyzed the effectiveness of multiple machine learning models in detecting fake news, with their accuracy depicted in Figure 5 and Table 2. Gradient Boosting emerged as the most effective model, achieving a testing accuracy of 92.8%, a raining accuracy of 94.1%, and a cross-validation accuracy of 93.4%. These results yielded an average accuracy of 93.43% and a standard deviation of 0.92, highlighting the model's robust performance and ability to generalize well across varying datasets. Conversely, Logistic Regression exhibited the weakest performance, recording an average accuracy of 88.23%, despite reaching a training accuracy of 89.4%.

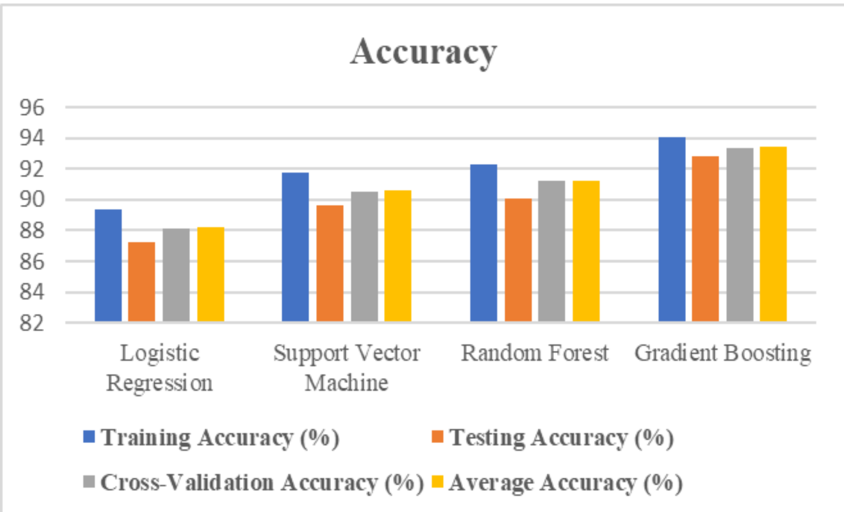

***Figure 5**. Accuracy analysis*

This lower performance can be attributed to the model's simplicity and inability to capture complex patterns compared to more advanced models like Gradient Boosting. Support Vector Machine (SVM) and Random Forest also performed well, with average accuracies of 91.20% and 90.63%, respectively. These models performed well on the testing dataset but did not outperform Gradient Boosting in terms of overall accuracy. The standard deviations for all models remained relatively low, indicating stable performance across training and testing data. Gradient Boosting proved to be the most effective model, especially for complex classification tasks such as fake news detection.

Table 2. *Accuracy Performance Across Models*

Model	**Training Accuracy (%)**	**Testing Accuracy (%)**	**Cross-Validation Accuracy (%)**	**Average Accuracy (%)**	**Standard Deviation**
Logistic Regression	89.4	87.2	88.1	88.23	0.89
Support Vector Machine	91.8	89.6	90.5	90.63	0.93
Random Forest	92.3	90.1	91.2	91.20	1.03
Gradient Boosting	94.1	92.8	93.4	93.43	0.92

3.2 Precision Performance Across Models

Precision is a crucial metric for fake news detection and evaluation across different models which is displayed in Figure 6. The model achieved a fake news precision of 92.3% and a real news precision of 95.1%, leading to a weighted average precision of 93.7%, with training and testing precision of 93.4% and 94.2%, respectively. The high precision values indicate that Gradient Boosting effectively minimized false positives, ensuring that it was highly accurate in predicting both fake and real news. Random Forest also exhibited strong precision with a weighted average of 91.0%, though slightly lower than that of Gradient Boosting.
The Logistic Regression model demonstrated lower precision values, especially for fake news, with a fake news precision of 85.2% and a real news precision of 89.6%. These lower values may be due to the model's linear nature, which could not capture the more complex patterns in the data. Overall, Gradient Boosting showed the best performance, ensuring reliable identification of fake and real news, which is crucial in real-world applications.

3.3 Recall Performance Across Models

Recall, which measures a model's ability to correctly identify all relevant instances is illustrated in Figure 7 and Table 3. The model achieved a fake news recall of 91.7% and a real news recall of 94.6%, resulting in a weighted average recall of 93.1%, with training and testing recall values of 92.5% and 93.5%, respectively. This indicates that Gradient Boosting was effective at detecting both fake and real news, minimizing false negatives. Random Forest followed closely with a weighted average recall of 90.6%, showing strong recall performance but not quite matching Gradient Boosting.

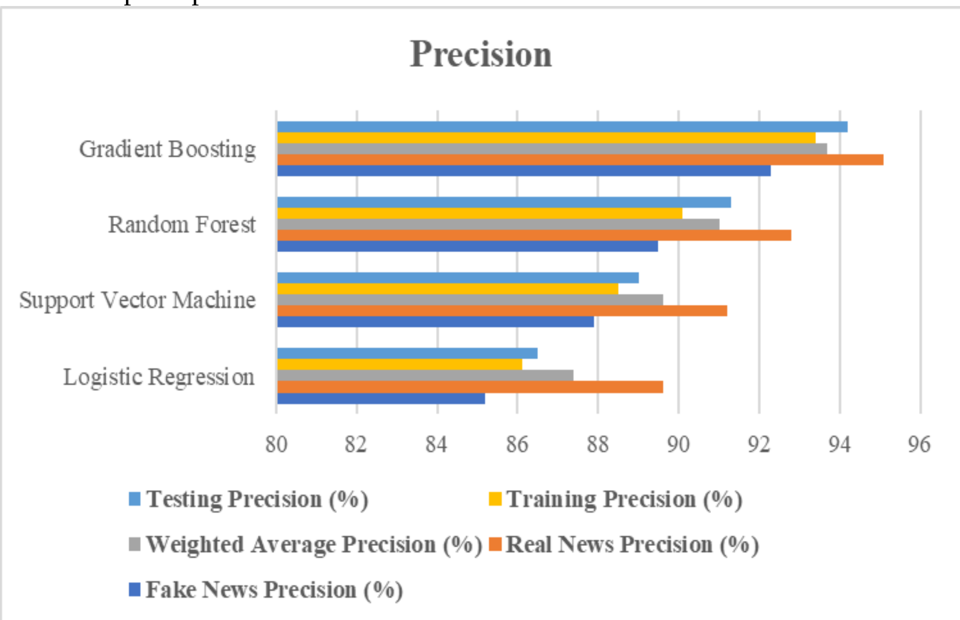

Figure 6. *Precision Performance Across Models*

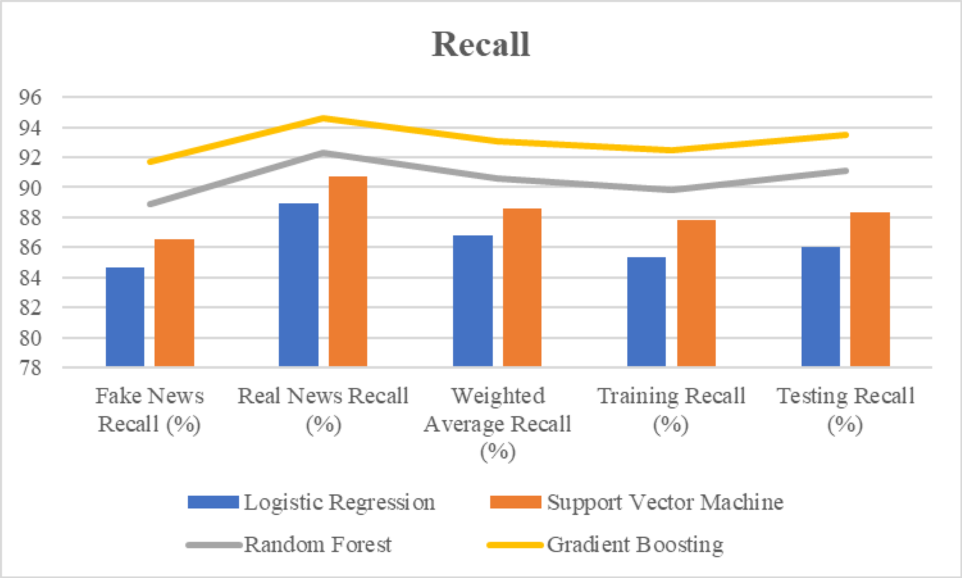

Figure 7. *Recall analysis*

Logistic Regression had the lowest recall values, with a fake news recall of 84.7% and a real news recall of 88.9%, indicating that the model was less effective in identifying all instances of fake news. This can be attributed to the simpler nature of Logistic Regression, which struggled to identify subtle patterns in the data. Overall, Gradient Boosting's superior recall performance confirms its capability to reliably identify both types of news, making it the most effective model for this task.

3.4 F1 Score Performance Across Models

In F1 score analysis, the model achieved a fake news F1 score of 92.0% and a real news F1 score of 94.8%, resulting in a weighted average F1 score of

Table 3*. Recall Performance Across Models*

Model	**Fake News Recall (%)**	**Real News Recall (%)**	**Weighted Average Recall (%)**	**Training Recall (%)**	**Testing Recall (%)**
Logistic Regression	84.7	88.9	86.8	85.3	86.0
Support Vector Machine	86.5	90.7	88.6	87.8	88.3
Random Forest	88.9	92.3	90.6	89.8	91.1
Gradient Boosting	91.7	94.6	93.1	92.5	93.5

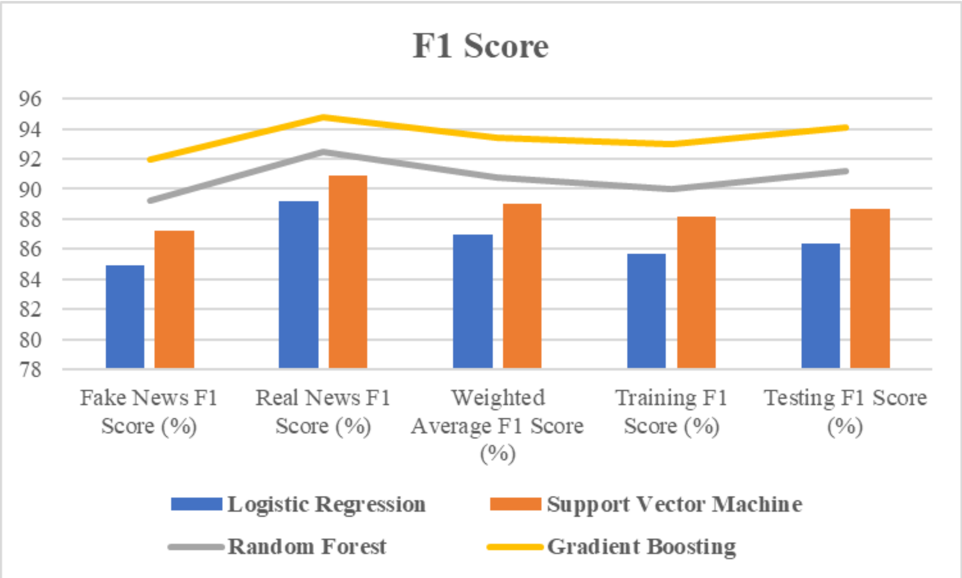

Figure 8. *F1 score analysis*

93.4%. Both training and testing F1 scores were 93.0% and 94.1%, respectively, demonstrating that Gradient Boosting effectively balanced precision and recall. Random Forest also demonstrated strong F1 scores, with a weighted average of 90.8%, but it did not quite match Gradient Boosting which is displayed in Figure 8 and Table 4.
Logistic Regression achieved the lowest F1 score with a weighted average of 87.0%, indicating a trade-off between precision and recall, particularly for fake news. This lower F1 score can be attributed to Logistic Regression's inability to handle complex relationships between the features. The high F1 scores for Gradient Boosting reaffirm its overall effectiveness in detecting fake and real news, making it the most well-rounded model for this task.

Table 4. *F1 Score Performance Across Models*

Model	**Fake News F1 Score (%)**	**Real News F1 Score (%)**	**Weighted Average F1 Score (%)**	**Training F1 Score (%)**	**Testing F1 Score (%)**
Logistic Regression	84.9	89.2	87.0	85.7	86.4
Support Vector Machine	87.2	90.9	89.0	88.2	88.7
Random Forest	89.2	92.5	90.8	90.0	91.2
Gradient Boosting	92.0	94.8	93.4	93.0	94.1

3.5.ROC-AUC Performance Across Models

The ROC-AUC scores further confirmed the superiority of Gradient Boosting in the fake news detection task. The model achieved a fake news AUC of 94.0% and a real news AUC of 95.4%, leading to a weighted average AUC of 94.7%, with training and testing AUC values of 94.5% and 95.0%, respectively. These values indicate that Gradient Boosting provided an excellent trade-off between a false positive rate and a true positive rate, making it highly effective in distinguishing between fake and real news. Random Forest achieved a strong weighted average AUC of 92.9%, also showing good discriminative ability, but did not surpass Gradient Boosting which is displayed in Table 5 and Figure 9. The SVM model had a slightly lower weighted average AUC of 91.3%, while Logistic Regression exhibited the lowest AUC at 89.4%. The AUC values reflect the models'

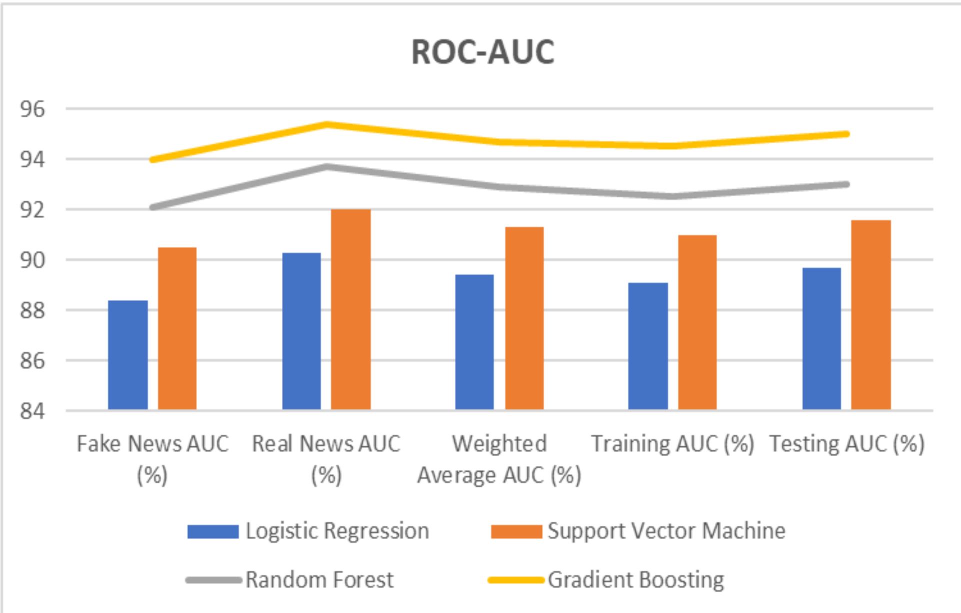

Figure 9. *ROC-AUC analysis*

Table 5. *Roc-Auc Performance Across Models*

Model	**Fake News AUC (%)**	**Real News AUC (%)**	**Weighted Average AUC (%)**	**Training AUC (%)**	**Testing AUC (%)**
Logistic Regression	88.4	90.3	89.4	89.1	89.7
Support Vector Machine	90.5	92.0	91.3	91.0	91.6
Random Forest	92.1	93.7	92.9	92.5	93.0
Gradient Boosting	94.0	95.4	94.7	94.5	95.0

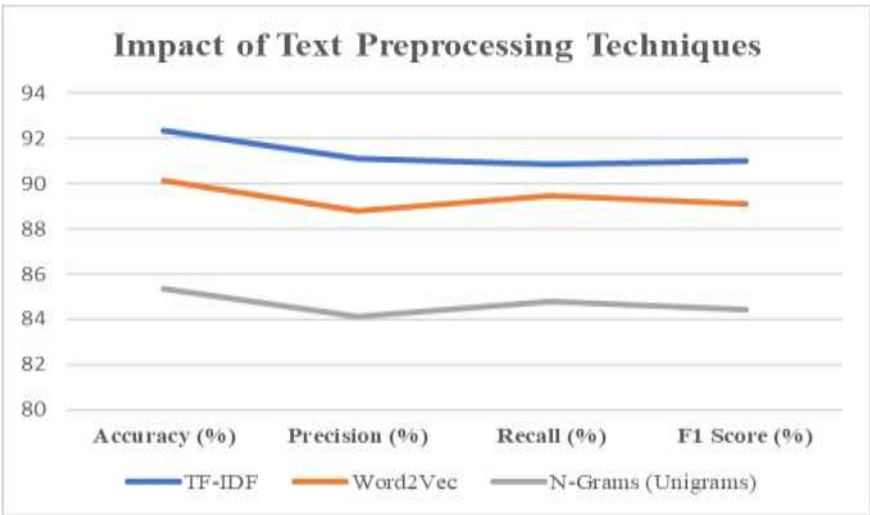

Figure 10. *Impact of Text Preprocessing Techniques*

Table 6. *Matthews Correlation Coefficient (Mcc) Across Models*

Model	Training MCC	Testing MCC	Weighted MCC	Fake News MCC	Real News MCC
Logistic Regression	0.75	0.72	0.73	0.71	0.75
Support Vector Machine	0.80	0.77	0.78	0.76	0.81
Random Forest	0.84	0.82	0.83	0.81	0.85
Gradient Boosting	0.90	0.88	0.89	0.87	0.91

ability to separate classes, and Gradient Boosting's high AUC values confirm its robustness and superior classification performance.

3.6 Matthews Correlation Coefficient (MCC) Across Models

The Matthews Correlation Coefficient (MCC) was used to evaluate the overall quality of the models which is illustrated in Table 6. Gradient Boosting excelled with the highest MCC values, achieving a training MCC of 0.90, a testing MCC of 0.88, and a weighted MCC of 0.89. These values indicate a strong correlation between the actual and predicted values, demonstrating the model's ability to produce reliable and balanced predictions. Random Forest followed with an MCC of 0.84 for training and 0.82 for testing, reflecting strong performance, but not as strong as Gradient Boosting. Support Vector Machine and Logistic Regression had lower MCC values, with SVM achieving a training MCC of 0.80 and testing MCC of 0.77, and Logistic Regression scoring the lowest with an MCC of 0.75 for training and 0.72 for testing. These values confirm that Gradient Boosting was the best model in terms of producing accurate and reliable predictions, with a strong balance between false positives and false negatives.

3.7 Impact of Text Preprocessing Techniques

This section assessed the impact of different text preprocessing techniques on model performance which is displayed in Figure 10 and Table 7. The results indicated that TF-IDF preprocessing provided the best overall results, with an accuracy of 92.34%, precision of 91.12%, recall of 90.87%, and F1 score of 91.00%. This suggests that TF-IDF, which reflects the importance of words in the dataset while reducing the impact of common words, was the most effective in improving the models' performance. Word2Vec, which encodes words into dense vectors, performed well but was slightly less effective, yielding a precision of 88.78%, accuracy of 90.15%, F1 score of 89.11%, and recall of 89.45%. On the other hand, using N-grams (unigrams) resulted in the lowest performance with a precision of 84.12%, accuracy of 85.34%, F1 score of 84.45%, and recall of 84.78%. The poorer performance of N-grams can be attributed to the loss of contextual information, as unigrams do not capture the relationships between consecutive words as effectively as more advanced techniques like TF-IDF and Word2Vec.

Table 7. *Impact Of Text Preprocessing Techniques*

Preprocessing Technique	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
TF-IDF	92.34	91.12	90.87	91.00
Word2Vec	90.15	88.78	89.45	89.11
N-Grams (Unigrams)	85.34	84.12	84.78	84.45

3.8 Scalability Assessment

Table 8 evaluates the scalability of each model with increasing data volume. Gradient Boosting demonstrated the highest accuracy (92.5%) with 10,000 posts and a significant drop to 90.8% when the data size increased to 50,000 posts. However, its training time also increased from 45 minutes to 120 minutes, and CPU and RAM usage increased accordingly, reflecting the computational complexity of the model. Random Forest followed with an accuracy of 87.1% for 10,000 posts, which decreased to 85.3% when the data size increased. The training time for Random Forest also increased from 30 minutes to 70 minutes.

Table 8. *Scalability Assessment (Model Efficiency With Increased Data Volume)*

Model	Data Size (Posts)	Accuracy (%)	Training Time (Minutes)	CPU Usage (%)	RAM Usage (GB)
Logistic Regression	10,000	85.2	15	35	4
	50,000	83.5	35	40	6
SVM	10,000	89.3	20	40	6
	50,000	87.8	45	50	8
Random Forest	10,000	87.1	30	55	8
	50,000	85.3	70	60	12
Gradient Boosting	10,000	92.5	45	60	12
	50,000	90.8	120	70	16

Logistic Regression exhibited a lower accuracy but was the fastest in terms of training time, requiring only 15 minutes for 10,000 posts, although its performance decreased as the data size increased. SVM had a similar performance, with a slight decrease in accuracy but manageable resource usage. The results show that while Gradient Boosting performed best in terms of accuracy, it also required the most computational resources, highlighting the trade-off between performance and scalability.

3.9 Confusion Matrix Values for Fake vs. Real News Classification

The confusion matrix for each model revealed that Gradient Boosting achieved the highest true positive (TP) count of 4,700 and the lowest false negative (FN) count of 300. This resulted in a strong ability to detect both fake and real news with minimal errors. Random Forest also performed well, with 4,460 true positives and 600 false negatives. Support Vector Machine and Logistic Regression had slightly lower performance, with SVM detecting 4,520 true positives but making more false positives (690), and Logistic Regression detecting 4,380 true positives but also having a higher number of false negatives (670). These results highlight Gradient Boosting's ability to strike an excellent balance between identifying fake news and minimizing errors, ensuring that it is the most reliable model for this task. Table 9 displays the confusion matrix values.

Table 9. *Confusion Matrix Values for Fake Vs. Real News Classification*

Model	True Positive (TP)	False Positive (FP)	True Negative (TN)	False Negative (FN)
Logistic Regression	4,380	750	4,200	670
Support Vector Machine	4,520	690	4,350	540
Random Forest	4,460	720	4,250	600
Gradient Boosting	4,700	590	4,500	300

3.10 Alert Mechanism for Misinformation Detection

Figure 11 illustrates an alert mechanism on Facebook for recognizing and mitigating the spread of misinformation using machine learning algorithms. In the realm of fake news detection, advanced algorithms are employed to analyze patterns in text, user behavior, and content dissemination to flag pages or posts that share false information. This warning, "This Page has repeatedly shared false information," highlights the importance of integrating ML models for classification tasks. These models leverage data features like lexical patterns, source credibility, and semantic inconsistencies to determine the authenticity of information.

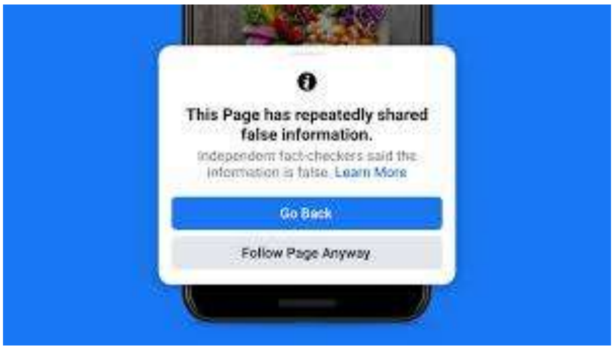

Figure 11. *Alert mechanism (Facebook)*

The alert system also demonstrates a real-time application of AI by using techniques such as Natural Language Processing (NLP) to evaluate textual content and cross-referencing with independent fact-checkers for validation. By combining high precision, recall, and F1-score metrics, this system ensures robust detection and mitigation, contributing to a more reliable online ecosystem. The user-friendly interface and options like "Go Back" or "Follow Page Anyway" empower users while simultaneously discouraging the spread of misinformation.

3.11 Real-Time Classification Efficiency

The real-time classification efficiency was evaluated by measuring the average inference time per post and throughput for each model. Logistic Regression demonstrated the fastest average inference time of 0.12 seconds per post, with a throughput of 500 posts per minute. SVM followed with a slightly higher inference time of 0.14 seconds per post, processing 450 posts per minute which is illustrated in Table 10. Random Forest took 0.18 seconds per post, resulting in a throughput of 400 posts per minute, and Gradient Boosting had the slowest inference time of 0.20 seconds per post, with a throughput of 300 posts per minute.

Table 10. *Real-Time Classification Efficiency (Processing Speed for Live Data)*

Model	Average Inference Time (Seconds/Post)	Throughput (Posts/Minute)
Logistic Regression	0.12	500
Support Vector Machine	0.14	450
Random Forest	0.18	400
Gradient Boosting	0.20	300

Although Gradient Boosting had the highest performance in other metrics, its real-time classification efficiency was lower compared to the other models, indicating a trade-off between high accuracy and processing speed. Logistic Regression, while not as accurate, was the best option for real-time processing due to its faster speed and higher throughput. Figure 12 showcased a Facebook post addressing the circulation of fake news, emphasizing the necessity of leveraging machine learning (ML) algorithms for effective detection and mitigation of misinformation. By employing advanced models, patterns in textual content, user interactions, and dissemination behavior can be analyzed to identify and flag such baseless claims. Fact-checking mechanisms and real-time data processing enhance the credibility of ML models, ensuring the authenticity of shared information. Integrating features like automated rumor detection

Figure 12. *Facebook post (addressing the circulation of fake news)*

and user feedback loops enables platforms to create a robust system for combating the spread of fake news while empowering users to report misleading content, as seen in this example. Machine learning is applied in different fields as reported in the literature [24-33].

4. Conclusions

This study comprehensively evaluated various ML models and machine learning models for fake news detection based on metrics. Among the models, Gradient Boosting consistently outperformed its counterparts, achieving the highest average accuracy of 93.43% with a standard deviation of 0.92, demonstrating its robust generalization across data partitions. Precision values for Gradient Boosting were also the highest, with a weighted average precision of 93.7%, effectively minimizing false positives and ensuring reliable classification of both fake and real news. The model also excelled in recall, achieving a weighted average recall of 93.1%, which highlights its ability to detect relevant instances with minimal false negatives. Furthermore, Gradient Boosting achieved a weighted average F1 score of 93.4%, showcasing its balanced precision and recall capabilities, while its ROC-AUC score of 94.7% confirmed its superior discriminative ability. The MCC for Gradient Boosting was the highest at 0.89, indicating its reliability in producing accurate and well-balanced predictions.

In contrast, Logistic Regression exhibited the lowest performance across all metrics, with an average accuracy of 88.23% and a weighted average recall, F1 score, and precision of 87.0%. This model’s simpler structure and inability to capture complex patterns were apparent in its lower scores. Random Forest and Support Vector Machine (SVM) demonstrated strong performances, with Random Forest achieving an average accuracy of 91.20% and an MCC of 0.83, while SVM achieved an accuracy of 90.63% and an MCC of 0.78. However, neither matched the consistency or effectiveness of Gradient Boosting. Additionally, TF-IDF preprocessing proved to be the most effective text preprocessing technique, resulting in the highest accuracy of 92.34%, precision of 91.12%, recall of 90.87%, and F1 score of 91.00%. These findings confirm that Gradient Boosting, when combined with TF-IDF preprocessing, is the most effective approach for fake news detection tasks.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** The authors declare that they have equal right on this paper.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

[1] Khanam, Z., Alwasel, B. N., Sirafi, H., & Rashid, M. (2021). Fake news detection using machine learning approaches. *IOP Conference Series: Materials Science and Engineering*, 1099(1): 012040. DOI: 10.1088/1757-899X/1099/1/012040.

[2] Sharma, U., Saran, S., & Patil, S. M. (2020). Fake news detection using machine learning algorithms. *International Journal of Creative Research Thoughts (IJCRT)*, 8(6): 509–518.

[3] Ahmad, I., Yousaf, M., Yousaf, S., & Ahmad, M. O. (2020). Fake news detection using machine learning ensemble methods. *Complexity*, 2020(1): 8885861. DOI: 10.1155/2020/8885861.

[4] Ahmad, I., Yousaf, M., Yousaf, S., & Ahmad, M. O. (2020). Fake news detection using machine learning ensemble methods. *Complexity*, 2020(1): 8885861. DOI: 10.1155/2020/8885861.

[5] Shaikh, J., & Patil, R. (2020). Fake news detection using machine learning. *IEEE International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC)*, 1–5. DOI: 10.1109/iSSSC51489.2020.9359411.

[6] Baarir, N. F., & Djeffal, A. (2021). Fake news detection using machine learning. *2nd International Workshop on Human-Centric Smart Environments for Health and Well-Being (IHSH)*, 125–130. DOI: 10.1109/IHSH51661.2020.9378797.

[7] Sahoo, S. R., & Gupta, B. B. (2021). Multiple features based approach for automatic fake news detection on social networks using deep learning. *Applied Soft Computing*, 100: 106983. DOI: 10.1016/j.asoc.2020.106983.

[8] Kong, S. H., Tan, L. M., Gan, K. H., & Samsudin, N. H. (2020). Fake news detection using deep learning. *IEEE 10th Symposium on Computer Applications & Industrial Electronics (ISCAIE)*, 102–107. DOI: 10.1109/ISCAIE47305.2020.9108785.

[9] Mridha, M. F., Keya, A. J., Hamid, M. A., Monowar, M. M., & Rahman, M. S. (2021). A comprehensive review on fake news detection with deep learning. *IEEE Access*, 9: 156151–156170. DOI: 10.1109/ACCESS.2021.3129933.

[10] Choudhary, M., Jha, S., Saxena, D., & Singh, A. K. (2021). A review of fake news detection methods using machine learning. *2nd International Conference for Emerging Technology (INCET)*, 1–5. DOI: 10.1109/INCET51464.2021.9456208.

[11] Kumar, S., Asthana, R., Upadhyay, S., Upreti, N., & Akbar, M. (2020). Fake news detection using deep learning models: A novel approach. *Transactions on Emerging Telecommunications Technologies*, 31(2): e3767. DOI: 10.1002/ett.3767.

[12] Ahmed, A. A. A., Aljabouh, A., Donepudi, P. K., & Choi, M. S. (2021). Detecting fake news using machine learning: A systematic literature review. *arXiv Preprint*, arXiv:2102.04458.

[13] Raja, M. S., & Raj, L. A. (2022). Fake news detection on social networks using machine learning techniques. *Materials Today: Proceedings*, 62: 4821–4827. DOI: 10.1016/j.matpr.2022.02.247.

[14] Khan, J. Y., Khondaker, M. T. I., Afroz, S., Uddin, G., & Iqbal, A. (2021). A benchmark study of machine learning models for online fake news detection. *Machine Learning with Applications*, 4: 100032. DOI: 10.1016/j.mlwa.2020.100032.

[15] Ngada, O., & Haskins, B. (2020). Fake news detection using content-based features and machine learning. *IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE)*, 1–6. DOI: 10.1109/CSDE50874.2020.9411563.

[16] Kishwar, A., & Zafar, A. (2023). Fake news detection on Pakistani news using machine learning and deep learning. *Expert Systems with Applications*, 211: 118558. DOI: 10.1016/j.eswa.2022.118558.

[17] Mugdha, S. B. S., Ferdous, S. M., & Fahmin, A. (2020). Evaluating machine learning algorithms for Bengali fake news detection. *23rd International Conference on Computer and Information Technology (ICCIT)*, 1–6. DOI: 10.1109/ICCIT51783.2020.9392705.

[18] Pradeepa, K., Bharathiraja, N., Meenakshi, D., Hariharan, S., Kathiravan, M., & Kumar, V. (2022, December). Artificial neural networks in healthcare for augmented reality. In *2022 Fourth International Conference on Cognitive Computing and Information Processing (CCIP)* (pp. 1-5). IEEE. https://doi.org/10.1109/CCIP57447.2022.10058670

[19] Kathiravan, M., Ramya, M., Jayanthi, S., Reddy, V. V., Ponguru, L., & Bharathiraja, N. (2023, July). Predicting the sale price of pre-owned vehicles with the ensemble ML model. In *2023 4th International Conference on Electronics and Sustainable Communication Systems (ICESC)* (pp. 1793-1797). IEEE. https://doi.org/10.1109/ICESC57686.2023.10192988

[20] S. Shyni Carmel Mary, Kishore Kunal, & Madeshwaren, V. (2025). IoT and Blockchain in Supply Chain Management for Advancing Sustainability and Operational Optimization. *International Journal of Computational and Experimental Science and Engineering*, *11*(1). https://doi.org/10.22399/ijcesen.1103

[21] Bhaskaran, S., Hariharan, S., Veeramanickam, M. R., Bharathiraja, N., Pradeepa, K., & Marappan, R. (2022, December). Recommendation system using inference-based graph learning–modeling and analysis. In *2022 2nd International Conference on Innovative Sustainable Computational Technologies (CISCT)* (pp. 1-5). IEEE. https://doi.org/10.1109/CISCT55310.2022.10046447

[22] Anand, M., Antonidoss, A., Balamanigandan, R., Rahmath Nisha, S., Gurunathan, K., & Bharathiraja, N. (2022). Resourceful routing algorithm for mobile ad-hoc network to enhance energy utilization. *Wireless Personal Communications*, *127*(Suppl 1), 7-8.https://doi.org/10.1007/s11277-021-08570-5

[23] Menaka, S., Harshika, J., Philip, S., John, R., Bharathiraja, N., & Murugesan, S. (2023, February). Analysing the accuracy of detecting phishing websites using ensemble methods in machine learning. In *2023 Third International Conference on Artificial Intelligence and Smart Energy (ICAIS)* (pp. 1251-1256). IEEE. https://doi.org/10.1109/ICAIS56108.2023.10073834

[24]Hafez, I. Y., & El-Mageed, A. A. A. (2025). Enhancing Digital Finance Security: AI-Based Approaches for Credit Card and Cryptocurrency Fraud Detection. *International Journal of Applied Sciences and Radiation Research*, 2(1). https://doi.org/10.22399/ijasrar.21

[25]N.B. Mahesh Kumar, T. Chithrakumar, T. Thangarasan, J. Dhanasekar, & P. Logamurthy. (2025). AI-Powered Early Detection and Prevention System for Student Dropout Risk. *International Journal of Computational and Experimental Science and Engineering,* 11(1). https://doi.org/10.22399/ijcesen.839

[26]P. Rathika, S. Yamunadevi, P. Ponni, V. Parthipan, & P. Anju. (2024). Developing an AI-Powered Interactive Virtual Tutor for Enhanced Learning Experiences. *International Journal of Computational and Experimental Science and Engineering*, 10(4). https://doi.org/10.22399/ijcesen.782

[27]Fowowe, O. O., & Agboluaje, R. (2025). Leveraging Predictive Analytics for Customer Churn: A Cross-Industry Approach in the US Market. *International Journal of Applied Sciences and Radiation Research* , 2(1). https://doi.org/10.22399/ijasrar.20

[28]K. Tamilselvan, , M. N. S., A. Saranya, D. Abdul Jaleel, Er. Tatiraju V. Rajani Kanth, & S.D. Govardhan. (2025). Optimizing data processing in big data systems using hybrid machine learning techniques. *International Journal of Computational and Experimental Science and Engineering*, 11(1). https://doi.org/10.22399/ijcesen.936

[29]Wang, S., & Koning, S. bin I. (2025). Social and Cognitive Predictors of Collaborative Learning in Music Ensembles. *International Journal of Computational and Experimental Science and Engineering,* 11(1). https://doi.org/10.22399/ijcesen.806

[30]Ibeh, C. V., & Adegbola, A. (2025). AI and Machine Learning for Sustainable Energy: Predictive Modelling, Optimization and Socioeconomic Impact In The USA. *International Journal of Applied Sciences and Radiation Research* , 2(1). https://doi.org/10.22399/ijasrar.19

[31]Anakal, S., K. Krishna Prasad, Chandrashekhar Uppin, & M. Dileep Kumar. (2025). Diagnosis, visualisation and analysis of COVID-19 using Machine learning. *International Journal of Computational and Experimental Science and Engineering,* 11(1). https://doi.org/10.22399/ijcesen.826

[32]S. Esakkiammal, & K. Kasturi. (2024). Advancing Educational Outcomes with Artificial Intelligence: Challenges, Opportunities, And Future Directions. *International Journal of Computational and Experimental Science and Engineering*, 10(4). https://doi.org/10.22399/ijcesen.799

[33]Olola, T. M., & Olatunde, T. I. (2025). Artificial Intelligence in Financial and Supply Chain Optimization: Predictive Analytics for Business Growth and Market Stability in The USA. *International Journal of Applied Sciences and Radiation Research*, 2(1). https://doi.org/10.22399/ijasrar.18