ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR

CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW

Vol.03 No.02 (2025)

FAKE NEWS IDENTIFICATION AND CLASSIFICATION USING MACHINE LEARNING

1. Faiqa Maqsood
Department of Computer Science
Superior University Lahore, Pakistan.
2. Muhammad Azam
Department of Computer Science
Superior University Lahore, Pakistan.
3. Muhammad Mahtab
Department of Computer Science
Superior University Lahore, Pakistan
4. Sadia Sahar
Department of Computer Science
Superior University Lahore, Pakistan

Corresponding Author: Department of Computer Science Superior University Lahore, Pakistan.
Email: *faiqa.maqsood57@gmail.com*

Abstract

This paper investigates the application of traditional machine learning algorithms for the detection of fake news using the "Fake News Detection: The Battle Against Misinformation" dataset. The study implements and evaluates the performance of Naive Bayes, Support Vector Machine (SVM), Logistic Regression, K-Nearest Neighbors, and Support Vector Classifier (SVC) models on this binary classification task. The dataset comprises Fake - 5000, Real - 4900 news items labeled as either "Fake" or "Real." The performance of each model is assessed using key evaluation metrics including accuracy, precision, recall, and F1-score, based on experimental results obtained on a held-out test set. The findings reveal that SVC is the best performing models based on results, SVM, Logistic Regression, and SVC achieve high accuracy around 99%, demonstrating their effectiveness in distinguishing between real and fake news within this dataset. The study provides a comparative analysis of these classical machine learning approaches, highlighting their strengths and limitations for fake news detection and offering insights for future research in this critical area.

Keywords: *Fake News Detection, Machine Learning, Text Classification, Natural Language Processing, Naive Bayes, Support Vector Machine, Logistic Regression, K-Nearest Neighbors, Support Vector Classifier.*

1. Introduction

In an era defined by the pervasive influence of digital platforms, the rapid dissemination of information has become both a boon and a bane [1][2]. While the internet and social media have democratized access to news and facilitated global connectivity, they have also inadvertently provided fertile ground [3][4] for the proliferation of misleading or entirely fabricated information[5][6], commonly referred to as "fake news". This phenomenon poses a significant threat to the integrity of media[7], erodes public trust in established institutions, and can have profound societal consequences, influencing public opinion, political discourse, and even public health[8]. The urgency of developing effective mechanisms for the detection and classification of fake news has thus become paramount in contemporary society .

Machine learning (ML) [9][10][11] has emerged as a powerful paradigm in the endeavor to combat the spread of misinformation. Its ability to analyze vast quantities of textual data and identify subtle patterns that distinguish genuine news from fabricated content makes it an indispensable tool for researchers and practitioners alike. Both traditional machine learning models[12], such as Naïve Bayes and Support Vector Machines (SVMs) , and more advanced deep learning architectures[13], including Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) [14]have been explored for their efficacy in this domain. These approaches leverage various features extracted from news articles to classify them as either authentic or deceptive.

However, the field of fake news detection is not without its challenges and controversies. Issues such as the inherent bias in training datasets , the dynamic and evolving nature of misinformation tactics , and the complexities of capturing nuanced linguistic cues continue to pose significant hurdles. Furthermore, the ethical implications of deploying automated detection systems, including potential censorship and the amplification of existing societal biases, necessitate careful consideration .

This research aims to contribute to the ongoing efforts in fake news detection. Specifically, this study seeks to answer the following research questions:

- **Research Question 1:** How accurately can traditional machine learning algorithms, namely Naive Bayes, SVM, Logistic Regression, K-Nearest Neighbors, and Support Vector Classifier (SVC), classify fake and real news articles within the provided dataset?
- **Research Question 2:** What are the comparative strengths and weaknesses of these models in terms of precision, recall, and F1-score for the task of fake news detection on this specific dataset?

To address these questions, this paper presents an empirical evaluation of the aforementioned machine learning models using the "Fake News Detection: The Battle Against Misinformation" dataset. The primary objectives of this study are:

- **Objective 1:** To implement and train Naive Bayes, SVM, Logistic Regression, K-Nearest Neighbors, and SVC models for binary classification of news articles as "Fake" or "Real."
- **Objective 2:** To evaluate and compare the performance of these models using relevant evaluation metrics, including accuracy, precision, recall, and F1-score.
- **Objective 3:** To analyze the experimental results and discuss the implications of the findings for the task of fake news detection.

The contribution of this study lies in providing a comparative analysis of the performance of several widely used classical machine learning models on a specific, publicly available fake news dataset. The findings offer insights into the suitability of these models for this task and can serve as a benchmark for future research.

The remainder of this paper is structured as follows: Section II provides a comprehensive review of the existing literature on fake news detection and the application of machine learning techniques. Section III details the proposed methodology, including the dataset used and the experimental setup. Section IV presents the experimental results. Section V offers a discussion of the findings and their implications. Finally, Section VI concludes the paper and outlines potential directions for future research.

2. Literature Review

The proliferation of fake news in the digital age has spurred a significant body of research aimed at understanding its dynamics and developing effective detection strategies. This section provides a review of the relevant literature, encompassing the nature and impact of fake news, the application of machine learning techniques for its detection, commonly used datasets, evaluation metrics, and the inherent challenges in this field.

Fake news, characterized by intentionally false or misleading information presented as legitimate news, has emerged as a critical societal concern. Its rapid spread through social media platforms and online news outlets can have far-reaching consequences, influencing public opinion on political matters, eroding trust in journalistic integrity, and even posing risks to public health by disseminating misinformation about medical issues[15]. The motivations behind the creation and dissemination of fake news are varied, ranging from financial gain through clickbait to deliberate attempts at political manipulation and social disruption [16]. Understanding the multifaceted nature of fake news, including its various forms and the psychological factors that contribute to its acceptance and sharing, is crucial for developing effective detection mechanisms [17], [18].

Machine learning has become a central pillar in the fight against fake news, offering the ability to analyze large volumes of text and identify patterns indicative of deception. Researchers have explored a wide array of algorithmic approaches, leveraging the power of Natural Language Processing (NLP) to extract meaningful features from news articles.

Several traditional machine learning models have been extensively applied to the task of fake news classification. Naïve Bayes, based on Bayes' theorem, assumes independence between features and has proven effective in various text classification tasks, including identifying spam and classifying documents [19]. Its simplicity and efficiency make it suitable for large datasets. Support Vector Machines (SVMs) are powerful classifiers that aim to find the optimal hyperplane to separate data points of different classes [20]. SVMs excel in high-dimensional spaces and have been shown to be effective in distinguishing between real and fake news based on textual features. Random Forest, an ensemble learning method, constructs multiple decision trees and aggregates their predictions, often leading to improved accuracy and robustness compared to individual decision trees [21] [22]. Its ability to handle noisy data and identify important features makes it valuable for fake news detection. Logistic Regression, a linear model, estimates the probability of a binary outcome (e.g., fake or real) based on input features. Despite its linearity, it can perform surprisingly well in text classification tasks, especially when combined with effective feature engineering. Gradient Boosting is another ensemble technique that builds a strong predictive model by iteratively combining weak learners, often achieving high accuracy in various classification problems, including fake news detection.

More recent research has explored the use of advanced machine learning models, particularly deep learning techniques, to capture more complex patterns in text data. Convolutional Neural Networks (CNNs), originally designed for image processing, have been adapted for NLP tasks, including fake news detection, by their ability to learn hierarchical representations of text and identify local patterns. Recurrent Neural Networks (RNNs), and their more sophisticated variants like Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRUs), are designed to process sequential data and can capture long-range dependencies in text, making them suitable for understanding the context and flow of news articles [23]. Furthermore, Transformer-based models, such as BERT (Bidirectional Encoder Representations from Transformers), have achieved state-of-the-art results in various NLP tasks due to their ability to learn contextualized word embeddings and capture intricate semantic relationships within text. While your initial write-up doesn't explicitly mention BERT, it's a significant advancement in the field.

The performance of machine learning models heavily relies on the quality of the features extracted from the text data. Various techniques are employed to transform raw text into numerical representations that can be fed into the models. Word embeddings, such as Word2Vec, GloVe, Doc2Vec, and the contextual embeddings from models like BERT, capture the semantic meaning of words and documents, allowing models to understand the nuances of language [24], [25]. Traditional methods like Bag-of-Words (BoW) and TF-IDF

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.03 No.02 (2025)

(Term Frequency-Inverse Document Frequency) represent text based on the frequency of words, although they lack semantic information [26]. Beyond lexical features, researchers have also explored stylistic features (e.g., sentiment, readability scores), syntactic features (e.g., part-of-speech tags), and even network-based features that consider the spread of information on social media [27]. Preprocessing steps, such as text cleaning (removing punctuation, special characters), tokenization (splitting text into words), and stemming or lemmatization (reducing words to their root form), are crucial for preparing the data for feature extraction.

The availability of high-quality, labeled datasets is essential for training and evaluating fake news detection models. Several datasets have been created for this purpose. The ISOT Fake News Dataset [28] provides a large collection of real and fake news articles across various topics. The LIAR dataset [29] consists of short, manually labeled statements from PolitiFact, offering a different perspective on misinformation. Multimodal datasets like Fakeddit [30][39] and FakeNewsNet include both textual and visual information, reflecting the multimedia nature of online news. Each dataset has its own characteristics in terms of size, topic diversity, and labeling process, which can influence the performance of the models trained on them. Researchers often employ data augmentation techniques to increase the size and diversity of training data and address issues of class imbalance [31].

The performance of fake news detection models is typically evaluated using standard classification metrics such as accuracy (the overall percentage of correctly classified instances), precision (the proportion of correctly identified fake news articles out of all articles classified as fake), recall (the proportion of correctly identified fake news articles out of all actual fake news articles), and the F1-score (the harmonic mean of precision and recall) [32]. For imbalanced datasets, where the number of real and fake news articles differs significantly, precision, recall, and F1-score provide a more nuanced understanding of the model's performance than accuracy alone. The AUC-ROC is another important metric that assesses the model's ability to distinguish between the two classes across different classification thresholds.

Despite the progress made, fake news detection remains a challenging task due to several factors [33]. The rapid and widespread dissemination of information on social media platforms makes it difficult to contain the spread of misinformation before it reaches a large audience . The dynamic and evolving nature of fake news content, with creators constantly adapting their strategies to evade detection, requires models to be robust and adaptable [34]. Data quality and bias in training datasets can lead to models that perform poorly on unseen data or exhibit unfairness towards certain types of news or sources[34]. The reliance on specific linguistic features can be circumvented by sophisticated manipulation techniques. Furthermore, incorporating user engagement and contextual information to improve detection accuracy is an ongoing area of research[34]. The effective deployment of advanced machine learning models in real-world scenarios also presents technological and scalability challenges [35].

This literature review provides a foundation for understanding the landscape of fake news detection and the role of machine learning in addressing this critical challenge. The subsequent sections of this paper will detail our proposed methodology, experimental results, and discussion of the findings in the context of existing research.

3. Methodology

Dataset Description:

The dataset utilized in this study is the "Fake News Detection: The Battle Against Misinformation" dataset. This dataset is specifically designed for research in fake news

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.03 No.02 (2025)

detection and provides a balanced collection of news articles labeled as either "Fake" or "Real." As detailed in the dataset description, it contains 5000 "Fake" and 4900 "Real" thousands of news items with a binary label.

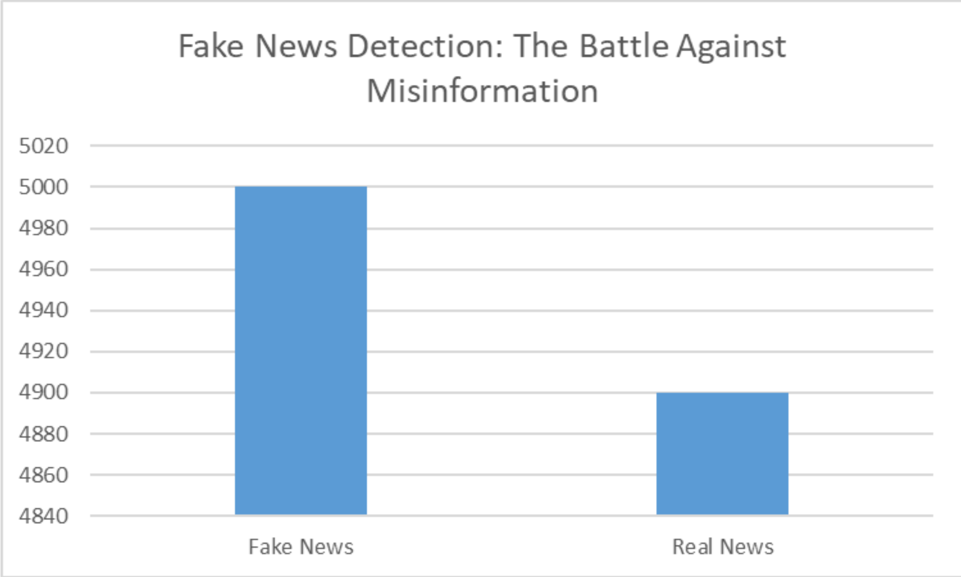

Figure 1: *Fake News Detection: The Battle Against Misinformation Dataset*

The dataset consists of two primary fields:

- **Text:** This field contains the full text of the news article, encompassing the main body of the content.
- **Label:** This categorical field indicates the authenticity of the news article, with values of "Fake" or "Real."

This near-balanced class distribution helps mitigate potential biases that can arise from significantly imbalanced datasets.

Data Preprocessing:

The text preprocessing pipeline involved several key steps to prepare the data for machine learning models. First, text cleaning was performed, which involved removing punctuation, special characters, and any potential HTML tags or URLs. Following this, all text was converted to lowercase to ensure consistency across the dataset. Next, tokenization was applied, splitting the text into individual words or tokens. To further refine the data, stop word removal was carried out, eliminating common words such as "the," "a," and "is," which typically do not add significant meaning. After this, stemming or lemmatization was utilized to reduce words to their root form—such as converting "running," "ran," and "runs" to "run". Once the text was cleaned and normalized, feature extraction techniques were applied to convert the textual data into numerical representations suitable for machine learning models. One of the primary methods used was TF-IDF (Term Frequency-Inverse Document Frequency), which assigns weights to words based on their frequency within a document and across the entire corpus.

Machine Learning Models:

In this study, several traditional machine learning models were implemented and evaluated. Naive Bayes was utilized as a probabilistic classifier based on Bayes' theorem, assuming independence between features. Support Vector Machine (SVM), a powerful classifier that determines the optimal hyperplane to separate data points of different classes, was also employed. We used Linear kernel SVM implementation from scikit-learn, In addition, Logistic Regression was applied, a linear model that leverages a sigmoid function to predict the probability of binary outcomes. We utilized the Logistic Regression implementation from scikit-learn, incorporating. K-Nearest Neighbors (KNN), a non-parametric algorithm that classifies data points based on the majority class of their 'k' nearest neighbors, was also tested. We used the K-Neighbors Classifier from scikit-learn. Finally, we included the Support Vector Classifier (SVC), another implementation of Support Vector Machines available in scikit-learn.

4. Experimental Setup:

The experimental procedure involved the following steps:

Data Splitting: The dataset was split into training and testing sets to evaluate the generalization performance of the models on unseen data. We used an 80% for training and 20% for testing split, resulting in training samples and testing samples.

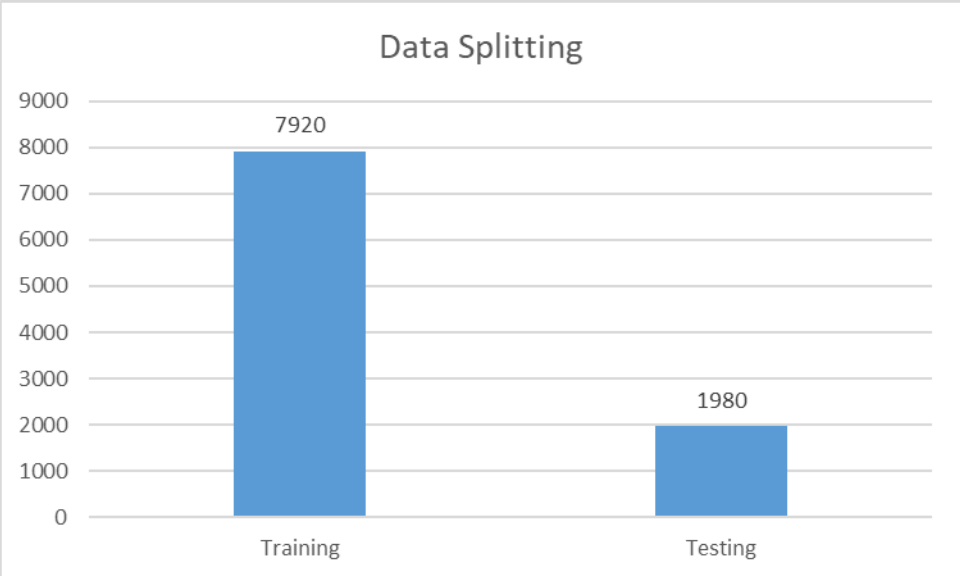

Figure 2: *Data Splitting*

Each of the five machine learning models was trained on the training dataset using the extracted TF-IDF features.

The trained models were evaluated on a held-out testing dataset to measure their performance. Several evaluation metrics were employed to provide a comprehensive assessment. Accuracy was calculated as the proportion of correctly classified instances out of the total number of instances. Precision was measured for each class (Fake and Real) as the ratio of correctly predicted instances to the total number of instances predicted as that class. Recall was determined as the ratio of correctly predicted instances to the total number of actual instances of each class. Finally, the F1-score was computed as the harmonic mean of precision and recall for each class, offering a balanced metric that accounts for both false positives and false negatives.

The experiments were conducted using Python with the scikit-learn library and many others on a Kaggle platform.

EXPERIMENTAL RESULTS

This section presents the results obtained from the experiments conducted using the five machine learning models on the fake news detection dataset. The performance of each model is reported based on the evaluation metrics described in the previous section.

Naive Bayes:

The classification report for the Naive Bayes model on the testing dataset is as follows:

Table 1: Naive Bayes Results

	Precision	**Recall**	**F1-Score**	**Support**
0	0.95	0.94	0.95	973

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.03 No.02 (2025)

1	0.94	0.96	0.95	1007
Accuracy			0.95	1980
Macro avg	0.95	0.95	0.95	1980
Weighted Avg	0.95	0.95	0.95	1980

The confusion matrix shown in the given below image summarizes the performance of the classification model. It indicates that 914 real instances and 966 fake instances were correctly classified, with minimal misclassifications. The low number of false positives (61) and false negatives (59) reflects the model's strong predictive ability.

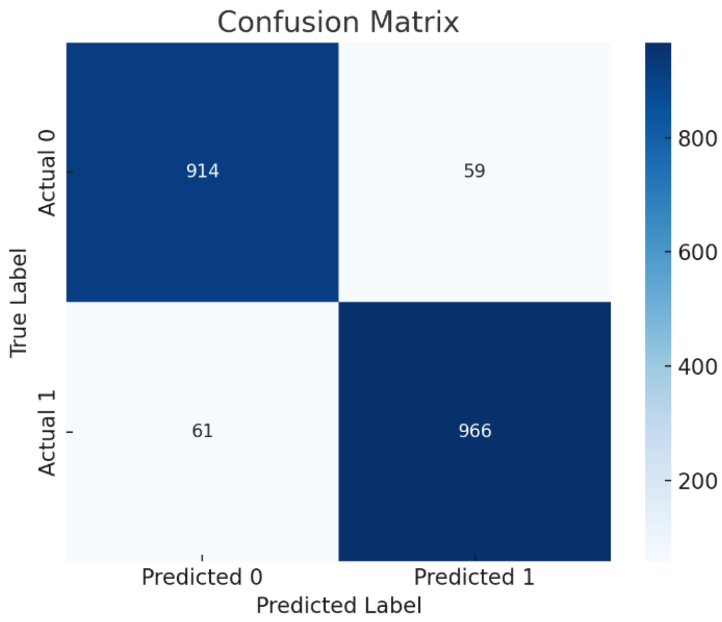

Figure 3: Naive Bayes Confusion Matrix

Support Vector Machine (SVM):
The performance of the SVM model on the testing dataset is summarized below:

Table 2 : Support Vector Machine (SVM) Results

	Precision	**Recall**	**F1-Score**	**Support**
0	0.99	1.00	0.99	1000
1	1.00	0.99	0.99	980
Accuracy			0.99	1980
Macro avg	0.99	0.99	0.99	1980
Weighted avg	0.99	0.99	0.99	1980

The classification report shows that the SVM model achieved very high performance, with an overall accuracy and macro F1-score of 0.9939. Both classes (Fake and Real) were classified with excellent precision and recall, indicating a highly effective and balanced model.

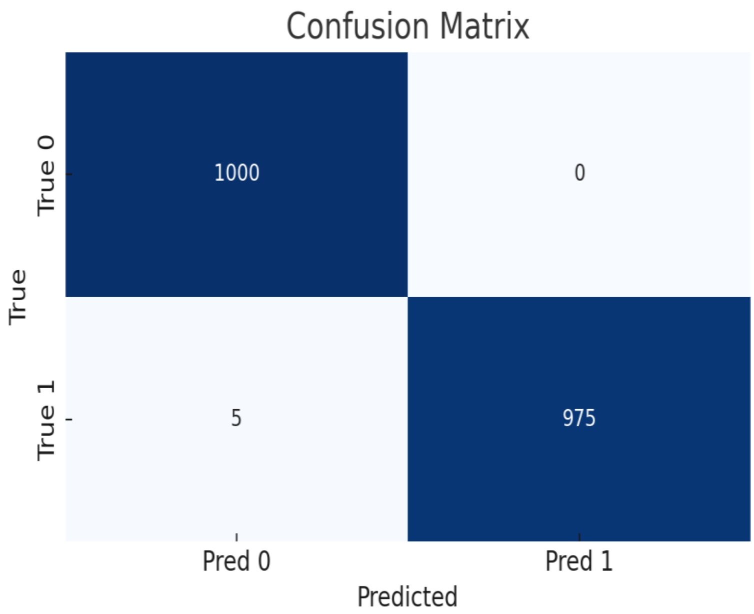

Figure 4: *Support Vector Machine (SVM) Confusion Matrix*

Logistic Regression:
The classification report Excel file provides a detailed performance summary of the Logistic Regression model. It outlines key evaluation metrics including precision, recall, F1-score, and support for each class (0 and 1). Additionally, it records the overall model accuracy, macro average, and weighted average, giving a comprehensive view of the model's effectiveness.

Table 3: Logistic Regression Results

Class	**Precision**	**Recall**	**F1-Score**	**Support**
0	0.99	0.99	0.99	996
1	0.99	0.99	0.99	984
accuracy			0.99	1980
macro avg	0.99	0.99	0.99	1980
weighted avg	0.99	0.99	0.99	1980

The confusion matrix image visually represents the performance of the Logistic Regression model. It illustrates the number of correct and incorrect predictions for each class, helping to easily identify the model's strengths and areas for improvement. By analyzing this matrix, one can better understand the classification results and any misclassifications made by the model.

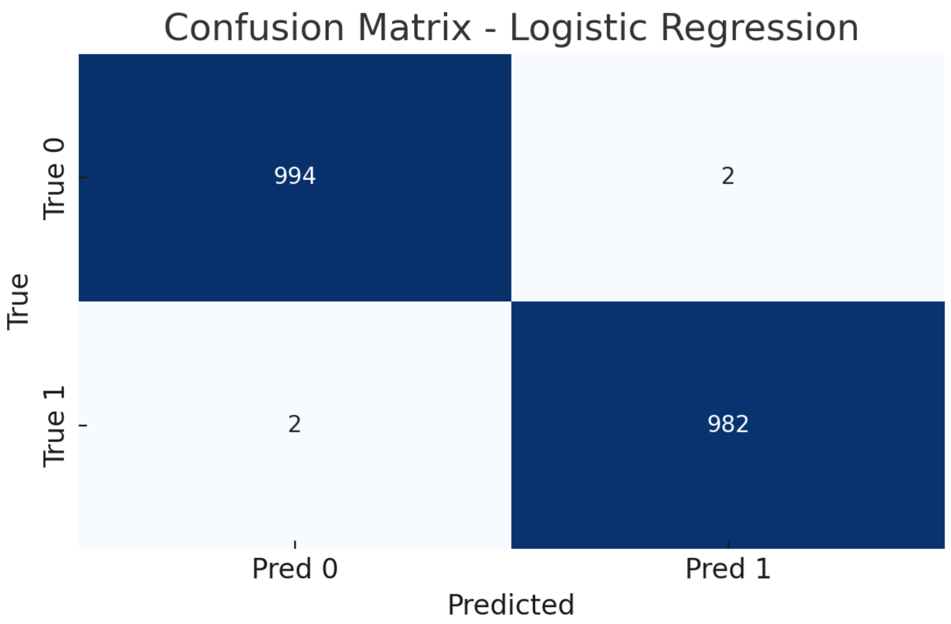

Figure 5: Logistic Regression Confusion Matrix

K-Nearest Neighbors (KNN):

The classification report Excel file presents the performance metrics of the K-Nearest Neighbors (KNN) model. It highlights the precision, recall, F1-score, and support for both classes (0 and 1) in the testing dataset. The overall model accuracy, along with macro and weighted averages, is also included to provide a complete overview of the model's behavior.

Table 4 : K-Nearest Neighbors (KNN) Results

	Precision	Recall	F1-score	Support
0	0.94	0.75	0.84	996
1	0.79	0.95	0.87	984
Accuracy			0.85	1980
Macro Avg	0.87	0.85	0.85	1980
Weighted Avg	0.87	0.85	0.85	1980

The confusion matrix image visually represents the classification results of the KNN model. It details the number of correct and incorrect predictions for each class, clearly illustrating the model's tendency to favor certain predictions. This matrix helps in assessing both the strengths and weaknesses of the KNN model.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR

CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW

Vol.03 No.02 (2025)

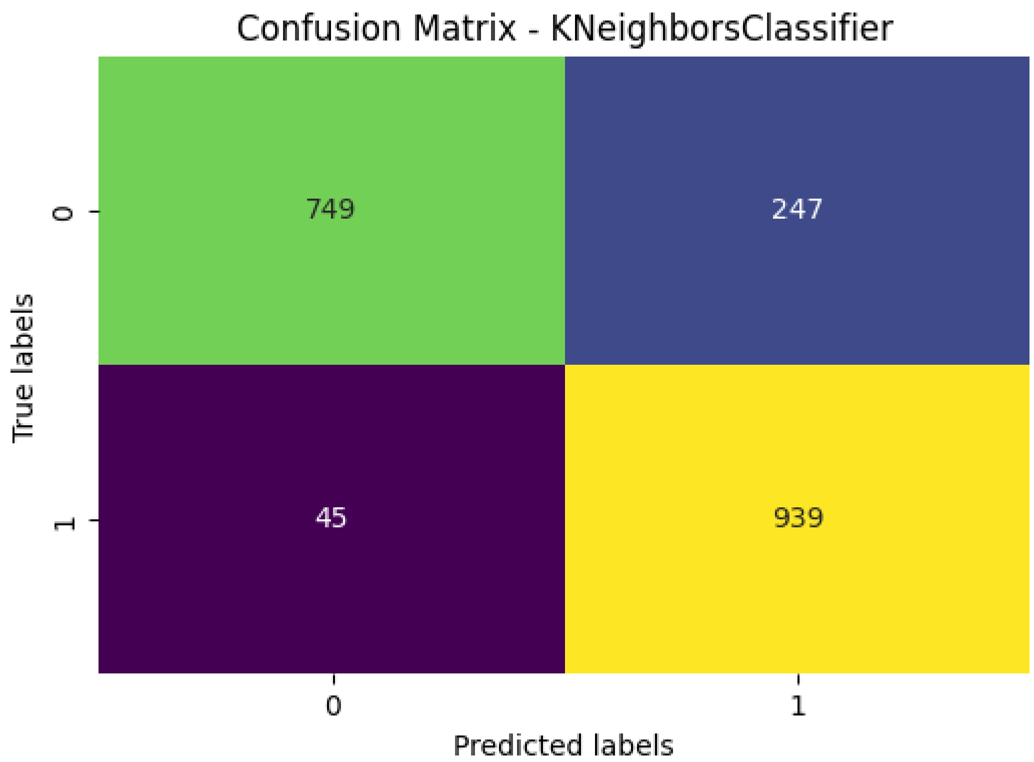

Figure 6: K-Nearest Neighbors (KNN) Confusion Matrix

Support Vector Classifier (SVC):

The classification report Excel file captures the evaluation results of the Support Vector Classifier (SVC) model. It lists precision, recall, F1-score, and support values for each class, reflecting the model's highly consistent and accurate performance. Overall metrics like accuracy, macro average, and weighted average are also provided for a complete performance summary.

Table 5: Support Vector Classifier (SVC) Results

	Precision	**Recall**	**F1-score**	**Support**
0	0.99	0.99	0.99	996
1	0.99	0.99	0.99	984
Accuracy			0.99	1980
Macro Avg	0.99	0.99	0.99	1980
Weighted Avg	0.99	0.99	0.99	1980

The confusion matrix for the SVC model graphically displays the classification outcomes. It shows how accurately the model classified each class, with minimal misclassifications. This matrix offers a clear and effective way to understand the excellent predictive ability of the SVC model.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR

CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW

Vol.03 No.02 (2025)

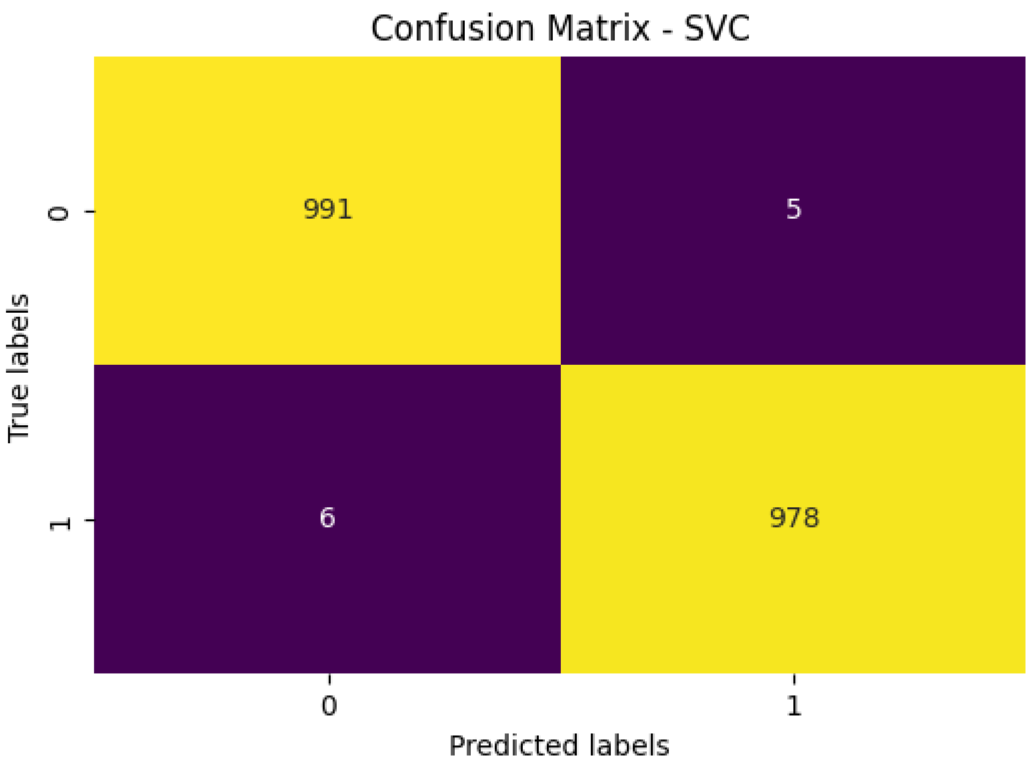

Figure 7: Support Vector Classifier (SVC) Confusion Matrix

Comparative Summary:

Model	Accuracy	Macro Avg F1-Score	Class 0 F1-Score	Class 1 F1-Score
Naive Bayes	0.95	0.95	0.95	0.95
Support Vector Machine (SVM)	0.9939	0.9939	0.99	0.99
Logistic Regression	0.99	0.99	0.99	0.99
K-Nearest Neighbors (KNN)	0.85	0.85	0.84	0.87
Support Vector Classifier (SVC)	0.99	0.99	0.99	0.99

The comparative analysis reveals a clear distinction in performance among the evaluated models. Support Vector Machine (SVM), Logistic Regression, and Support Vector Classifier (SVC) demonstrate superior accuracy, all achieving approximately 99%, with consistently high F1-scores across both classes, indicating their effectiveness in accurately classifying fake and real news. Naive Bayes presents a competitive performance with an accuracy of 95% and balanced F1-scores, showcasing its robustness in text classification. In contrast, K-Nearest Neighbors (KNN) exhibits a notable decrease in accuracy to 85%, accompanied by lower F1-scores, suggesting its relative inadequacy for this specific task compared to the other models.

5. RESULTS DISCUSSION

The experimental results presented in the previous section provide a comparative evaluation of five traditional machine learning models for the task of fake news detection on the "Fake News Detection: The Battle Against Misinformation" dataset.

The **Naive Bayes** model achieved a respectable accuracy of 95%, with balanced precision, recall, and F1-scores for both the "Fake" and "Real" news classes. This demonstrates the effectiveness of this probabilistic model for text classification, particularly given its simplicity and efficiency.

The **Support Vector Machine (SVM)**, **Logistic Regression**, and **Support Vector Classifier (SVC)** models exhibited remarkably high performance, all achieving an accuracy of approximately 99%. The precision, recall, and F1-scores for both classes were also consistently high (around 0.99), indicating that these models are highly effective in correctly identifying both fake and real news articles within this dataset. The ability of SVM and SVC to find optimal separating hyperplanes in high-dimensional feature spaces, coupled with the capacity of Logistic Regression to model the probability of class membership effectively, likely contributes to their strong performance.

In contrast, the **K-Nearest Neighbors (KNN)** model showed a comparatively lower accuracy of 85%. While the recall for the "Real" news class was relatively high (0.95), the precision for this class (0.79) and the overall performance metrics suggest that KNN is less effective than the other models for this specific task and dataset. The performance of KNN can be sensitive to the choice of 'k' and the distance metric, and it might struggle with high-dimensional text data where the notion of "nearest neighbors" becomes less clear.

The high accuracy achieved by SVM, Logistic Regression, and SVC highlights the potential of classical machine learning techniques for building effective fake news detection systems. Future research could explore the application of more advanced techniques, such as deep learning models and transformer networks, to further enhance detection accuracy and robustness. Additionally, investigating the impact of different feature engineering methods and addressing the challenges of data bias and the evolving nature of fake news remain crucial areas for future work.

6. Conclusion

This study has presented an empirical evaluation of five traditional machine learning models – Naive Bayes, Support Vector Machine (SVM), Logistic Regression, K-Nearest Neighbors, and Support Vector Classifier (SVC) – for the task of fake news detection using the "Fake News Detection: The Battle Against Misinformation" dataset. The experimental results demonstrate that SVM, Logistic Regression, and SVC achieve superior performance, with accuracy rates of approximately 99%, indicating their strong capability in distinguishing between real and fake news articles within this dataset. The Naive Bayes model also showed competitive performance with an accuracy of 95%. In contrast, the K-Nearest Neighbors model exhibited lower accuracy, suggesting it may not be as well-suited for this specific text classification task.The findings of this study contribute to the growing body of research on fake news detection by providing a comparative analysis of widely used classical machine learning algorithms on a relevant dataset. The high performance of SVM, Logistic Regression, and SVC underscores the effectiveness of these techniques as a baseline for future investigations. Future research should focus on exploring more advanced machine learning models, such as deep learning architectures, and investigating the impact of sophisticated feature engineering techniques to further improve the accuracy and robustness of fake news detection systems. Addressing the challenges of data bias, the evolving nature of fake news, and the interpretability of detection models are also critical directions for future work in this important and rapidly evolving field.

References

[1] Jabeen, Tayyba, Yasir Mehmood, Hamayun Khan, Muhammad Fawad Nasim, and Syed Asad Ali Naqvi. "Identity Theft and Data Breaches How Stolen Data Circulates on the Dark Web: A Systematic Approach." Spectrum of engineering sciences 3, no. 1 (2025): 143-161.

[2] I. Samadov, “Fake News Detection. Challenges and Machine Learning Models to Combat Disinformation,” Medium. Accessed: Apr. 28, 2025. [Online]. Available: https://ismatsamadov.medium.com/fake-news-detection-challenges-and-machine-learning-models-to-combat-disinformation-775864f84940

[3] Tariq, Muhammad Arham, Muhammad Ismaeel Khan, Aftab Arif, Muhammad Aksam Iftikhar, and Ali Raza A. Khan. "Malware Images Visualization and Classification With Parameter Tunned Deep Learning Model." Metallurgical and Materials Engineering 31, no. 2 (2025): 68-73.https://doi.org/10.63278/1336.

[4] Arif, A., A. Khan, and M. I. Khan. "Role of AI in Predicting and Mitigating Threats: A Comprehensive Review." JURIHUM: Jurnal Inovasi dan Humaniora 2, no. 3 (2024): 297-311.

[5] Khan, M. I., A. Arif, and A. R. A. Khan. "AI-Driven Threat Detection: A Brief Overview of AI Techniques in Cybersecurity." BIN: Bulletin of Informatics 2, no. 2 (2024): 248-61.

[6] Arif, Aftab, Muhammad Ismaeel Khan, and Ali Raza A. Khan. "An overview of cyber threats generated by AI." International Journal of Multidisciplinary Sciences and Arts 3, no. 4 (2024): 67-76.

[7] Zainab, Hira, Ali Raza A. Khan, Muhammad Ismaeel Khan, and Aftab Arif. "Ethical Considerations and Data Privacy Challenges in AI-Powered Healthcare Solutions for Cancer and Cardiovascular Diseases." Global Trends in Science and Technology 1, no. 1 (2025): 63-74.

[8] Zainab, Hira, Muhammad Ismaeel Khan, Aftab Arif, and Ali Raza A. Khan. "Deep Learning in Precision Nutrition: Tailoring Diet Plans Based on Genetic and Microbiome Data." Global Journal of Computer Sciences and Artificial Intelligence 1, no. 1 (2025): 31-42.

[9] Khan, Muhammad Ismaeel, Aftab Arif, and Ali Raza A. Khan. "AI's Revolutionary Role in Cyber Defense and Social Engineering." International Journal of Multidisciplinary Sciences and Arts 3, no. 4 (2024): 57-66.

[10] Khan, Ali Raza A., Muhammad Ismaeel Khan, and Aftab Arif. "AI in Surgical Robotics: Advancing Precision and Minimizing Human Error." Global Journal of Computer Sciences and Artificial Intelligence 1, no. 1 (2025): 17-30.

[11] Zainab, Hira, Ali Raza A. Khan, Muhammad Ismaeel Khan, and Aftab Arif. "Innovative AI Solutions for Mental Health: Bridging Detection and Therapy." Global Journal of Emerging AI and Computing 1, no. 1 (2025): 51-58.

[12] Zainab, Hira, Muhammad Ismaeel Khan, Aftab Arif, and Ali Raza A. Khan. "Development of Hybrid AI Models for Real-Time Cancer Diagnostics Using Multi-Modality Imaging (CT, MRI, PET)." Global Journal of Machine Learning and Computing 1, no. 1 (2025): 66-75.

[13] Khan, Muhammad Ismaeel, Aftab Arif, and Ali Raza A. Khan. "The Most Recent Advances and Uses of AI in Cybersecurity." BULLET: Jurnal Multidisiplin Ilmu 3, no. 4 (2024): 566-578.

[14]Khan, Shaz Mumtaz, Fawad Nasim, Jawad Ahmad, and Sohail Masood. "Deep Learning-Based Brain Tumor Detection." Journal of Computing & Biomedical Informatics 7, no. 02 (2024).

[15] "Fake news detection: recent trends and challenges | Social Network Analysis and Mining." Accessed: Apr. 28, 2025. [Online]. Available: https://link.springer.com/article/10.1007/s13278-024-01344-4

[16] "How Deepfakes Are Impacting Public Trust in Media | Pindrop." Accessed: Apr. 28, 2025. [Online]. Available: https://www.pindrop.com/article/deepfakes-impacting-trust-media/

[17] H. Padalko, V. Chomko, and D. Chumachenko, "A novel approach to fake news classification using LSTM-based deep learning models," *Front. Big Data*, vol. 6, p. 1320800, Jan. 2024, doi: 10.3389/fdata.2023.1320800.

[18] A. E. Eltoukhy, "An Advanced Framework for Fake News Detection: Integrating Feature Extraction and Optimal Selection to Combat Misinformation," *J. Electr. Syst.*, vol. 20, no. 3, Art. no. 3, Aug. 2024.

[19] "Addressing Algorithms in Disinformation." Accessed: Apr. 28, 2025. [Online]. Available: https://crestresearch.ac.uk/comment/addressing-algorithms-in-disinformation/

[20] "Fake News Datasets | MediaFutures." Accessed: Apr. 28, 2025. [Online]. Available: https://mediafutureseu.github.io/fakenewsdatasets.html

[21] "Papers with Code - Machine Learning Datasets." Accessed: Apr. 28, 2025. [Online]. Available: https://paperswithcode.com/datasets?task=fake-news-detection&page=1

[22] "Papers with Code - Machine Learning Datasets." Accessed: Apr. 28, 2025. [Online]. Available: https://paperswithcode.com/datasets?task=fake-news-detection

[23] D. Mouratidis, A. Kanavos, and K. Kermanidis, "From Misinformation to Insight: Machine Learning Strategies for Fake News Detection," *Information*, vol. 16, no. 3, Art. no. 3, Mar. 2025, doi: 10.3390/info16030189.

[24] S. K. Hamed, M. J. Ab Aziz, and M. R. Yaakub, "A review of fake news detection approaches: A critical analysis of relevant studies and highlighting key challenges associated with the dataset, feature representation, and data fusion," *Heliyon*, vol. 9, no. 10, p. e20382, Sep. 2023, doi: 10.1016/j.heliyon.2023.e20382.

[25] "A Review of Image and Text Feature Extraction Methods in Fake News Detection Tasks | IIETA." Accessed: Apr. 28, 2025. [Online]. Available: https://www.iieta.org/journals/isi/paper/10.18280/isi.290202

[26] "Fairness issues, current approaches, and challenges in machine learning models | International Journal of Machine Learning and Cybernetics." Accessed: Apr. 28, 2025. [Online]. Available: https://link.springer.com/article/10.1007/s13042-023-02083-2

[27] Z. Boukouvalas and A. Shafer, "Role of Statistics in Detecting Misinformation: A Review of the State of the Art, Open Issues, and Future Research Directions," *Annu. Rev. Stat. Its Appl.*, vol. 11, no. 1, pp. 27–50, Apr. 2024, doi: 10.1146/annurev-statistics-040622-033806.

[28] S. Maham, A. Tariq, M. U. G. Khan, F. S. Alamri, A. Rehman, and T. Saba, "ANN: adversarial news net for robust fake news classification," *Sci. Rep.*, vol. 14, no. 1, p. 7897, Apr. 2024, doi: 10.1038/s41598-024-56567-4.

[29] "A novel approach for assessing fairness in deployed machine learning algorithms | Scientific Reports." Accessed: Apr. 28, 2025. [Online]. Available: https://www.nature.com/articles/s41598-024-68651-w

[30] A. Galli, E. Masciari, V. Moscato, and G. Sperlí, "A comprehensive Benchmark for fake news detection," *J. Intell. Inf. Syst.*, vol. 59, no. 1, pp. 237–261, 2022, doi: 10.1007/s10844-021-00646-9.

[31] "Full article: Optimized Fake News Classification: Leveraging Ensembles Learning and Parameter Tuning in Machine and Deep Learning Methods." Accessed: Apr. 28, 2025. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/08839514.2024.2385856

[32] B. William, "Machine Learning Techniques for Fake News Detection," Mar. 05, 2025, *Preprints*: 2025030325. doi: 10.20944/preprints202503.0325.v1.

[33] "Real-time fake news detection in online social networks: FANDC Cloud-based system | Scientific Reports." Accessed: Apr. 28, 2025. [Online]. Available: https://www.nature.com/articles/s41598-024-76102-9

[34] F. Farhangian, R. M. O. Cruz, and G. D. C. Cavalcanti, "Fake news detection: Taxonomy and comparative study," *Inf. Fusion*, vol. 103, p. 102140, Mar. 2024, doi: 10.1016/j.inffus.2023.102140.

[35] V. Gunasekaran and D. Karthikeyan, "Real Time Fake news Detection Web App Enhanced by Machine Learning Algorithms," in *2024 4th International Conference on Intelligent Technologies (CONIT)*, Jun. 2024, pp. 1–7. doi: 10.1109/CONIT61985.2024.10627483.