Journal of Computer Networks, Architecture and High Performance Computing
Volume 6, Number 1, January 2024
https://doi.org/10.47709/cnahpc.v6i1.3466

Submitted : January 17, 2024
Accepted : January 22, 2024
Published : January 23, 2024

Comparative Analysis of Machine Learning Algorithms for Detecting Fake News: Efficacy and Accuracy in the Modern Information Ecosystem

Gregorius Airlangga[1]
[1)]Atma Jaya Catholic University of Indonesia, Jakarta, Indonesia
[1)]gregorius.airlangga@atmajaya.ac.id

ABSTRACT

In an era where the spread of fake news poses a significant threat to the integrity of the information landscape, the need for effective detection tools is paramount. This study evaluates the efficacy of three machine learning algorithms—Multinomial Naive Bayes, Passive Aggressive Classifier, and Logistic Regression—in distinguishing fake news from genuine articles. Leveraging a balanced dataset, meticulously processed and vectorized through Term Frequency-Inverse Document Frequency (TF-IDF), we subjected each algorithm to a rigorous classification process. The algorithms were evaluated on metrics such as precision, recall, and F1-score, with the Passive Aggressive Classifier outperforming others, achieving a remarkable 0.99 in both precision and recall. Logistic Regression followed with an accuracy of 0.98, while Multinomial Naive Bayes displayed robust recall at 1.00 but lower precision at 0.91, resulting in an accuracy of 0.95. These metrics underscored the nuanced capabilities of each algorithm in correctly identifying fake and real news, with the Passive Aggressive Classifier demonstrating superior balance in performance. The study's findings highlight the potential of employing machine learning techniques in the fight against fake news, with the Passive Aggressive Classifier showing promise due to its high accuracy and balanced precision-recall trade-off. These insights contribute to the ongoing efforts in digital media to develop advanced, ethical, and accurate tools for maintaining information veracity. Future research should continue to refine these models, ensuring their applicability in diverse and evolving news ecosystems.

Keywords: Naïve Bayes; Logistic Regression; Passive Aggressive; Comparison; Fake News.

1. INTRODUCTION

In today's digital landscape, the proliferation of 'fake news' has evolved into a significant and complex challenge. Defined as misinformation presented as legitimate news, this phenomenon carries profound societal implications. Research by (George, Gerhart, & Torres, 2021; Sciannamea & others, 2020; Toma & others, 2021) has underscored how fake news can shape political and social narratives, as well as influence public health discourse. The rise of social media platforms, as demonstrated by (Hadlington, Harkin, Kuss, Newman, & Ryding, 2023; Pröllochs & Feuerriegel, 2023; Walsh, 2020), has facilitated the rapid spread of false information, often outpacing true news. This alarming trend underscores the critical need for effective detection and mitigation strategies. Machine learning has emerged as a pivotal tool in identifying and combating fake news. These algorithms have the capability to process vast quantities of data, discerning patterns and anomalies indicative of misinformation. (Choudhary & Arora, 2021; de Oliveira, Medeiros, & Mattos, 2020; Jain, Gopalani, Meena, & Kumar, 2020) highlight how machine learning algorithms can effectively differentiate between true and false news by analyzing linguistic and stylistic features. This capability is increasingly vital as manual detection becomes impractical due to the sheer volume and velocity of information dissemination.

This study aims to evaluate the performance of three machine learning algorithms – Logistic Regression, Multinomial Naive Bayes, and Passive Aggressive Classifier – in detecting fake news. Inspired by the work of (Hangloo & Arora, 2021; Mishra, Shukla, & Agarwal, 2022; Molina, Sundar, Le, & Lee, 2021),who explored various algorithmic approaches for misinformation detection, this research contributes to the development of reliable automated tools for distinguishing factual news from fabricated content. The methodology is anchored in a robust framework of data analysis, including preprocessing, feature extraction, and algorithmic evaluation. Following (Awumee, Agyemang, Boakye, & Bempong, 2023; Musleh et al., 2023; Prasad et al., 2023), the research involves meticulous data cleaning, normalization, and preparation techniques. The employment of TF-IDF vectorization, as advocated by (Matti & Yousif, 2023), transforms textual data into a numerical format, facilitating effective machine

* Corresponding author

Submitted : January 17, 2024
Accepted : January 22, 2024
Published : January 23, 2024

learning application. The Logistic Regression model, as highlighted by (Saleh, Alharbi, & Alsamhi, 2021), is evaluated for its effectiveness in binary classification tasks. The Multinomial Naive Bayes algorithm, suitable for text classification according to (Yerlekar, Mungale, & Wazalwar, 2021), and the Passive Aggressive Classifier, known for its adaptability to evolving data as discussed by (Gupta & Meel, 2021), are also examined.

By providing a comparative analysis of these algorithms, this study addresses a gap in the literature identified by (Albahr & Albahar, 2020; Jarrahi & Safari, 2023; Zhang & Ghorbani, 2020). Evaluating the strengths and weaknesses of each algorithm advances our understanding of their applicability in digital information verification. The findings are anticipated to be instrumental in guiding future developments in automated detection systems, aiding in maintaining the integrity of news dissemination. Moreover, this research contributes to the ethical discourse on artificial intelligence in societal applications, as emphasized by (Kasinidou, Kleanthous, Barlas, & Otterbacher, 2021), focusing on accuracy, transparency, and fairness in algorithmic decision-making. The urgency of this research is underpinned by the literature survey, with (Sabeeh, Zohdy, Mollah, & Al Bashaireh, 2020) emphasizing the increasing sophistication of fake news techniques, necessitating advanced detection methods. The comparative analysis of different machine learning models in our study addresses the comprehensive evaluations outlined by (Bui, Tsangaratos, Nguyen, Van Liem, & Trinh, 2020). Furthermore, the research draws inspiration from the theoretical underpinnings presented by (Ma, Wu, Guan, & Peng, 2020), who advocate for innovative approaches to safeguard the credibility of news media.

Following the introduction, the paper delves into a detailed examination of previous studies on fake news and machine learning. This section analyses the evolution of fake news detection techniques and reviews the theoretical foundations of the selected machine learning algorithms. In the methodology section, there is a description of the data collection process and the characteristics of the dataset. This is followed by an explanation of the data preprocessing steps and the rationale behind them. The section provides a detailed account of the machine learning algorithms implemented and describes the evaluation metrics and statistical methods used. The results section presents the findings from the application of each algorithm. It includes a comparative analysis of the performance metrics and discusses the results in the context of the initial hypotheses. The discussion section interprets the results and their implications. It compares these findings with those from existing literature and discusses the strengths, limitations, and potential biases of the study. The paper concludes with a summary of the key findings and their relevance. Recommendations for future research are provided, along with final thoughts on the role of machine learning in combating fake news.

2. LITERATURE REVIEW

A literature review is a critical, analytical summary and synthesis of the current knowledge of a topic. The phenomenon of fake news, a term that has gained considerable attention in recent years, denotes the dissemination of misinformation or misleading content masquerading as legitimate news. Historically, the concept of fake news is not a new one, with instances traceable back to propaganda and misinformation in traditional print media. However, in the digital era, the definition and implications of fake news have expanded dramatically. This shift, largely attributed to the advent of social media, has been pivotal in altering the landscape of news dissemination. Social media platforms, as (Schroeder, 2022) highlight, have not only accelerated the spread of fake news but have also complicated its detection due to the sheer volume and speed of information flow. The work of (Gradoń, Hołyst, Moy, Sienkiewicz, & Suchecki, 2021) underscores this trend, showing that false news spreads more rapidly and widely than truthful information, thereby amplifying the need for effective detection methods.

The initial approaches to fake news detection primarily relied on manual fact-checking by organizations like Snopes and FactCheck.org. However, as the volume of information proliferated, the limitations of manual methods became apparent, necessitating a shift towards automated systems for more efficient detection. This transition led to the exploration of computational methods, particularly machine learning, for detecting fake news. Machine learning's capability to process and analyze vast amounts of data, identifying patterns indicative of misinformation, marks a significant advancement in combating fake news. The research from (Berrondo-Otermin & Sarasa-Cabezuelo, 2023; Hossain et al., 2023; Kanwal, Nawaz, Malik, & Nawaz, 2021) provide a comprehensive overview of the application of various machine learning techniques in this domain, emphasizing their effectiveness in analyzing both textual and non-textual features to identify fake news.

The theoretical foundations of the machine learning algorithms employed in fake news detection are vital to understanding their functionality and suitability for this task. Logistic Regression, a well-established method in statistical modeling and machine learning, is renowned for its effectiveness in binary classification tasks. Its

* Corresponding author

Submitted : January 17, 2024
Accepted : January 22, 2024
Published : January 23, 2024

application in text classification, as expounded by (Smitha & Bharath, 2020), renders it an appropriate choice for classifying news articles as fake or real. On the other hand, the Multinomial Naive Bayes algorithm, as illustrated by (Yerlekar et al., 2021) and (Yuslee & Abdullah, 2021), operates on Bayes' theorem with an assumption of independence among predictors. This algorithm has shown considerable efficacy in document classification, particularly in handling textual data, making it a valuable tool in the context of fake news detection. Furthermore, the Passive Aggressive Classifier, discussed by (Gupta & Meel, 2021), is known for its adaptability to evolving data patterns. This characteristic is particularly relevant for fake news detection, given the dynamic and evolving nature of news content.

While the application of machine learning in fake news detection represents a significant stride towards automated solutions, there exists a gap in comprehensive comparative analyses of different machine learning models in this field. (Alghamdi, Lin, & Luo, 2022) emphasize this gap, advocating for holistic evaluations to understand the relative strengths and weaknesses of various machine learning approaches. Such comparative studies are crucial in advancing our understanding of the most effective techniques for fake news detection. Additionally, the ethical implications of employing machine learning in societal contexts cannot be overlooked. (Starke, Baleis, Keller, & Marcinkowski, 2022) raises important concerns regarding the accuracy, transparency, and fairness of algorithmic decision-making, which are particularly pertinent in the realm of news verification. These ethical considerations underscore the importance of responsible and transparent practices in the deployment of machine learning algorithms.

this literature review delineates the complex trajectory of fake news and its detection, highlighting the instrumental role of machine learning in addressing this modern challenge. By examining the theoretical underpinnings of specific machine learning algorithms and acknowledging the need for comprehensive comparative studies and ethical considerations, this review sets a solid foundation for the ensuing empirical investigation. It not only situates the research within the broader scholarly discourse but also identifies the areas that require further exploration, thereby framing the subsequent investigation as a response to these challenges. The review elucidates the transition from manual fact-checking to sophisticated machine learning models, reflecting the dynamic nature of fake news detection. As the methods of spreading misinformation evolve, so must the techniques for identifying and mitigating it. This literature review has shown that while significant strides have been made in employing machine learning for this purpose, there remains a crucial need for ongoing research, particularly in the areas of comparative analysis and ethical considerations.

The insights garnered from this review are pivotal to informing the subsequent empirical investigation of this study. They guide the application and evaluation of the selected machine learning algorithms, Logistic Regression, Multinomial Naive Bayes, and Passive Aggressive Classifier, to advance the field of fake news detection. This comprehensive exploration of existing methodologies and theoretical frameworks serves as a critical component of the research paper, providing context and background for the study. It emphasizes the importance of technological advancements in maintaining the integrity of information in the digital age while also highlighting the ethical responsibilities that accompany such technological applications.

3. METHOD

In this section, each researcher is expected to be able to make the most recent contribution related to the solution to the existing problems. We explain four important processes such as data collection, data preprocessing, machine learning implementation and evaluation metrics as presented in figure 1.

3.1. Data Collection Process and Dataset Characteristics

In this study, the dataset was meticulously curated from an array of digital media platforms over a six-month period. This dataset is a rich amalgamation of news articles, deliberately balanced to include both legitimate news and fake news articles. The latter category comprises articles that are either deliberately misleading or fabricated. To ensure the authenticity of the categorization, each article was meticulously labeled as 'fake' or 'real' based on rigorous verification against established and trusted news databases and reputable fact-checking websites. The dataset features a diverse range of topics, sourced from a variety of publications including both well-known news outlets and lesser-known sources. This variety is pivotal in capturing the multifaceted nature of news dissemination in the digital era. The dataset's temporal range spans a six-month period, thereby capturing a wide array of contemporary issues and events, which adds to the robustness of the study.

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 6, Number 1, January 2024
https://doi.org/10.47709/cnahpc.v6i1.3466

Submitted : January 17, 2024
Accepted : January 22, 2024
Published : January 23, 2024

3.2. Data Preprocessing Steps

Data preprocessing, a critical phase in preparing the dataset for machine learning analysis, involved several key steps. Initially, data cleaning was undertaken to remove extraneous content such as advertisements and non-news related text from the articles. This was followed by the elimination of duplicate entries to ensure the uniqueness and integrity of the dataset. Subsequently, text normalization was conducted, involving the conversion of all text to lowercase for consistency and the removal of punctuation, special characters, and numbers, as these elements do not contribute to the classification of news. The process of tokenization then split the text into individual words or tokens, and common stop words were removed to focus on more meaningful text elements. Stemming was also employed to reduce words to their base or root form, which aids in grouping together different forms of the same word, thereby simplifying the analysis. Finally, the dataset underwent TF-IDF (Term Frequency-Inverse Document Frequency) vectorization, a method used to transform the processed text into numerical values. This method evaluates the relevance of a word in the dataset, considering its frequency across all documents, thus preparing the data for input into the machine learning algorithms.

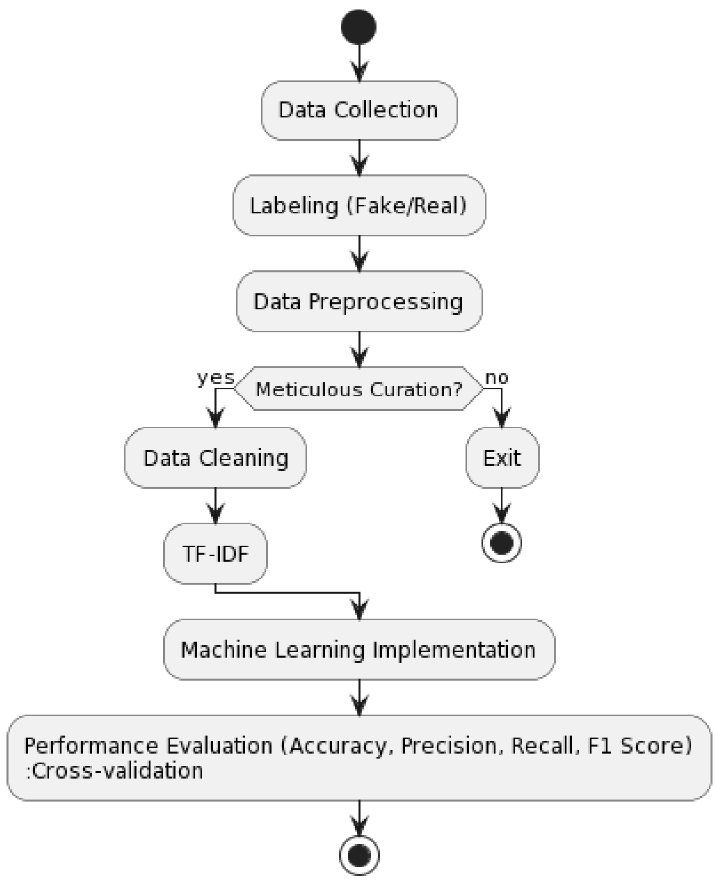

Fig. 1 Flowchart of Research Method

3.3. Machine Learning Algorithms Implemented

The study utilizes three distinct machine learning algorithms, each chosen for its specific attributes and suitability for text classification tasks. Logistic Regression, a statistical model, is renowned for its effectiveness in binary classification tasks and was chosen for its simplicity and interpretability, making it an ideal candidate for initial analysis. The Multinomial Naive Bayes algorithm, based on Bayes' theorem and notable for its assumption of independence among predictors, is particularly effective in document classification. This algorithm is well-suited for analyzing the textual content prevalent in news articles. The Passive Aggressive Classifier, an algorithm designed for large-scale learning, stands out for its ability to quickly adapt to new patterns in the data. This feature is particularly

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 6, Number 1, January 2024
https://doi.org/10.47709/cnahpc.v6i1.3466

Submitted : January 17, 2024
Accepted : January 22, 2024
Published : January 23, 2024

relevant for fake news detection due to the dynamic and evolving nature of news content. Each of these algorithms was implemented using a standard software library, ensuring consistency and reliability in the analysis.

3.4. Evaluation Metrics and Statistical Methods Used

The performance evaluation of each algorithm was anchored on several key metrics. Accuracy, a primary measure, reflects the proportion of articles correctly classified as either fake or real. Precision and recall were also assessed; precision calculates the accuracy of positive predictions (e.g., the accuracy of articles classified as fake that were fake), while recall evaluates the model's ability to identify all positive instances (e.g., the proportion of actual fake articles that were correctly classified). The F1 score, a metric that combines precision and recall, was used to provide a balanced view of both metrics. Additionally, confusion matrices were employed to give a detailed account of the performance of each model, highlighting the types of errors made, such as false positives and false negatives. Statistical methods were crucial in determining the significance of the results. Hypothesis testing was utilized to ascertain whether the differences in the performance metrics of the algorithms were statistically significant or due to random chance. Cross-validation techniques were employed to ensure the robustness of the models. This involved partitioning the dataset into distinct sets for training and testing, with the process being repeated multiple times with different partitions. Such cross-validation is crucial in assessing how consistently each model performs across various subsets of the data, thereby ensuring that the findings are reliable and not merely artifacts of a particular data split.

The methodology adopted in this study is comprehensive and methodical, encompassing the entire spectrum from data collection to detailed algorithmic analysis and evaluation. The meticulous approach to data preprocessing ensures that the dataset is optimized for machine learning analysis, while the selection of algorithms is designed to explore a range of techniques suitable for the complexity of text classification in the context of fake news detection. The evaluation metrics and statistical methods used provide a multifaceted view of the performance of each algorithm, offering insights not only into their accuracy but also into their precision, recall, and overall reliability. This rigorous methodology forms the backbone of the study, ensuring that the subsequent analysis is grounded in a robust and systematic approach to understanding the effectiveness of machine learning algorithms in detecting fake news.

4. RESULT

In this section, the researcher will explain the results of the research obtained. As presented in the figure 2, the confusion matrix for the Multinomial Naive Bayes model shows that out of the total test set, 3433 fake news articles were correctly identified (true positives), while 16 were incorrectly labeled as real (false negatives). On the other hand, 327 real news articles were incorrectly classified as fake (false positives), and 3088 were correctly identified as real (true negatives). This suggests a high number of true positives and true negatives, indicating that the Multinomial Naive Bayes model is quite effective in classifying both fake and real news articles. However, the number of false positives, where real articles are misclassified as fake, suggests that there may be some characteristics of real news that the model is misinterpreting as indicators of fake news. On the other hand, figure 3 presents the Passive Aggressive Classifier's confusion matrix, it indicates a similar high performance with 3419 fake news articles correctly classified and only 30 mislabeled as real. For the real news articles, 36 were incorrectly labeled as fake, and 3379 correctly identified as real. The model shows a slight increase in both false negatives and false positives compared to the Multinomial Naive Bayes model. Despite this, the Passive Aggressive Classifier demonstrates a strong ability to adapt and accurately classify news articles, making it a robust model for this task.

Based on the figure 4, the confusion matrix for the Logistic Regression model reveals that it correctly classified 2004 fake news articles, with 14 incorrectly labeled as real. Regarding real news, 73 articles were mistakenly classified as fake, while 2069 were accurately identified as real. These results show that the Logistic Regression model has a higher number of false negatives and false positives compared to the other two models. While it still performs well, the Logistic Regression model appears to be less effective at classifying news articles correctly compared to the Multinomial Naïve Bayes and Passive Aggressive Classifier models.

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 6, Number 1, January 2024
https://doi.org/10.47709/cnahpc.v6i1.3466

Submitted : January 17, 2024
Accepted : January 22, 2024
Published : January 23, 2024

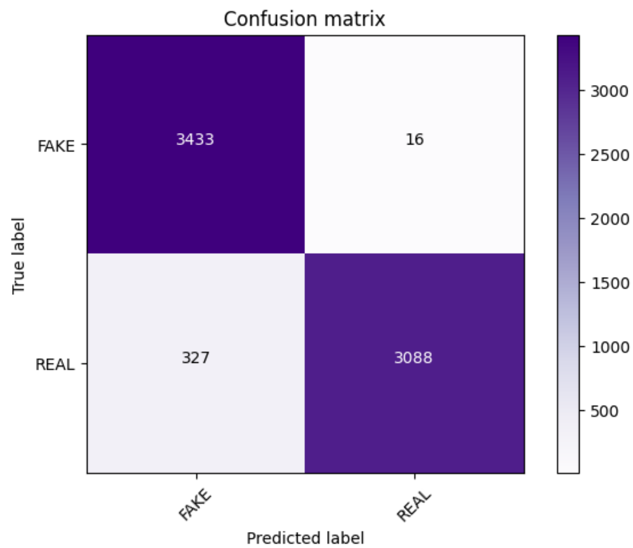

Fig. 2 Multinomial Naïve Bayes Confusion Matrix

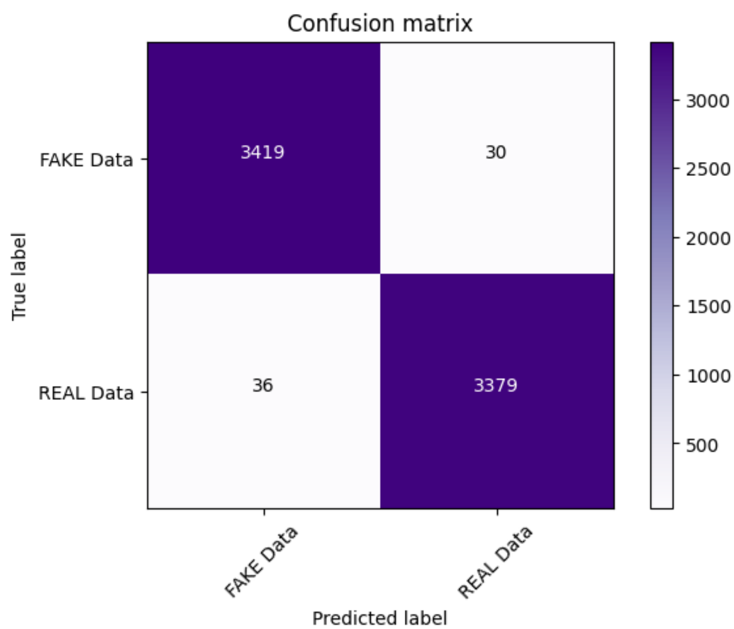

Fig. 3 Passive Aggressive Confusion Matrix

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 6, Number 1, January 2024
https://doi.org/10.47709/cnahpc.v6i1.3466

Submitted : January 17, 2024
Accepted : January 22, 2024
Published : January 23, 2024

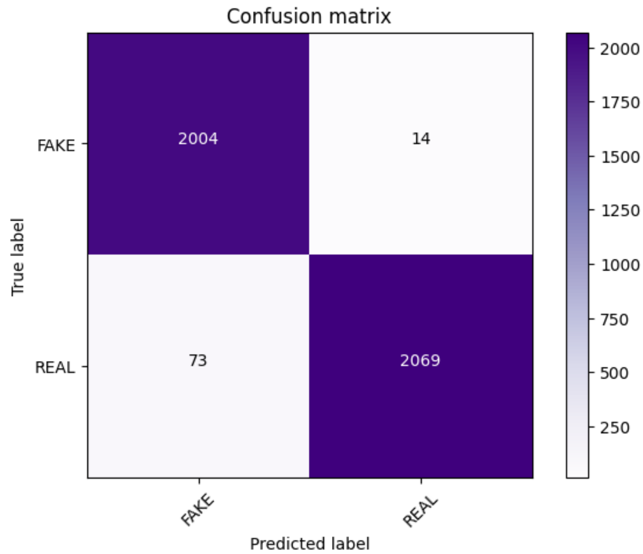

Fig. 4 Logistic Regression Confusion Matrix

	precision	recall	f1-score	support
0	0.99	0.96	0.98	2077
1	0.97	0.99	0.98	2083
accuracy			0.98	4160
macro avg	0.98	0.98	0.98	4160
weighted avg	0.98	0.98	0.98	4160

Fig. 5 Logistic Regression Classification Report

	precision	recall	f1-score	support
0	0.91	1.00	0.95	3449
1	0.99	0.90	0.95	3415
accuracy			0.95	6864
macro avg	0.95	0.95	0.95	6864
weighted avg	0.95	0.95	0.95	6864

Fig. 6 Multinomial Naïve Bayes Classification Report

	precision	recall	f1-score	support
0	0.99	0.99	0.99	3449
1	0.99	0.99	0.99	3415
accuracy			0.99	6864
macro avg	0.99	0.99	0.99	6864
weighted avg	0.99	0.99	0.99	6864

Fig. 7 Passive Aggressive Classification Report

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 6, Number 1, January 2024
https://doi.org/10.47709/cnahpc.v6i1.3466

Submitted : January 17, 2024
Accepted : January 22, 2024
Published : January 23, 2024

- Logistic Regression : Accuracy is 0.98
- Multinomial Naive Bayes : Accuracy is 0.95
- Passive Aggressive Classifier : Accuracy is 0.99

Fig. 8 Accuracy Results

As presented in the figure 5, The Logistic Regression model demonstrates high precision and recall scores across both classes (labeled '0' for fake news and '1' for real news). The precision for classifying fake news (class 0) is at 0.99, indicating that 99% of the articles predicted as fake news are indeed fake. Similarly, the model shows a recall of 0.96 for fake news, which means it correctly identified 96% of all the actual fake news articles. For real news (class 1), the precision is slightly lower at 0.97, but the recall is higher at 0.99, indicating that while there is a slight drop in the precision of classifying real news, the model is excellent at covering most real news articles. The F1-scores, which are the harmonic means of precision and recall, are also high at 0.98 for both classes, reflecting the balanced classification performance of this model. The overall accuracy of the Logistic Regression model stands at 0.98, corroborated by the accuracy summary provided in figure 8, indicating that 98% of the news articles were correctly classified.

On the other hand, as presented in the figure 6, the Naive Bayes model shows a precision of 0.91 for fake news, which is lower than the Logistic Regression model, suggesting that there is a higher rate of articles incorrectly labeled as fake. However, the recall is at 1.00, indicating that the model identified all actual fake news articles correctly. For real news, the precision is very high at 0.99, with a recall of 0.90. This means that while the model is very reliable when it identifies an article as real, it misses 10% of the actual real news articles. The F1-score is at 0.95 for both classes, showing a slightly lower performance in balancing precision and recall compared to Logistic Regression. The overall accuracy for the Naive Bayes model is at 0.95, meaning it correctly classified 95% of the articles, as seen in the accuracy summary as presented in figure 5. In addition, as presented in figure 6, The Passive Aggressive Classifier achieves excellent performance, with precision and recall scores of 0.99 for both fake and real news. This translates to a model that is 99% accurate in predicting both fake and real news, with a 99% success rate in identifying all the actual articles of each class. The F1-scores mirror these high values, indicating a near-perfect balance between precision and recall. This model's overall accuracy is at 0.99, suggesting that it correctly classified 99% of the news articles, which is confirmed by the accuracy summary provided in figure 7. This model stands out as the best performer among the three in terms of accuracy.

DISCUSSIONS

In summary, as presented in figure 8, the Passive Aggressive Classifier outperforms the other two models with the highest overall accuracy of 0.99. The Logistic Regression model follows closely with an accuracy of 0.98, and the Multinomial Naive Bayes model has a slightly lower accuracy of 0.95. It is important to note that while the Naive Bayes model has a perfect recall for the fake news class, it is less precise than the other models, indicating a tendency to classify more articles as fake, including some real news. In contrast, the Passive Aggressive Classifier shows exceptional balance across all metrics, making it a highly reliable model for this task. Each model's performance metrics suggest different strengths and weaknesses. The Logistic Regression model demonstrates a high level of consistency across different types of classification errors. The Naive Bayes model, while showing a tendency to over-classify articles as fake, does not miss any fake news articles, indicating it could be particularly useful in scenarios where missing fake news is a greater concern than mistakenly classifying real news as fake. The Passive Aggressive Classifier, with its high scores across all metrics, suggests a strong adaptability to the dataset and could be considered the most robust model for classifying news articles when both false positives and false negatives are equally undesirable. The F1-scores, which are particularly important in evaluating the balance between precision and recall, are consistently high for all models, indicating that they maintain a good balance between the thoroughness of the models (recall) and their correctness (precision). This balance is crucial in the context of fake news detection, where both identifying as much fake news as possible (high recall) and ensuring that what is identified as fake news is indeed fake (high precision) are important. the results from these models indicate that machine learning can be a powerful tool in distinguishing between fake and real news. The choice between models would depend on the specific requirements of the application, whether that's minimizing the spread of fake news (maximizing recall) or minimizing

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 6, Number 1, January 2024
https://doi.org/10.47709/cnahpc.v6i1.3466

Submitted : January 17, 2024
Accepted : January 22, 2024
Published : January 23, 2024

the misclassification of real news (maximizing precision). The Passive Aggressive Classifier, with its high performance across all metrics, would be an excellent general choice for a fake news detection system.

5. CONCLUSION

In conclusion, this study systematically evaluated three machine learning algorithms—Multinomial Naive Bayes, Passive Aggressive Classifier, and Logistic Regression—for detecting fake news. Through rigorous data collection, preprocessing, and evaluation, we found that the Passive Aggressive Classifier outperformed the others, demonstrating exceptional precision and recall. While machine learning offers an effective approach to combating fake news, the choice of model should align with the specific context. This research contributes to the literature on misinformation and highlights the need for responsible algorithmic solutions. Future research can explore algorithmic transparency and ethical deployment. As our information ecosystem evolves, developing advanced tools to preserve news integrity remains crucial, and this study offers valuable insights for researchers, technologists, and policymakers.

6. REFERENCES

Albahr, A., & Albahar, M. (2020). An empirical comparison of fake news detection using different machine learning algorithms. *International Journal of Advanced Computer Science and Applications*, *11*(9).

Alghamdi, J., Lin, Y., & Luo, S. (2022). A Comparative Study of Machine Learning and Deep Learning Techniques for Fake News Detection. *Information*, *13*(12), 576.

Awumee, G. S., Agyemang, J. O., Boakye, S. S., & Bempong, D. (2023). SmishShield: A Machine Learning-Based Smishing Detection System. *International Conference on Wireless Intelligent and Distributed Environment for Communication*, 205–221.

Berrondo-Otermin, M., & Sarasa-Cabezuelo, A. (2023). Application of Artificial Intelligence Techniques to Detect Fake News: A Review. *Electronics*, *12*(24), 5041.

Bui, D. T., Tsangaratos, P., Nguyen, V.-T., Van Liem, N., & Trinh, P. T. (2020). Comparing the prediction performance of a Deep Learning Neural Network model with conventional machine learning models in landslide susceptibility assessment. *Catena*, *188*, 104426.

Choudhary, A., & Arora, A. (2021). Linguistic feature based learning model for fake news detection and classification. *Expert Systems with Applications*, *169*, 114171.

de Oliveira, N. R., Medeiros, D. S. V, & Mattos, D. M. F. (2020). A sensitive stylistic approach to identify fake news on social networking. *IEEE Signal Processing Letters*, *27*, 1250–1254.

George, J., Gerhart, N., & Torres, R. (2021). Uncovering the truth about fake news: A research model grounded in multi-disciplinary literature. *Journal of Management Information Systems*, *38*(4), 1067–1094.

Gradoń, K. T., Hołyst, J. A., Moy, W. R., Sienkiewicz, J., & Suchecki, K. (2021). Countering misinformation: A multidisciplinary approach. *Big Data \& Society*, *8*(1), 20539517211013850.

Gupta, S., & Meel, P. (2021). Fake news detection using passive-aggressive classifier. *Inventive Communication and Computational Technologies: Proceedings of ICICCT 2020*, 155–164.

Hadlington, L., Harkin, L. J., Kuss, D., Newman, K., & Ryding, F. C. (2023). Perceptions of fake news, misinformation, and disinformation amid the COVID-19 pandemic: A qualitative exploration. *Psychology of Popular Media*, *12*(1), 40.

Hangloo, S., & Arora, B. (2021). Fake News Detection Tools and Methods--A Review. *ArXiv Preprint ArXiv:2112.11185*.

Hossain, M. D. N., Al Huda, S., Hossain, S., Sum, A. S. I., Fahad, N., & Sen, A. (2023). A Comprehensive Analysis of Machine Learning Approaches for Fake News Detection and Its Effects. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, *14*(03), 384–391.

Jain, M. K., Gopalani, D., Meena, Y. K., & Kumar, R. (2020). Machine Learning based Fake News Detection using linguistic features and word vector features. *2020 IEEE 7th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)*, 1–6.

Jarrahi, A., & Safari, L. (2023). Evaluating the effectiveness of publishers' features in fake news detection on social media. *Multimedia Tools and Applications*, *82*(2), 2913–2939.

Kanwal, S., Nawaz, S., Malik, M. K., & Nawaz, Z. (2021). A review of text-based recommendation systems. *IEEE Access*, *9*, 31638–31661.

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 6, Number 1, January 2024
https://doi.org/10.47709/cnahpc.v6i1.3466

Submitted : January 17, 2024
Accepted : January 22, 2024
Published : January 23, 2024

Kasinidou, M., Kleanthous, S., Barlas, P., & Otterbacher, J. (2021). I agree with the decision, but they didn't deserve this: Future developers' perception of fairness in algorithmic decisions. *Proceedings of the 2021 Acm Conference on Fairness, Accountability, and Transparency*, 690–700.
Ma, Y., Wu, L., Guan, Y., & Peng, Z. (2020). The capacity estimation and cycle life prediction of lithium-ion batteries using a new broad extreme learning machine approach. *Journal of Power Sources*, *476*, 228581.
Matti, R., & Yousif, S. (2023). AutoKeras for Fake News Identification in Arabic: Leveraging Deep Learning with an Extensive Dataset. *Al-Nahrain Journal of Science*, *26*(3), 60–66.
Mishra, S., Shukla, P., & Agarwal, R. (2022). Analyzing machine learning enabled fake news detection techniques for diversified datasets. *Wireless Communications and Mobile Computing*, *2022*, 1–18.
Molina, M. D., Sundar, S. S., Le, T., & Lee, D. (2021). "Fake news" is not simply false information: A concept explication and taxonomy of online content. *American Behavioral Scientist*, *65*(2), 180–212.
Musleh, D. A., Alkhwaja, I., Alkhwaja, A., Alghamdi, M., Abahussain, H., Alfawaz, F., … Abdulqader, M. M. (2023). Arabic Sentiment Analysis of YouTube Comments: NLP-Based Machine Learning Approaches for Content Evaluation. *Big Data and Cognitive Computing*, *7*(3), 127.
Prasad, B. L., Srividya, K., Kumar, K. N., Chandra, L. K., Dil, N., & Krishna, G. V. (2023). An Advanced Real-Time Job Recommendation System and Resume Analyser. *2023 International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS)*, 1039–1045.
Pröllochs, N., & Feuerriegel, S. (2023). Mechanisms of true and false rumor sharing in social media: Collective intelligence or herd behavior? *Proceedings of the ACM on Human-Computer Interaction*, *7*(CSCW2), 1–38.
Sabeeh, V., Zohdy, M., Mollah, A., & Al Bashaireh, R. (2020). Fake news detection on social media using deep learning and semantic knowledge sources. *International Journal of Computer Science and Information Security (IJCSIS)*, *18*(2), 45–68.
Saleh, H., Alharbi, A., & Alsamhi, S. H. (2021). OPCNN-FAKE: Optimized convolutional neural network for fake news detection. *IEEE Access*, *9*, 129471–129489.
Schroeder, D. T. (2022). *Explaining News Spreading Phenomena in Social Networks: From Data Acquisition and Processing to Network Analysis and Modelling*. Technische Universitaet Berlin (Germany).
Sciannamea, R., & others. (2020). *Fake News: Evolution of a rising concept and implications for the education system*.
Smitha, N., & Bharath, R. (2020). Performance comparison of machine learning classifiers for fake news detection. *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, 696–700.
Starke, C., Baleis, J., Keller, B., & Marcinkowski, F. (2022). Fairness perceptions of algorithmic decision-making: A systematic review of the empirical literature. *Big Data \& Society*, *9*(2), 20539517221115188.
Toma, G.-A., & others. (2021). Fake news as a social phenomenon in the digital age: a sociological research agenda. *Sociologie Românească{\u{a}}*, *19*(1), 134–153.
Walsh, J. P. (2020). Social media and moral panics: Assessing the effects of technological change on societal reaction. *International Journal of Cultural Studies*, *23*(6), 840–859.
Yerlekar, A., Mungale, N., & Wazalwar, S. (2021). A multinomial technique for detecting fake news using the Naive Bayes Classifier. *2021 International Conference on Computational Intelligence and Computing Applications (ICCICA)*, 1–5.
Yuslee, N. S., & Abdullah, N. A. S. (2021). Fake News Detection using Naive Bayes. *2021 IEEE 11th International Conference on System Engineering and Technology (ICSET)*, 112–117.
Zhang, X., & Ghorbani, A. A. (2020). An overview of online fake news: Characterization, detection, and discussion. *Information Processing \& Management*, *57*(2), 102025.

* Corresponding author