JOURNAL OF GOVERNANCE AND INTEGRITY
ISSN: 2600-7479 e-ISSN: 2600-786X
VOLUME 8, ISSUE 1, 2025, 867 - 873
DOI: https://doi.org/10.15282/jgi.8.1.2025.11289

RESEARCH ARTICLE

Machine learning technique for online fake news detection

Nur Khairunnisa Kamaruddin*, Mohd Faizal Ab Razak, Ahmad Firdaus Zainal Abidin, Salwana Mohamad@Asmara, Anishah Muhammad Syafiq

Faculty of Computing, Universiti Malaysia Pahang Al-Sultan Abdullah, 26600 Pekan, Pahang, Malaysia

ABSTRACT – The Internet has become a fundamental part of our daily lives, influencing social, political, and other domains. In the modern era, social media has emerged as a powerful tool for individuals and organizations globally. While social media offers significant benefits, such as the ability to disseminate news globally, it also poses the risk of spreading false information. Fake news can be particularly harmful as it spreads rapidly and without limits, potentially triggering political disagreements, harming mental health, and impacting the economy. This research delves into the phenomenon of fake news in detail, proposing detecting fake news through machine learning techniques. Eminie Bozkus provided a dataset for the training and testing processes. This research aims to evaluate various algorithms' accuracy and compare their effectiveness in detecting fake news.

ARTICLE HISTORY
Received : 16-09-2024
Revised : 12-01-2025
Accepted : 20-05-2025
Published : 30-07-2025

KEYWORDS
Fake news
Support vector machine
Naïve Bayes
Decision tree
Governance

1. INTRODUCTION

Social media has completely changed how we interact, share information, and have conversations with one another in the 21st century. It encompasses diverse digital tools designed to facilitate networking, sharing, and social interaction. Hence, social media growth encompasses websites such as Facebook, Twitter, Instagram, Pinterest, TikTok, and the like. This platform can help individuals and organizations grow their circles of influence, share their ideas, and spread information to a global audience. In today's digital era, social media platforms are useful in the global dissemination of information, including those related to the world economy, education, politics, current news, and the like. The emergence of social media has made the way for information to be shared, used, and spread in the current day with greater velocity than ever before. Information dissemination offers numerous advantages, including the capacity to elevate public consciousness of the importance of knowledge; nevertheless, it also presents the risk of disseminating inaccurate information. These platforms may be exploited to spread false information to malicious ends, such as inciting hatred based on extremist ideologies, influencing public opinion for political purposes, and forming biased opinions to win elections or financial gain (Hamed et al., 2023). The problem of fake news on social media can affect politics and democracy by representing significant risks that vary and undermine the trust of everyone in media sources. This may cause people to use slanderous or abusive language, place blame on certain individuals, and start racial arguments on social media.

Furthermore, the dissemination of misinformation can harm mental health because people are prone to panicking and making assumptions based on what they know, which brings them in danger of harm. In April 2020, over 4,000 fake news articles circulated, including misinformation regarding the COVID-19 pandemic, increasing fear among the panic-stricken public (Hamed et al., 2023). Fake news propagation can potentially cause the global economy's collapse. Consequently, economic impacts may have a detrimental effect on consumers and business communities. They also can potentially increase the risk of monetary challenges and fraudulent business syndicates. Another example is a fake tweet from Eli Lilly, which promised free insulin on November 10, 2022. The pharmaceutical multinational denied the news, but by the next day, shares plunged by approximately $22 billion (Arcuri et al., 2023). Hence, a machine-learning approach is used to analyze or detect fake news. Jain and Kasbe conducted a study evaluating the accuracy of various machine learning algorithms for making predictions (Gupta et al., 2023). In this study, the tool used to generate the data was Phyton. Governments and regulatory agencies must establish and enforce policies that reduce the dissemination of fabrication. Establishing specific guidelines on what defines fake news and the penalties for spreading it requires collaboration between social media organizations

2. LITERATURE REVIEW

Types of fake news

i) Manipulated Content: Edited images, videos, or audio (e.g., deepfakes) designed to deceive and mislead audiences.

ii) Misleading Content: Twisted or incomplete information presented as sensational headlines, often used to incite conflict or spread propaganda.

*CORRESPONDING AUTHOR | N. K. Kamaruddin | ✉ nisaakmrdn@gmail.com

iii) Fabricated Content: Completely false information created to confuse or harm, often designed to appear credible and attract curiosity.

iv) Satire/Parody: Fictional stories are shared humorously but sometimes mistaken as factual, often criticizing politicians, celebrities, or society.

v) False Connection: Headlines, images, or backgrounds that do not match the actual content, leading to misinterpretation.

vi) False Context: Accurate facts with misleading background details distort the overall meaning.

vii) Imposter Content: Fake websites or profiles that mimic legitimate news sources to spread false or misleading information.

There are several methods for proposing a system that can accurately detect online fake news.

Table 1 summarizes research related to fake news detection, demonstrating the effectiveness of various machine-learning techniques and models. Jose et al. (2021) utilized Logistic Regression and Bi-directional LSTM RNNs, achieving accuracy of 90.37 and 93.46%, respectively, by leveraging the natural language processing (NLP) technique. Kai Xuan et al. (2023) achieved remarkable accuracy rates with decision trees (DTs) (99.9%), support machine vectors (SVM) (99.6%), and Naïve Bayes (95.9%). Sharma and Garg (2021) explored multiple models, with SVM reaching 91% accuracy. Matheven and Kumar (2022) introduced a hybrid deep learning and NLP approach, achieving at least 90% accuracy with LSTM networks. Krishna and Adimoolam (2022) found DTs to be highly effective, with an accuracy of 97.67% and precision of 94.60%, whereas SVM had 91.74% accuracy and 90.12% precision. These studies collectively highlight the significance of combining advanced machine-learning models with NLP techniques to enhance the accuracy and reliability of fake news detection systems.

The implications of exploring misinformation involve intricate moral dilemmas about the objectivity of the research methodology, consequences, and responsibilities of the researcher. Maintaining integrity requires accurate examination, openness to methodology and conflicts of interest, and solid, objective procedures. Public trust is a major concern because research results can affect public confidence in the media; therefore, responsible communication is necessary to avoid fear or mistrust. The influence on policymaking, preventing stigmatization, safeguarding data privacy, and obtaining informed permission are additional ethical factors to consider. Researchers must balance safeguarding free expression and preventing disinformation, maintaining political objectivity, and interacting with various stakeholders. Ethics committees should assess proposals and using artificial intelligence (AI) and machine intelligence responsibly is crucial. Sustaining integrity and participating in a responsible conversation about fake news requires ongoing ethical thought and compliance with publication restrictions.

3. METHODOLOGY

A detailed discussion of the methodology, encompassing multiple techniques and approaches, is provided in Figure 1 to identify and assess fake news. The synthesis of fake news may involve many forms, such as text, photos, audio, and video, all of which can be identified using the suggested techniques and approaches. To identify fake news, this research project went through various phases.

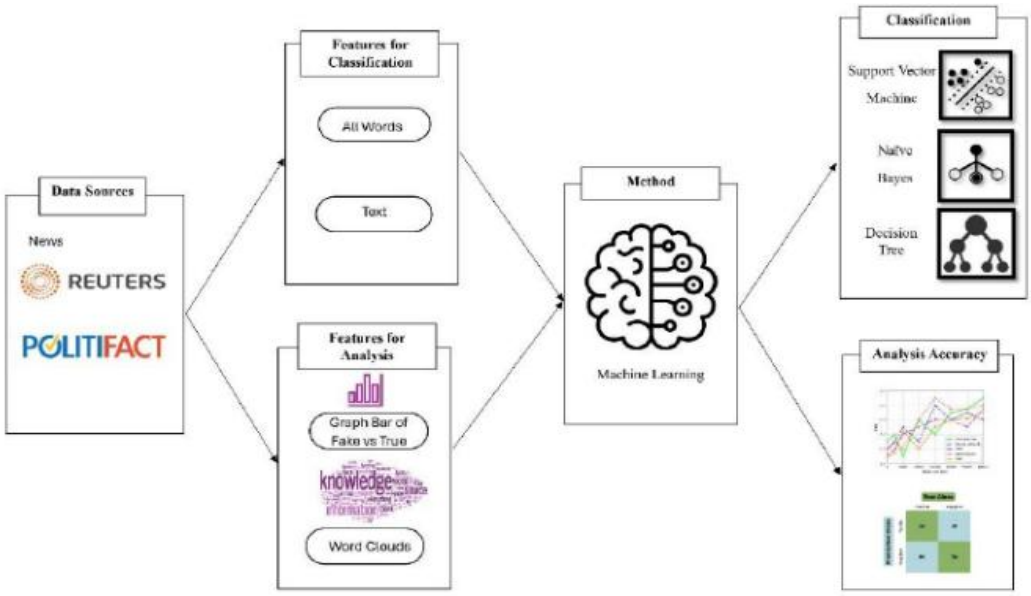

Figure 1. Framework of fake news

Table 1. Comparison of fake news detection

Element	Previous Work 1	Previous Work 2	Previous Work 3	Previous Work 4	Previous Work 5
References	(Jose et al., 2021)	(Kai Xuan et al., 2023)	(Sharma & Garg, 2021)	(Matheven & Kumar, 2022)	(Krishna & Adimoolam, 2022)
Aim	Focuses on detecting false connections and fabricated content.	Investigate using social media data to identify fake news about the COVID-19 pandemic.	Identifying fake news based on headlines Hindi	Study the effectiveness of a deep learning model combined with an NLP model in identifying fake news.	Present an insight into the classification of news articles in the modern community, as well as the various content types of news articles and their impact on readers.
Problem	Identifying relevant features that distinguish fake news from real news makes fake news detection a challenging problem.	To detect the COVID-19 fake news.	The challenge is to design a real-time tool to detect fake news.	Research the practicality of developing an algorithm to identify and detect fake news accurately and efficiently.	The comparison of decision tree with the support vector machine (SVM)
Technique	Using the National Language Processing (NLP) technique based on Google News API and the tweet given to the machine learning model for classification.	NLP and the SVM.	SVM, Gaussian Na¨ıve Bayes, K-neighbors, Random Forest, Decision Tree.	NLP techniques.	Decision tree (DT) algorithm and SVM algorithm
Accuracy	LSTM to reach an accuracy of 93.46%.	DT achieves the highest accuracy of 99.9%, SVM has 99.6% accuracy, and Na¨ıve Bayes has an accuracy of 95.9%	SVM is 91%.	LSTM is 90% accurate.	The DT method has 97.67% accuracy, and the SVM method has 91.74% accuracy.

3.1 Research Framework

Developing algorithms using machine learning to detect fake news requires collecting raw data from various social media sources, known as the data collection process. NLP techniques, such as tokenization and stopword removal, are used in preprocessing to clean and prepare data for analysis. By reducing the number of features while maintaining important data, dimensionality reduction improves the computational efficiency. Several machine learning techniques are used: DTs use a tree-structured prediction model to identify fake news by making decisions based on feature values, SVM build an optimal hyperplane to classify data, and Naïve Bayes classifiers use probabilistic learning for text classification.

3.2 Project Requirement

The research on fake news detection utilizes Eminie Bozkus's dataset and employs three top classifiers: Naïve Bayes, SVM, and DT. The process involved data collection, preprocessing, feature extraction, dataset division, and model evaluation. Key constraints include lengthy data filtering and increased training time with larger datasets owing to the computational demands of the SVM. A case study highlighted the pervasive issue of fake news spread via social and traditional media, which can harm individuals and organizations. This study compares the accuracy of the three classifiers to identify the most effective model for detecting fake news on social media platforms, ultimately selecting the best-performing algorithm.

Based on Figure 2, this research involves extracting data from a fake news dataset to train, test, and evaluate machine learning models for fake news detection. Data preprocessing includes cleaning, organizing, and transforming raw text by removing duplicates, non-alphabetic characters, and stop words, followed by tokenization. Term Frequency-Inverse Document Frequency (TF-IDF) numerically represents word significance in documents. The dataset is split into training and testing sets for building and evaluating the Naïve Bayes, SVM, and DT models. This process ensures that the models can make accurate predictions of new data by preventing overfitting and comparing the accuracy of each classifier.

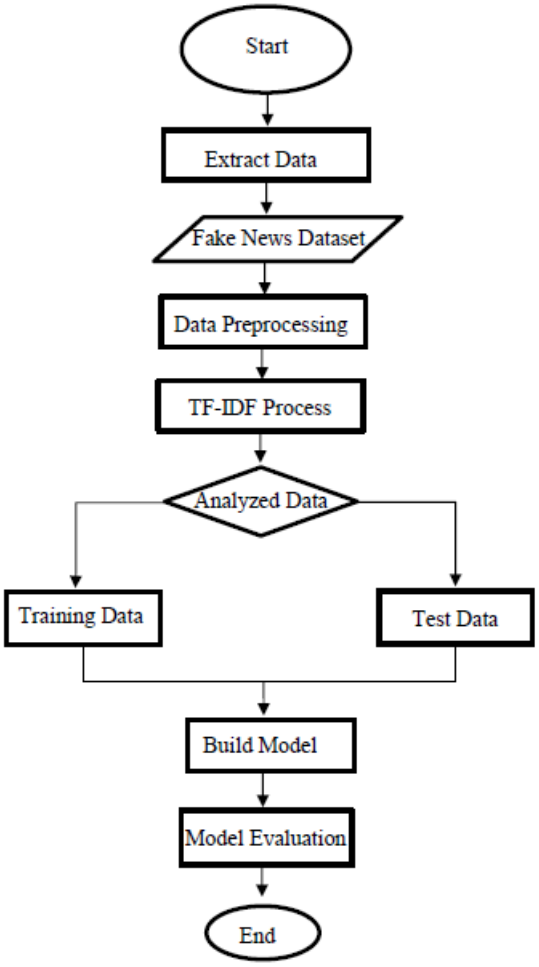

Figure 1. Research framework

3.3 Data Design

The data used for this research were provided by Eminie Bozkus, which are real and fake news, and the dataset was split into two: real and fake news. The fake news dataset comprised 23446 news, and the real news dataset contained 21417 news from the propagating fake news issue. This dataset was combined into one dataset, and the total dataset was 44863.

3.4 Implementation

The research began with data extraction from a fake news dataset, followed by data preprocessing to clean and transform the raw text. This included removing duplicates, non-alphabetic characters, stopwords, converting text to lowercase, tokenizing, and lemmatizing words. The data were divided into training and testing sets. TF-IDF was used to represent word significance numerically. The analysis included visualizing word frequencies and creating word clouds to highlight key topics in fake news. Dimensionality reduction compressed data while retaining essential information, improving computational efficiency and model generalization. The dataset, containing titles, texts, subjects, and fake news labels, was analyzed for biases and imbalances. The resulting TF-IDF matrices were used to train Naïve Bayes, SVM, and DT models, with each algorithm's accuracy compared to determine the best model for detecting fake news.

4. RESULTS AND DISCUSSION

The results from the three machine learning classifiers, SVM, Naïve Bayes, and DT, are presented below. The outcome consisted of five parameters: accuracy, precision, recall, f1-score, and support.

Table 2. Result accuracy

Classifier	Accuracy (%)	Precision	Recall	F1-Score	Support
Support Vector Machine	99.42	1.00	99.0	99.0	7041
Naïve Bayes	76.4	73.0	86.0	79.0	514
Decision Tree	94.00	94.0	94.0	94.0	4729

Regarding the accuracy of the results in Table 2, the SVM achieved the highest accuracy of 99.42% compared to the DT, which was 94.00% in identifying fake news. According to observations of the accuracy of the outcome classifiers, the SVM is more effective than the Naïve Bayes and DT in detecting fake news. Furthermore, it indicates that the efficacy of fake news detection depends on feature selection. However, the precision rate can affect the classifier's ability to produce significant and accurate results.

The ratio of true positives to the total number of true positives and false positives is used to determine the precision, which is a measure of the accuracy of the positive predictions. A high precision means the classifier predicts positive cases with low error rates. Called the ratio of true positives to the total of true positives and false negatives, recall, also called sensitivity, evaluates the classifier's capacity to detect all positive instances. A high recall rate indicates that the majority of real positive examples are correctly captured by the classifier. The F1-score, particularly helpful in situations with an unequal class distribution, is the harmonic mean of the precision and recall. This single statistic balances these two values. A high F1 score suggests that recall and precision are well-balanced. By displaying the dataset's distribution, support which is the number of real occurrences in each class provides context for the other metrics. When combined, these indicators provide a thorough understanding of the performance of a classifier.

3.5 Model Evaluation

Model evaluation of the process to assess the performance and effectiveness of SVM, Naïve Bayes, and DT for classifying online fake news data.

Precision

$$Precision = \frac{true\ positive}{true\ positives + false\ positives}$$

Recall

$$Recall = \frac{true\ positive}{true\ positives + false\ negatives}$$

F1-Score Formula

$$F1 = 2.\ \frac{precision\ .recall}{precision + recall}$$

These metrics allow analysts to evaluate predictions for completeness and accuracy, comprehend the effects of class imbalances, and determine how effectively a model generalizes to actual data.

5. CONCLUSIONS

The conclusion of fake news research presented an understanding of machine learning techniques for online fake news. The dataset utilized in this research comprised several fake news articles that have been collected and features selected to detect fake news. Machine learning algorithms were implemented on the model and the accuracy of classifiers, such as SVM, naïve Bayes, and DT. From this approach, the best outcome for each machine learning method was that SVM obtained 99.42%. Hence, the SVM is the most recommended algorithm for fake news detection in this study.

FUNDING STATEMENT

This study was not supported by any grants from funding bodies in the public, private, or not-for-profit sectors.

AUTHORS CONTRIBUTION

Nur Khairunnisa Binti Kamaruddin (Conceptualization; Formal analysis; Visualization; Methodology; Data curation; Writing - original draft; Writing – review & editing)

Ts. Dr. Mohd Faizal Ab Razak (Supervision; Project administration; Guidance)

Ts. Dr. Ahmad Firdaus Bin Zainal Abidin (Supervision; Guidance)

Ts. Dr. Salwana Bt Mohamad @ Asmara (Supervision; Guidance)

Anishah Binti Muhammad Syafiq (Support)

AVAILABILITY OF DATA AND MATERIALS

The data supporting this study's findings are available on request from the corresponding author.

ETHICAL STATEMENT

Not applicable.

CONFLICT OF INTEREST

The author(s) declare there is NO conflict of interest, financial or non-financial, that could have influenced the content of this manuscript.

REFERENCES

Arcuri, M. C., Gandolfi, G., & Russo, I. (2023). Does fake news impact stock returns? Evidence from US and EU stock markets. *Journal of Economics and Business*, *125*, 106130.

Brahim, G. Ben, Tomar, G. S., & Institute of Electrical and Electronics Engineers Saudi Arabia Section. (n.d.). 2022 14th IEEE International Conference on Computational Intelligence and Communication Networks (CICN 2022): proceedings

Emine Bozkuş. (2016). *Fake News Detection Datasets*. Kaggle.com. https://www.kaggle.com/datasets/emineyetm/fake-news-detection-datasets

Gupta, A., Batla, A., Kumar, C., & Jain, G. (2023, June). Comparative analysis of machine learning models for fake news classification. In *2023 3rd International Conference on Intelligent Technologies (CONIT)* (pp. 1-5). IEEE.

Hamed, S. K., Ab Aziz, M. J., & Yaakub, M. R. (2023). 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*, *9*(10).

Hoti, A. H., Hoti, M. H., Hoti, H., & Salihu, A. (2022, June). Identifying Fake News written on Albanian language in social media using Naïve Bayes, SVM, Logistic Regression, Decision Tree and Random Forest algorithms. In *2022 11th Mediterranean Conference on Embedded Computing (MECO)* (pp. 1-6). IEEE.

Jindal, K., Bhardwaj, V., Ray, S., Parvez, U., & Raj, V. (2022, December). Compare The Performance of Machine Learning Classifiers for Misinformation Detection. In *2022 5th International Conference on Contemporary Computing and Informatics (IC3I)* (pp. 1284-1289). IEEE.

Jose, X., Kumar, S. M., & Chandran, P. (2021, October). Characterization, classification and detection of fake news in online social media networks. In *2021 IEEE Mysore Sub Section International Conference (MysuruCon)* (pp. 759-765). IEEE.

Krishna, N. L. S. R., & Adimoolam, M. (2022, February). Fake News Detection system using Decision Tree algorithm and compare textual property with Support Vector Machine algorithm. In *2022 International conference on business analytics for technology and security (ICBATS)* (pp. 1-6). IEEE.

Matheven, A., & Kumar, B. V. D. (2022, November). Fake news detection using deep learning and natural language processing. In *2022 9th International Conference on Soft Computing & Machine Intelligence (ISCMI)* (pp. 11-14). IEEE.

Mladenova, T., & Valova, I. (2022, June). Research on the ability to detect fake news in users of social networks. In *2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)* (pp. 1-4). IEEE.

Pisner, D. A., & Schnyer, D. M. (2019). Support vector machine. In *Machine Learning: Methods and Applications to Brain Disorders* (pp. 101–121). Elsevier.

Sharma, D. K., & Garg, S. (2021, July). Machine learning methods to identify Hindi fake news within social-media. In *2021 12th International conference on computing communication and networking technologies (ICCCNT)* (pp. 1-6). IEEE.

Vadlamudi, P. S., Gunasekaran, M., & Nagalakshmi, T. J. (2023, January). An analysis of the effectiveness of the naive bayes algorithm and the support vector machine for detecting fake news on social media. In *2023 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE)* (pp. 726-731). IEEE.

Wang, L., Wang, Y., De Melo, G., & Weikum, G. (2019). Understanding archetypes of fake news via fine-grained classification. *Social Network Analysis and Mining*, *9*(1), 37.

Xuan, K. L. K., Bhuiyan, M. I., Kamarudin, N. S., Nasir, A. F. A., & Abdullah, M. Z. T. (2023, August). COVID-19 fake news detection model on social media data using machine learning techniques. In *2023 IEEE 8th International Conference on Software Engineering and Computer Systems (ICSECS)* (pp. 28-34). IEEE.