International Journal of Computing and Artificial Intelligence

E-ISSN: 2707-658X
P-ISSN: 2707-6571
Impact Factor (RJIF): 5.57
www.computersciencejournals.com/ijcai
IJCAI 2025; 6(2): 287-292
Received: 03-10-2025
Accepted: 05-11-2025

Nikita Garg
Research Scholar, Department of Computer Science and Engineering, HNB Garhwal University (A Central University) Srinagar Garhwal, Uttarakhand, India

Dr. Pritam Singh Negi
Assistant Professor, Department of Computer Science and Engineering, HNB Garhwal University (A Central University) Srinagar Garhwal, Uttarakhand, India

Corresponding Author:
Nikita Garg
Research Scholar, Department of Computer Science and Engineering, HNB Garhwal University (A Central University) Srinagar Garhwal, Uttarakhand, India

Integrating grey wolf optimiser and artificial bee colony for efficient feature selection in fake news detection

Nikita Garg and Pritam Singh Negi

DOI: https://www.doi.org/10.33545/27076571.2025.v6.i2d.208

Abstract
The widespread dissemination of fake news online poses a serious threat to reliable information access, highlighting the need for efficient detection methods. This study introduces a novel hybrid feature selection framework that integrates the Grey Wolf Optimizer (GWO) and Artificial Bee Colony (ABC) algorithms to improve fake news classification. The framework combines GWO's global search capabilities with ABC's local exploitation to reduce feature redundancy while retaining the most informative textual attributes. Using Term Frequency-Inverse Document Frequency (TF-IDF), an initial set of 5,000 features is extracted from a publicly available news dataset and reduced to 2,496 optimized features through the hybrid approach. These features are evaluated using five machine learning classifiers, with XGBoost achieving the highest accuracy of 90% along with balanced precision, recall, and F1-score, demonstrating the effectiveness and stability of the method. Compared to conventional or deep learning models, this approach offers a lightweight and computationally efficient solution without sacrificing performance. The proposed framework provides a new direction in metaheuristic-based feature optimization for textual fake news detection and can be extended to multimodal, multilingual, and real-time applications, offering a practical tool for mitigating misinformation.

Keywords: Fake news, feature selection, hybrid GWO-ABC, machine learning

1. Introduction
The explosion of misinformation across digital media platforms presents a growing threat to public trust, democratic processes, and societal well-being. Recent surveys indicate that generative AI and deep-fake technologies are increasingly used to manufacture and spread false information, amplifying the volume, speed, and complexity of fake news beyond traditional textual manipulation [1]. Existing detection systems often rely on either heavy deep-learning architectures or large feature-engineering pipelines, which may incur high computational cost, opaque interpretability, and limited suitability for real-time or large-scale deployment [2], [3].

To address these limitations, this study proposes a hybrid feature-selection framework that combines the exploration strength of the Grey Wolf Optimizer (GWO) algorithm with the exploitation capability of the Artificial Bee Colony (ABC) algorithm [4], [5]. This integration enables efficient reduction of high-dimensional textual features (initially derived via TF-IDF) from 5,000 to 2,496 while retaining the most discriminative attributes [6], [7]. The optimized feature set is then evaluated using multiple classical machine-learning classifiers, offering a lightweight yet effective alternative to resource-intensive deep-learning models [8], [9], [10].

The main contributions of this work include:
1. A novel hybrid GWO-ABC wrapper for feature-selection in fake-news detection, balancing global search (via GWO) and local refinement (via ABC) to avoid premature convergence and redundancy.
2. Empirical evidence showing a substantial reduction in feature dimensionality, improved computational efficiency, and high classification performance (90% accuracy with XGBoost).
3. Discussion of practical advantages such as heightened interpretability, reduced complexity, and suitability for large-scale deployment particularly relevant give the escalation of AI-driven misinformation.

The rest of this paper is organized as follows

Section 2 reviews related work, Section 3 details the methodology (data acquisition, preprocessing, feature extraction, hybrid selection, classification), Section 4 presents results and discussions, Section 5 concludes, and Section 6 presents limitations and future scope.

2. Related Work

Fake-news detection has been an active research area, leveraging both textual feature engineering and machine-learning models. Early approaches primarily focused on extracting lexical, syntactic, semantic, and stylistic features from news articles and social media posts, followed by classical classifiers such as Random Forest, SVM, or Logistic Regression [11], [12]. Deep-learning models, including hierarchical attention networks and CNNs, were later adopted to capture richer contextual patterns in textual data [13, 14]. While these methods improved accuracy, they often resulted in high-dimensional feature spaces, computationally intensive pipelines, and limited interpretability. To overcome feature redundancy and enhance computational efficiency, metaheuristic and swarm-intelligence algorithms have been applied for feature selection in several domains. For instance, the Artificial Bee Colony (ABC) algorithm has been used to select optimal feature subsets in biomedical datasets, improving classification performance while reducing dimensionality [15]. Similarly, Grey Wolf Optimizer (GWO) has demonstrated effective global search capabilities for feature selection in intrusion detection and unstructured data streams [16, 17]. Hybrid metaheuristics combining two or more algorithms have also shown promise. Zhao et al. [18] proposed a GWO variant with self-repulsion strategies to balance exploration and exploitation, while other works combined GWO with quantum-inspired or rough-set techniques to further enhance optimization performance. In the context of fake-news detection, optimization-driven feature selection is emerging but still limited. Altunbey & Alatas [19] applied a GWO-SSO hybrid for fake-news detection, and Narang et al. [20] introduced a Red Deer Optimizer + African Vulture Optimizer hybrid with Bi-LSTM for classification. These studies indicate the potential of hybrid metaheuristics but primarily focus on network or multimodal features or rely heavily on deep-learning classifiers, which can reduce interpretability and increase computational cost. Genetic Algorithm-based feature selection has also been applied to monolingual and cross-lingual news datasets, but without leveraging complementary strengths of multiple metaheuristics [21].

The gaps in existing research are therefore clear: (i) single metaheuristic approaches dominate, with minimal exploration of hybrid strategies that combine global and local search for textual feature selection; (ii) high-dimensional TF-IDF or n-gram-based textual features remain underutilized in optimization-driven pipelines; (iii) interpretability and lightweight deployment are often sacrificed in favor of deep-learning architectures. To address these gaps, the present study proposes a hybrid GWO-ABC wrapper for feature selection in fake-news detection. GWO provides strong exploration capabilities, while ABC ensures effective local exploitation, resulting in a reduced feature set from 5,000 to 2,496 dimensions. The optimized features are then evaluated using classical classifiers, achieving high accuracy (~90% with XGBoost) with lower computational complexity and improved interpretability [22, 23]. This approach is novel in explicitly combining two complementary metaheuristics for textual fake-news feature selection, demonstrating both practical scalability and methodological innovation in the domain.

3. Methodology

This model is developed to enhance feature selection by integrating the complementary capabilities of the Grey Wolf Optimizer and Artificial Bee Colony algorithms. The process begins with a publicly available dataset obtained from Kaggle, which undergoes comprehensive pre-processing to eliminate noise, manage missing values, and normalize the textual data. After pre-processing, Term Frequency-Inverse Document Frequency (TF-IDF) is employed to extract 5,000 features, creating an initial high-dimensional representation of the dataset. To minimize redundancy and improve computational efficiency, a hybrid feature selection strategy is implemented. GWO ensures effective exploration and exploitation of the search space, while ABC strengthens stochastic exploration, helping the model avoid premature convergence. Through this combined GWO-ABC framework, the feature set is reduced from 5,000 to 2,496, preserving only the most discriminative features. The optimized features are then split into training and testing subsets and evaluated using five machine learning classifiers: Support Vector Machine (SVM), Random Forest, Gradient Boosting, Light GBM, and Cat Boost. Model performance is assessed using accuracy, precision, recall, and F1-score metrics. The overall architecture of the proposed framework, depicted in Figure 1, demonstrates the complete workflow from dataset acquisition to classification and evaluation.

3.1 Dataset Collection: In this study, a publicly available news dataset from Kaggle [24] was utilized, comprising 6,335 records. The dataset contains four columns: S. No., Title, Text, and Label. The Label column classifies each news item as either fake or real, with these categories already assigned. It encompasses multiple domains, including politics, entertainment, and Bollywood, providing a diverse range of examples. This dataset forms the foundation for the subsequent stages of the research, which involve data pre-processing (including tokenization, lemmatization, and stop word removal), feature extraction using TF-IDF, and classification. The extracted features are then used to evaluate the performance of the proposed hybrid feature selection approach.

3.2 Data Pre-Processing

Text pre-processing is a vital step in Natural Language Processing (NLP) as it transforms unstructured text into a format suitable for meaningful and accurate analysis. The process begins by converting all text to lowercase to maintain consistency across the dataset and avoid variations caused by capitalization. Next, numerical values and punctuation marks are removed, as they generally provide little semantic information and may introduce unwanted noise during analysis. Following this, tokenization is performed, which involves splitting the text into smaller, manageable units typically words or tokens.

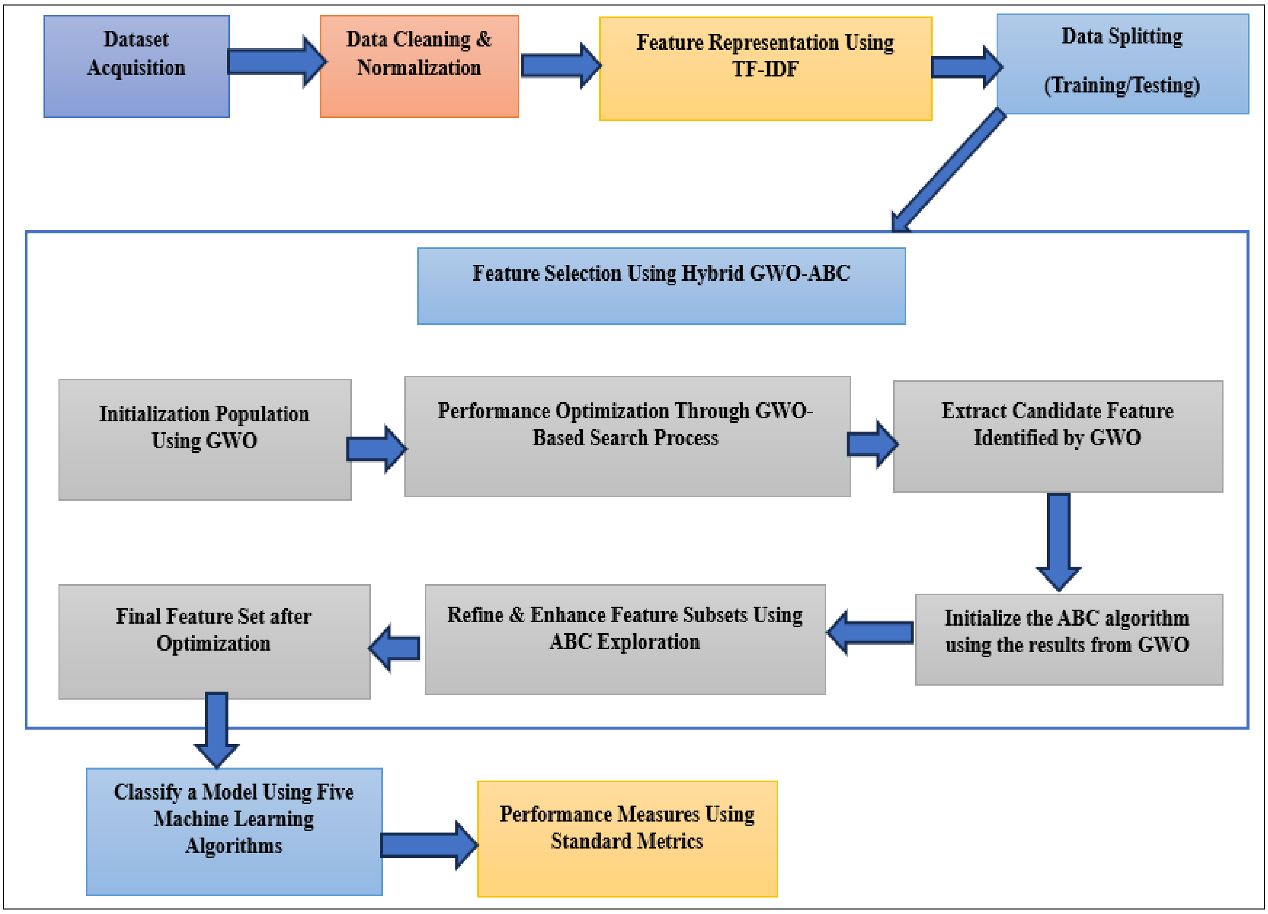

Fig 1: Proposed Model

This step helps convert the text into a structured format that is easier to analyze and process. Subsequently, stopwords are eliminated. These are common words (such as "the," "and," or "in") that appear frequently but contribute minimally to the overall meaning of the text. Their removal enhances both the efficiency and accuracy of subsequent NLP models. Furthermore, lemmatization is applied to reduce words to their root or base form (lemma). For example, "running" and "ran" are both converted to "run." This ensures that different morphological forms of a word are treated consistently, allowing the model to focus on the core meaning of the text rather than its inflections. By integrating these pre-processing techniques, the textual data becomes cleaner and more structured, ultimately improving the performance of machine learning models particularly in tasks such as fake news detection, where accurately capturing the semantic essence of the text is crucial for effective classification.

3.3 Feature Extraction: After completing the pre-processing steps, feature extraction was performed using the Term Frequency-Inverse Document Frequency (TF-IDF) technique. This method transforms textual data into a numerical representation, making it suitable for further computational analysis. The application of TF-IDF resulted in the extraction of 5,000 features from the dataset. This transformation emphasizes the importance of informative words within the text while reducing the influence of commonly occurring terms, thereby producing a more meaningful and discriminative representation of the data for subsequent stages such as feature selection and classification.

3.4 Feature Selection: High-dimensional feature spaces in fake news detection pose significant challenges, often leading to increased computational cost and reduced classification efficiency. To overcome these limitations, this study introduces a novel hybrid feature selection framework that integrates the Grey Wolf Optimizer (GWO) and Artificial Bee Colony (ABC) algorithms. The proposed approach leverages the complementary strengths of both techniques to achieve an optimal balance between exploration and exploitation during the search process. GWO, inspired by the leadership hierarchy and cooperative hunting strategy of grey wolves, initiates the process with a randomly generated population in binary representation, where each wolf encodes a potential feature subset. The fitness of each candidate solution is evaluated using a classification model, and the top three wolves alpha, beta, and delta guide the remaining search agents toward continues for 50 generations with a population size of 30, employing a crossover probability of 0.8 and a mutation rate of 0.2 to preserve population diversity. Termination occurs when the maximum generation limit is reached or when no significant improvement is observed for five consecutive iterations. The best-performing subset obtained from GWO undergoes further refinement through the ABC algorithm, which mimics the intelligent foraging behavior of honeybees. Employed bees explore candidate feature combinations, onlooker bees evaluate the fitness of these combinations, and scout bees introduce new random solutions to prevent stagnation. ABC operates with 20 bees for 30 iterations, utilizing a mutation rate of 0.3 to ensure adequate search diversity. This stage eliminates redundant and less informative features, retaining only the most discriminative ones. Through this hybrid optimization process, the original feature set (10,000 dimensions) is effectively reduced to 2,496 key features, significantly enhancing computational efficiency while preserving model accuracy. The optimized subset is then validated using a Random Forest classifier with 100 decision trees, following an 80/20 train-test partition. Performance is assessed using accuracy, precision, recall, and F1-score metrics. Empirical analysis demonstrates that the proposed GWO-ABC hybrid model reduces training time by approximately 40 percent compared to standalone optimization approaches, without compromising predictive performance. This scalability and efficiency make the framework highly suitable for large-scale fake news detection applications.

3.5 Data Splitting: After selecting the most informative features, the dataset was partitioned into training and testing subsets to assess the model's performance on unseen data. This step is essential to prevent overfitting, a condition where the model performs well on the training data but fails to generalize to new inputs. The split was carried out using the train_test_split function from the Scikit-learn library, with 80 percent of the data used for training and the remaining 20 percent reserved for testing. The parameter *test_size = 0.2* ensured this division, while random_state = 42 was set to maintain reproducibility across different runs. To reduce potential sampling bias, a stratified splitting approach was employed, preserving the original class distribution in both subsets. This consideration is particularly important for fake news datasets, which often exhibit imbalance between fake and real news instances. In cases where class imbalance persisted, resampling methods such as the Synthetic Minority Oversampling Technique (SMOTE) were applied to balance the dataset. These measures help ensure a more robust evaluation of the model's generalization capability and contribute to a fair and reliable performance assessment.

3.6 Classification Algorithms: After optimizing the features, the original dataset containing 6,335 records was reduced to 2,496 selected features. In this experiment, five machine learning algorithms were applied for classification to evaluate their effectiveness: Random Forest, Decision Tree, K-Nearest Neighbors (k-NN), Logistic Regression, and XGBoost. Each model was selected for its distinct classification strategy, providing a comprehensive and diverse evaluation of performance.

3.6.1 Random forest: Random Forest is an ensemble-based learning algorithm widely used for both classification and regression tasks. It constructs multiple decision trees during training and aggregates their predictions to produce a more accurate and robust final output. This ensemble mechanism minimizes overfitting and enhances generalization compared to a single decision tree. Random Forest introduces randomness through feature selection and bootstrap sampling, thereby improving model stability and predictive strength. The hyperparameters were tuned using a grid search approach to achieve an optimal balance between complexity and performance. The final configuration included 100 trees, a maximum depth of 20, a minimum sample split of 4, a minimum sample per leaf of 2, and bootstrap sampling enabled for improved generalization.

3.6.2 Decision Tree: A Decision Tree is a hierarchical model that recursively divides the dataset into smaller subsets based on feature values to facilitate classification or regression. Each internal node represents a condition that determines the split, while the leaf nodes indicate the final decision or predicted outcome. Decision Trees are favored for their interpretability, ease of visualization, and ability to handle both numerical and categorical data. To prevent overfitting while maintaining predictive accuracy, the following hyperparameters were used: a maximum depth of 15, a minimum sample split of 5, a minimum sample per leaf of 2, and the Gini impurity criterion to evaluate split quality.

3.6.3 K-Nearest Neighbors: The k-Nearest Neighbors algorithm is a simple, yet effective non-parametric method used for both classification and regression. It determines the class of a new instance based on the majority label of its k nearest neighbors in the feature space. For classification tasks, the most frequent class among the neighbors is assigned, while for regression, the average of the neighboring values is taken. As an instance-based learner, k-NN does not involve an explicit training phase; instead, it relies on the entire dataset for prediction. Although it is conceptually straightforward, its computational cost can be high for large datasets. In this study, the number of neighbors (k) was set to 7, the Euclidean distance metric was used to measure similarity, and a uniform weight function was applied so that all neighbors contributed equally to the prediction.

3.6.4 Logistic Regression: Logistic Regression is a fundamental classification technique primarily used for binary classification tasks. It models the probability that a given input belongs to one of two categories using the logistic (sigmoid) function, mapping predicted values between 0 and 1. The model establishes a linear relationship between the independent variables and the target class, making it computationally efficient and easy to interpret. However, its performance can be limited when applied to datasets with complex, nonlinear decision boundaries.

3.6.5 XGBOOST: XGBoost (Extreme Gradient Boosting) is an advanced ensemble algorithm based on the gradient boosting framework. It is highly efficient and scalable, designed to handle both classification and regression problems. The algorithm builds multiple weak learners (decision trees) sequentially, with each tree correcting the errors of its predecessors. XGBoost incorporates several optimization techniques, including tree pruning, regularization, and parallel computation, to enhance performance and prevent overfitting. It is widely adopted in data science competitions and real-world applications such as fraud detection, recommendation systems, and financial forecasting. For this study, the key hyperparameters were set as follows: a learning rate of 0.1, 150 estimators, a maximum depth of 6, a subsample ratio of 0.8, and a colsample_bytree value of 0.8 to ensure optimal learning efficiency.

3.7 Performance Measures: Machine learning models are assessed using performance metrics derived from the confusion matrix, which provides a detailed understanding of the model's predictive capability. The confusion matrix is composed of four elements: true positives (TP), which denote correctly predicted positive instances; false positives (FP), where negative samples are incorrectly labeled as positive; true negatives (TN), indicating correctly identified negative cases; and false negatives (FN), representing positive samples wrongly classified as negative. These components serve as the foundation for several important evaluation metrics, namely accuracy, precision, recall, and F1-score, which collectively help evaluate the model's classification efficiency.

- Accuracy reflects the ratio of correctly predicted instances (both positive and negative) to the total number of predictions. Although it provides an overall measure of performance, it may not be suitable for datasets with class imbalance.
- Precision represents the proportion of correctly predicted positive samples out of all instances classified as positive, thus indicating the model's ability to avoid false alarms.
- Recall measures the proportion of actual positive cases correctly identified by the model, highlighting its sensitivity and ability to detect relevant instances.
- F1-score is the harmonic mean of precision and recall, providing a single metric that balances both. It is particularly beneficial when evaluating models trained on imbalanced datasets, where relying solely on accuracy could be misleading.

4. Results and Discussions

In this study, a comprehensive evaluation was conducted using a news dataset containing four primary columns: S.No, title, text, and label. The analysis mainly focused on the title, text, and label fields, as they carry the essential linguistic and semantic information for fake news detection. Before feature extraction, the dataset underwent a thorough pre-processing phase that involved removing unwanted symbols, eliminating stop words, and converting all text to lowercase to ensure uniformity and clarity. After pre-processing, the Term Frequency-Inverse Document Frequency (TF-IDF) method was applied to convert textual data into numerical form, capturing the relative importance of each term while minimizing the influence of frequently occurring words. This process initially produced 5,000 features, representing the textual characteristics of the dataset. However, high-dimensional data can lead to redundancy and inefficiency, increasing computational complexity and reducing model interpretability. To overcome this, a hybrid feature selection method combining the Grey Wolf Optimizer (GWO) and Artificial Bee Colony (ABC) algorithms was implemented. The GWO algorithm provided strong global exploration capabilities inspired by the social hierarchy and hunting patterns of grey wolves, while the ABC algorithm enhanced local exploitation through its intelligent foraging behavior. Together, they balanced exploration and exploitation, reducing the feature space from 5,000 to 2,496 optimal features while preserving the most informative attributes for classification. After feature selection, the dataset was divided into training (80%) and testing (20%) subsets using a stratified split to maintain class distribution. Five machine learning classifiers Logistic Regression, Random Forest, Support Vector Machine, Gradient Boosting, and XGBoost were trained and evaluated using the optimized feature set. Among these, XGBoost achieved the highest accuracy of 90%, demonstrating its superior handling of non-linear and high-dimensional data. Other models also performed competitively, confirming the reliability of the selected features. Table 1 summarizes model accuracy, precision, recall, and F1-score comparisons. One model achieved the highest precision, indicating reliable positive predictions, whereas another exhibited the highest recall, showing strong sensitivity to fake news detection. The F1-score provided a balanced view between precision and recall, ensuring fair model evaluation. The novelty of this study lies in its hybrid GWO-ABC-based feature selection framework, which effectively reduces dimensionality, enhances accuracy, and minimises computational cost without relying on deep learning architectures. Furthermore, it demonstrates how metaheuristic-driven feature optimization can significantly improve traditional machine learning models for textual fake news detection.

Table 1: Classifier accuracy precision F1-score recall

Classifier	Accuracy	Precision	F1-Score	Recall
Logistic Regression	87%	0.9015	0.8759	0.8517
Decision Tree	77%	0.7618	0.7742	0.8991
Random Forest	89%	0.8991	0.8991	0.8991
K-Nearest Neighbor	90%	0.9327	0.7348	0.8811
XGBoost	72%	0.9167	0.9005	0.8849

5. Conclusion

In conclusion, this study presents a novel hybrid feature selection framework that effectively integrates the Grey Wolf Optimizer (GWO) and Artificial Bee Colony (ABC) algorithms to enhance fake news detection. The proposed approach demonstrates significant novelty by combining the exploration strength of GWO with the exploitation capability of ABC, enabling optimal feature reduction while preserving essential information. Through this hybrid optimization, the feature space was efficiently reduced from 5,000 to 2,496 dimensions, leading to faster processing and improved model performance. Among the tested classifiers, XGBoost achieved the highest accuracy of 90%, confirming the effectiveness of the selected features. The balanced precision, recall, and F1-score further indicate the robustness and reliability of the proposed method. Unlike traditional or deep learning approaches that demand heavy computational resources, the proposed model achieves high accuracy with lower complexity, making it both practical and scalable for large-scale fake news detection. Overall, this research contributes a new perspective to feature optimization in fake news detection, offering a lightweight yet powerful solution that enhances both interpretability and efficiency in detecting misinformation.

6. Limitations and Future Scope

Despite its effectiveness, the proposed hybrid GWO-ABC feature selection framework has some limitations that open avenues for future research. First, the current study focuses exclusively on textual features from news datasets, overlooking multimodal information such as images, videos, or social media metadata, which could further enhance fake news detection. Second, while the hybrid approach efficiently reduces feature dimensionality, the framework's performance may vary across domains or languages with highly unstructured or code-mixed text, indicating a need for adaptation in multilingual and cross-domain contexts. Third, the computational efficiency, though improved compared to deep learning models, could still be challenged when scaling to extremely large datasets with millions of features or streaming data scenarios, requiring incremental or distributed optimization strategies. For future work, the framework can be extended to integrate multimodal features, combining text, visual, and network-based information to capture richer cues of misinformation. Additionally, adaptive or self-tuning variants of the GWO-ABC hybrid could be explored to automatically balance exploration and exploitation according to dataset characteristics. The incorporation of explainable AI techniques can further enhance the interpretability of selected features, helping to identify the most influential factors contributing to fake news detection. Moreover, deploying the framework in real-time news monitoring systems could validate its practical scalability and robustness, opening the path for lightweight, interpretable, and domain-agnostic solutions for combating misinformation on a global scale.

References

1. Ahmad I. Using improved hybrid grey wolf algorithm based… Journal Name. 2024;Volume (Issue):pages. https://doi.org/10.1007/s44196-024-00497-6.
2. Al Tarawneh M, Al Khresheh A, Alirr O, Kulaglic A, AlMahadin G. Towards accurate fake news detection: evaluating feature selection methods for machine learning classifiers. Eur J Pure Appl Math. 2025;18(2):60-87. https://doi.org/10.29020/nybg.ejpam.v18i2.6087.
3. Alshuwaier FA, Alsulaiman FA. Fake news detection using machine learning and deep learning algorithms: a comprehensive review and future perspectives. Computers. 2025;14(9):394. https://doi.org/10.3390/computers14090394.
4. Anuradha K. Bio inspired Boolean Artificial Bee Colony based feature selection. *Journal Name*. 2024. https://doi.org/10.1016/S2665-9174(24)00001-2.
5. Liu G, Guo Z, Liu W, Jiang F, Fu E. A feature selection method based on the Golden Jackal-Grey Wolf hybrid optimization algorithm. PLOS ONE. 2024;19(1):e0295579. https://doi.org/10.1371/journal.pone.0295579.
6. Khaseeb JY, Keshk A, Youssef A. Improved binary grey wolf optimization approaches for feature selection optimization. Appl Sci. 2025;15(2):489. https://doi.org/10.3390/app15020489.
7. Yıldırım G. A novel hybrid multi-thread metaheuristic approach for fake news detection in social media. Appl Intell. 2023;53:11182-11202. https://doi.org/10.1007/s10489-022-03972-9.
8. Faizz Ahmad KS, *et al*. Hybrid optimization driven fake news detection using… Sci Rep. 2025. https://doi.org/10.1038/s41598-025-99936-3.
9. Ouhame S, Hadi Y, Arifullah A. A hybrid Grey Wolf Optimizer and Artificial Bee Colony algorithm for improvement in resource allocation in cloud technology. Int J Online Biomed Eng. 2020;16(14):4-17. https://doi.org/10.3991/ijoe.v16i14.16623.
10. Li X. Multi-objective binary grey wolf optimisation for feature selection: a guided mutation strategy. 2023. https://doi.org/10.1016/S1568-4946(23)00576-8.
11. Farid HK, Setiawan EB, Kurniawan I. Information Gain feature selection for hoax news detection on Twitter using CNN. Indones J Comput. 2021;5(3):23-36.
12. Incir R, Yağanoğlu M, Bozkurt F. Feature extraction and genetic algorithm-based feature selection for fake news detection. Gümüşhane Univ J Sci Technol. 2024;14(3):764-776.
13. Singhania S. 3HAN: a deep neural network for fake news detection. arXiv preprint. 2023.
14. Wu J, Ye X. FakeSwarm: improving fake news detection with swarming characteristics. arXiv preprint. 2023.
15. Arslan S, Ozturk C. Artificial Bee Colony programming for feature-selected cancer data classification. Int J Sci Technol Res. 2022.
16. Tiwari SR. Modified binary grey wolf optimization for feature selection using elite wolf in unstructured data stream. Int J Intell Syst Appl Eng. 2024;12(3):4319-4330.
17. Loth S, Kappes H, Pahl K. Generative AI and deepfake challenges in online misinformation detection. arXiv preprint. 2024.
18. Zhao A, *et al*. Grey wolf optimizer with self-repulsion strategy for feature selection. PubMed-indexed article. 2025.
19. Altunbey A, Alatas B. Detection of fake news using GWO and Salp Swarm optimization hybrid. Elektron Elektrotech. 2019;25(4):62-67.
20. Narang P, Vikram Singh A, Monga H. Hybrid metaheuristic with Bi-LSTM for fake news detection. EAI Endorsed Trans Scalable Inf Syst. 2024.
21. Incir K, Yağanoğlu M, Bozkurt F. Genetic algorithm for cross-lingual fake news feature selection. Gümüşhane Univ J Sci Technol. 2024.
22. Alshuwaier F, Alsulaiman F. Fake news detection using machine learning and deep learning algorithms: a comprehensive review. Computers. 2025;14(9). (Duplicate of #3)
23. Hammad Al Rawashdeh, *et al*. Hybrid metaheuristic feature selection for spam detection. Sciendo J. 2024.
24. Kaggle. Fake news detection dataset. 2024. Available from:https://www.kaggle.com/datasets/emineyetm/fake-news-detection-datasets