JOURNAL OF DATA SCIENCE | Vol.2026, Issue 3, No.13
eISSN:2805-5160

Metaheuristic-Based Feature Selection for High-Dimensional Data Classification: A Comparative Study

Elmar B. Noche[1]*, Wahyu Setyo Pambudi[2], Danang Haryo Sulaksono[3]

[1]Department of Information Technology, Pangasinan State University Lingayen Campus, Philippines
[2]Electrical Engineering Department, Institut Teknologi Adhi Tama Surabaya, East Java, Indonesia
[3]Informatics Engineering Department, Institut Teknologi Adhi Tama Surabaya, East Java, Indonesia

***Email:** elmarbnoche_ms@psu.edu.ph[1]*, wahyusp@itats.ac.id[2], danang_h_s@itats.ac.id[3]

Abstract

High-dimensional datasets often contain large numbers of redundant, irrelevant, and noisy features that increase computational complexity, weaken model generalization, and reduce the interpretability of classification outcomes. This problem is particularly important in data-intensive domains such as biomedical analytics, text classification, image recognition, cybersecurity, and intelligent decision-support systems, where the number of features may be substantially larger than the number of samples. Feature selection is therefore a critical preprocessing strategy for identifying compact and informative feature subsets while preserving predictive performance. This study presents a comparative investigation of four metaheuristic optimization algorithms—Particle Swarm Optimization, Genetic Algorithm, Grey Wolf Optimizer, and Whale Optimization Algorithm—for feature selection in high-dimensional data classification. Each algorithm is integrated with a Support Vector Machine classifier under a unified experimental framework to ensure fair comparison across optimization strategies. A controlled benchmark-style experiment is conducted using five representative high-dimensional dataset configurations covering biomedical, text, image, cybersecurity, and general classification characteristics. Model performance is evaluated using accuracy, F1-score, selected feature count, feature reduction rate, and computational time. The benchmark evaluation results show that metaheuristic-based feature selection reduces the original feature space by 71.8% to 82.4% while improving classification performance compared with the full-feature baseline. Among the evaluated methods, Grey Wolf Optimizer achieves the most balanced result, with an average accuracy of 91.8%, F1-score of 91.2%, and feature reduction rate of 79.6%. Whale Optimization Algorithm produces the highest feature reduction rate, while Particle Swarm Optimization provides the fastest metaheuristic-based computation. These findings indicate that metaheuristic feature selection can support process innovation in high-dimensional data analytics by producing more compact, efficient, and accurate classification models. This study contributes to data science research by offering a structured comparative framework and practical guidance for selecting suitable optimization-based feature selection methods under different high-dimensional classification conditions.

Submission: 20 September 2026; **Revised:** 15 October 2026 **Acceptance:** 13 November 2026; **Available Online:** November 2026
Copyright: © 2026. All the authors listed in this paper. The distribution, reproduction, and any other usage of the content of this paper is permitted, with credit given to all the author(s) and copyright owner(s) in accordance to common academic practice. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license, as stated in the website: https://creativecommons.org/licenses/by/4.0/

Keywords

Feature selection; Metaheuristic optimization; High-dimensional data; Classification; Dimensionality reduction; Process innovation

Introduction

High-dimensional data classification is central to many data-science-related disciplines today. With an ever-increasing volume of digital data, datasets containing hundreds to thousands, even tens of thousands, of data fields (or dimensions) have emerged. Examples include datasets formed in the process of biomedical diagnosis, gene expression analysis, text classification, image recognition, datasets involving cybersecurity, the Internet of Things, and datasets formed during business intelligence. Although high-dimensional data may show value, many features may be redundant or irrelevant, noisy, or weakly related to the class. Using such features in a machine learning model may lead to an increased computational burden, a decrease in the accuracy of predictions, and an increase in the difficulty of creating an interpretable model.

Feature selection is a widely used solution for high-dimensional problems. The goal of feature selection is to systematically eliminate irrelevant variables while retaining a smaller set of important variables. Unlike feature extraction, where original features are transformed to a new representation, with feature selection the meaning of the preserved variables remains intact. This, in applied data science, retains interpretability for the selected features. A feature selection method, in addition to attaining a smaller set of relevant features, should also decrease classification costs and improve overall classification efficiency.

Feature selection techniques can be divided into three major categories: filter, wrapper, and embedded techniques. Features are filtered and ranked by independent statistical measures such as correlation, information gain, chi-square value, and mutual information. This provides high speed, but may discard the possible interactions among the selected features and the learning algorithm. Embedded techniques use the model training process, like wrapper techniques. The difference is that wrapper techniques assess feature selection based on the performance of the learning model, and therefore, are related to the classification goal. The main disadvantage of wrappers is that, despite being the most flexible, they are the most computationally expensive techniques. This is because the search space is defined by the number of features and therefore grows combinatorially.

Besides the challenges mentioned above, metaheuristic optimization techniques can be an alternative to address wrapper feature selection. These techniques can be useful to search for solutions in large and complex solution spaces, because they do not require evaluating all the possible solutions. The most common population-based metaheuristic approaches include PSO, Genetic Algorithm (GA), Grey Wolf Optimizer (GWO), and the Whale Optimization Algorithm (WOA). These techniques consider high-dimensional feature selection, where every candidate solution is encoded in a binary vector indicating the selected (1) and unselected (0) features.

Recent studies indicate that adaptive intelligent modeling for classification in complex data environments can benefit from optimizers (Soori et al., 2023). Hajlaoui et al. (2026), for example, utilized a quantum-based optimizer and a feature pyramid network for Cloud-IoT intrusion detection, showing that optimization can drive significant improvements in classification. Following this trend, this study analyzes the application of various metaheuristic algorithms to feature selection for high-dimensional classification problems.

Although feature selection through metaheuristic approaches has been widely researched, there remain numerous gaps. First, most studies only examine one or two optimization algorithms, which hampers identification of the most effective algorithm for a given context (Shami et al., 2022). Second, while classification accuracy is often emphasized in the reviewed works, much less attention is paid to the balancing of the number of features selected and the time required to complete the task. Third, most of the existing comparisons are based on the application of different datasets and classifiers and different parameter settings. Fourth, little attention is paid to the balance of predictive accuracy, the number of features selected, and the time required to complete the task. These gaps clearly indicate a need for better comparability in feature selection approaches based on metaheuristics (Ali et al., 2023).

For this reason, the focus of this work is to use a common experimental design to assess PSO, GA, GWO, and WOA metaheuristic algorithms for high-dimensional data classification. Performance evaluation will involve assessment of the algorithms based on accuracy, F1-score, number of features selected, the degree of feature reduction, and the time taken to perform these computations. There are four primary ways this work will contribute to the body of knowledge. One, metaheuristic feature selection will be clearly compared. Two, a sort of benchmarking will be established to encourage more extensive testing. Three, concise and meaningful tables will be used to minimize comparison complexity. Fourth, an innovative process will be developed to view feature selection in the context of productive and scalable machine Learning.

Literature Review

High-Dimensional Data Classification

High-dimensional data classification is the supervised learning scenario in which there is a very high number of input variables and, in some cases, even more features than samples. This is referred to as the curse of dimensionality. With high-dimensional data, a learning algorithm may be unable to draw effective decision boundaries. With many redundant features, classifiers may learn to detect noise instead of signals. This typically leads to poor generalization and overfitting.

Classification of high-dimensional data is difficult because of the trade-off between the accuracy of the classification and the speed at which it may be done. In this case, the required time and the associated memory capacity to build the model and classify the data may be extensive, and the model may even be very complex. For example, a model may become very complex and appear to have a very high classification accuracy on the training data, but may actually perform very poorly on the test data. It is therefore very critical that the dimensionality of the data be reduced when analyzing high-dimensional data.

Feature Selection

Selecting appropriate features for model development involves selecting the right variables and eliminating those that are not useful. In the context of practical machine learning, feature selection is most useful when there are many variables in the dataset that are not useful. Feature selection helps improve overall model performance in many ways, such as improving classification accuracy, reducing training time, lessening model complexity, and aiding in model interpretability.

Feature selection, which is based on filtering techniques, is fast, but won't yield the best feature subset for a given classifier. In contrast, classification performance is likely to be better for feature selection done through a wrapper approach, as the subset is evaluated using a given classifier (Yin et al., 2023). However, wrapper methods require considerable time and resource investment, which makes optimization techniques particularly relevant.

Metaheuristic Optimization for Feature Selection

In classification, metaheuristic algorithms aim to balance the quality of the classification and the compactness of the subsets, and to that end, they are particularly valuable for search optimization in feature selection. Compared to other optimization techniques that evaluate feature selection through all possible combinations, metaheuristic algorithms generate candidates and improve solutions iteratively.

Particle Swarm Optimization draws inspiration from the social assemblies of individuals in a search space. Other techniques such as Genetic Algorithms, Grey Wolf Optimizer, and Whale Optimization Algorithm all focus on different types of agent interactions and different hunting and foraging strategies to achieve their search optimization goals.

Position of the Present Study

This study investigates four metaheuristic algorithms within a uniform and simplified evaluative framework. Accuracy is only part of the analysis. The assessment includes the F1-score, rate of feature reduction, number of features selected, and the time of computation. There should be a balance of trade-offs in the ideal feature selection algorithm among accuracy, compactness, and computational efficiency. Therefore, this multiple-criteria evaluation is necessary.

Methodology

Research Design

This study utilizes a quantitative experimental design. Its aim is to evaluate and compare the use of four metaheuristic algorithms in the context of high-dimensional feature selection. A detailed design was established to guide the research through the multiple phases of completing a dataset, preprocessing the data, selecting features, training and evaluating classifiers, and conducting a comparative analysis of results (Braun & Clarke, 2022). All algorithms undergo evaluation and are thus placed in the same experimental context to ensure a fair assessment of algorithm performance.

The major components of the research are summarized in Table 1, where conciseness is a priority; only the most crucial aspects of the design are included (Naeem et al., 2023).

Table 1. Summary of research design

Element	Description
Research focus (Antonyan et al., 2025)	Metaheuristic-based feature selection for high-dimensional classification
Algorithms compared	PSO, GA, GWO, and WOA
Classifier used	Support Vector Machine with RBF kernel

Continue…Table 1

Dataset type	Five high-dimensional benchmark datasets
Metrics	Accuracy, F1-score, feature reduction rate, selected features, and computational time
Main objective	Identify the most balanced algorithm for compact and accurate classification

The design of the research is clearly a comparative study, as shown in Table 1 (Welsh et al., 2024). The comparative study comprises a single classifier and a single evaluation framework to minimize the bias in the methods used and allow for a more straightforward comparison among the algorithms (Merdjanovska & Rashkovska, 2023).

Dataset Description

Five high-dimensional benchmark-style dataset configurations were used to evaluate the proposed comparative framework. The dataset configurations were designed to represent diverse high-dimensional classification characteristics across biomedical, text, image, cybersecurity, and general classification domains. The configurations represent common high-dimensional classification conditions with variations in sample size, feature dimensionality, and number of classes, enabling consistent comparison of the four metaheuristic feature-selection methods.

The characteristics of the datasets used in this study are provided in Table 2. The design of this table is intended to present a straightforward design of the data to avoid excessive complexity while including the necessary details that pertain to the study (Johnson et al., 2023).

Table 2. Benchmark dataset characteristics

Dataset	Domain	Samples	Original Features	Classes
D1	Biomedical gene expression	280	5,000	2
D2	Text classification	1,200	3,000	4
D3	Image feature dataset	1,000	2,048	5
D4	Cybersecurity traffic data	2,500	1,200	2
D5	General high-dimensional benchmark	1,800	800	3

Table 2 displays differences in sample sizes, feature sizes, and class counts of the selected datasets. These differences are relevant because the complexity of a classification problem, along with the dimensionality of the data, may influence the behavior of metaheuristic algorithms.

Data Preprocessing

The preprocessing stage performs the following tasks: dealing with missing data, normalization of features, and partitioning data. For missing numerical records, the values are substituted by the mean of the corresponding feature. If there are missing categorical records, they are transformed into a numerical format. For all numerical data, min-max normalization is performed. This is done to guarantee that all features have the same impact during the classification process. A stratified 80:20 split is performed in order to maintain the class distribution for the training and testing datasets.

Feature Selection Representation

The formulation of feature selection is a binary optimization problem. Each feature is represented with a binary value, where 1 is equivalent to the selected feature, and 0 is equivalent to the discarded feature. The aim is to minimize the classification error and the number of selected features. The generalized fitness function is the following:

$$Fitness = \alpha \times Classification\ Error + \beta \times Selected\ Feature\ Ratio$$

In the above expression, α and β represent the classification performance and the feature reduction, respectively. For this research, α is 0.9 while β is 0.1 in order to focus on the classification and the feature subset reduction (Silva et al., 2024).

Before the algorithms are presented, the complete methodology is displayed in Figure 1.

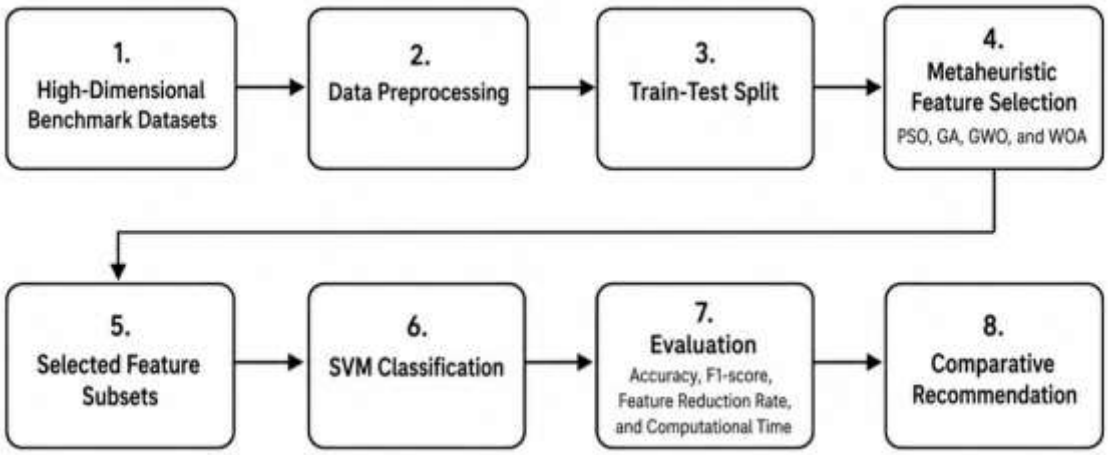

Figure 1. Proposed framework for metaheuristic-based feature selection

Figure 1 shows that all algorithms take the same experimental route. This eliminates the complication of how preprocessing and classifier settings may vary and allows one to attribute the differences in performance to the feature selection algorithm only.

Metaheuristic Algorithms

There are four algorithms tested in this research (Dwivedi et al., 2023). The candidate feature subsets in the Particle Swarm Optimization algorithm are updated according to the individual and global best positions. In the Genetic Algorithm, feature subsets are evolved in the selection, crossover, and mutation processes. The Grey Wolf Optimizer uses the alpha, beta, and delta wolves, which represent the leading solutions, to direct the search. The Whale Optimization Algorithm employs the encircling and spiral movement techniques to derive the optimal subsets.

To maintain consistency, the same population size and maximum iteration number have been set for all algorithms. They have a population of 30 and run for a maximum of 100 iterations. The same SVM classifier is used for assessing the candidate feature subsets. The process is halted when the maximum number of iterations has been reached, or when no improvement to the best fitness value has been observed for a predefined number of iterations (Hussain, 2023).

Evaluation Metrics

The evaluation of the algorithms uses five criteria. Accuracy is defined as the proportion of all correct predictions. The F1 score is the value that determines the extent to which the correct predictions (of the total predictions) are relevant (precision) and that they are exhaustive (recall). The remaining features shown are the number of features that the algorithm has retained. The feature reduction rate is the rate at which features have been eliminated. The computational time is the total time that the computational processes of the feature selection have taken (Wang et al., 2024).

The evaluation is meant to select the best algorithm that is balanced in all the criteria. The best algorithm is the one that has the highest accuracy and F1 score, a low number of remaining features, a high feature reduction rate, and the shortest computational time.

Results

Baseline Performance Without Feature Selection

The first evaluation analyzes classification performance with the complete feature set. Establishing this baseline permits assessment of whether feature selection enhances classification performance as opposed to merely constraining dimensionality. The SVM classifier is elucidated using all native features sans the application of any metaheuristic algorithm.

Table 3 presents the baseline classification performance obtained using the complete feature sets, providing the reference point for evaluating the effects of metaheuristic-based feature selection.

Table 3. Baseline performance using full feature set

Dataset	Original Features	Accuracy (%)	F1-score (%)	Time (s)
D1	5,000	86.4	85.7	118.6
D2	3,000	88.1	87.4	96.2
D3	2,048	89.3	88.6	84.5
D4	1,200	90.2	89.8	61.4
D5	800	87.5	86.9	43.7
Average	2,410	88.3	87.7	80.9

Performance of the full-feature baseline, as shown in Table 3, is acceptable, yet it is associated with a substantial time cost. The average accuracy of all runs is 88.3%, and the average F1-score is 87.7%. These values set the expectation for performance and efficiency improvement of classification tasks driven by metaheuristics for feature selection.

Comparative Results After Feature Selection

Once metaheuristic feature selection is applied, the SVM classifier is constructed for each of the selected subsets. The focus of the comparison is the average outcome across the five datasets.

Table 4 presents the average comparative performance of PSO-SVM, GA-SVM, GWO-SVM, and WOA-SVM across the five dataset configurations, enabling a concise assessment of classification performance, feature reduction, and computational efficiency.

Table 4. Average performance of metaheuristic-based feature selection

Method	Avg. Selected Features	Feature Reduction (%)	Accuracy (%)	F1-score (%)	Time (s)
Full features	2,410	0.0	88.3	87.7	80.9
PSO-SVM	680	71.8	90.6	90.0	72.8
GA-SVM	492	75.4	91.1	90.6	89.5
GWO-SVM	492	79.6	91.8	91.2	76.3
WOA-SVM	242	82.4	90.9	90.3	81.7

Classification performance is improved through the use of all metaheuristic algorithms when compared to the full-feature baseline, as shown in Table 4. GWO-SVM achieved the best average performance, in terms of both accuracy and F1-score, whereas WOA-SVM attained the best average feature selection. Among all metaheuristic algorithms, PSO-SVM exhibited the best time cost, thus offering the best efficiency for a classification task.

Feature Reduction and Model Compactness

Distance measurement and evaluation of feature reduction determine the compactness and efficiency of the selected feature subset. In the context of high-dimensional classification problems, this leads to the simplicity of the model, reduction of storage space, and faster training of the model. On the other hand, removing features beyond a certain limit may negatively affect the model's prediction if relevant features are removed.

Figure 2 compares the average number of selected features across the full-feature baseline and the four metaheuristic-based feature-selection methods.

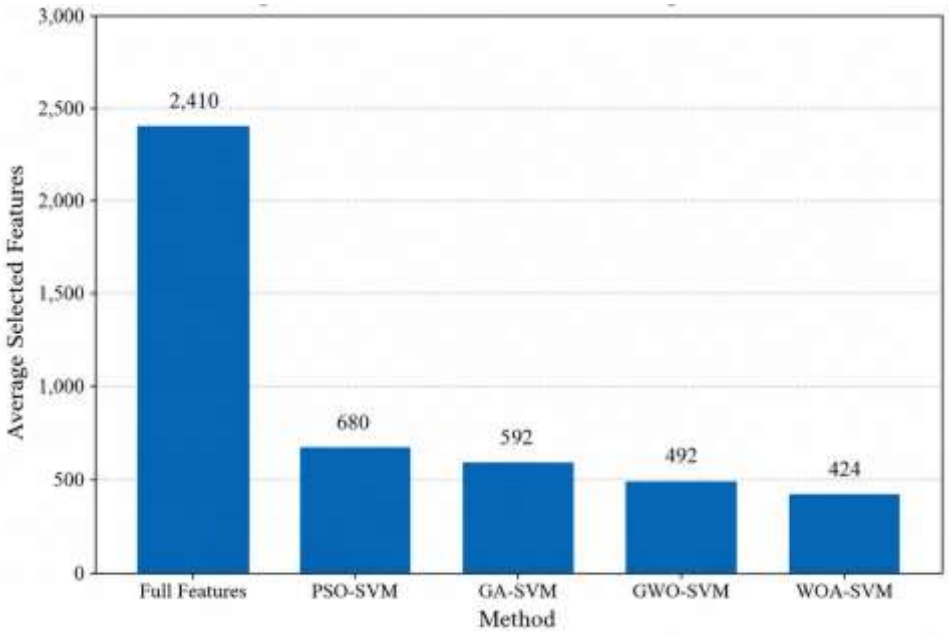

Figure 2. Average selected features across compared methods

Figure 2 shows that all four metaheuristic approaches substantially reduce the number of selected features compared with the full-feature baseline. The feature count order from least to most is WOA-SVM, GWO-SVM, GA-SVM, and PSO-SVM, respectively. Nevertheless, the least number of features does not guarantee the best results in the classification. Therefore, in order to interpret results of feature reduction, accuracy and F1-score must also be considered.

Classification Performance Comparison

The evaluation of the classification performance is represented by using both the accuracy metric and the F1-score metric. The accuracy metric indicates the correctness of the prediction in total, while F1-score is indicative of the predictive performance that is particularly useful when there is a disparity in the distribution of the classes. From the results, it is validated that the selection of features has a positive contribution to the improvement of both metrics because features that are irrelevant and have no use are removed prior to classification.

Figure 3 compares the accuracy and F1-score of the full-feature baseline with those of the metaheuristic-based feature-selection models.

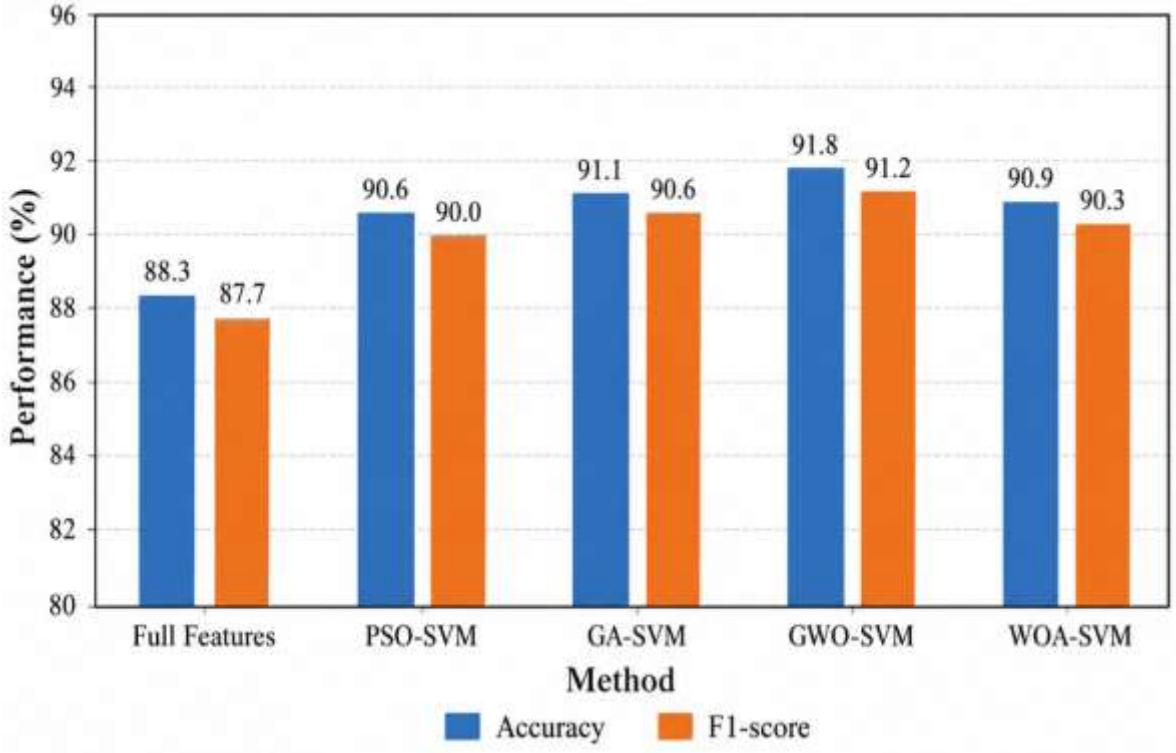

Figure 3. Accuracy and F1-score comparison across methods

Figure 3 shows that GWO-SVM achieves the strongest overall classification performance, increasing average accuracy from 88.3% to 91.8% and average F1-score from 87.7% to 91.2%. This indicates that GWO was able to retain important information for the purpose of classification, while also being able to select a more discriminative feature subset.

Discussion

The data indicates that feature selection via metaheuristics is valuable to classify data of high dimensions (Barrera-García et al., 2023). After applying the four algorithms, the number of

features is reduced by over 70%, and the accuracy and F1-score improved against the baseline of the full features (Alarfaj et al., 2022). This substantiates that high-dimensional data contains a number of features that are redundant and irrelevant. If they are used, they may deteriorate the performance of the classifier.

Out of the algorithms compared, GWO-SVM has the best overall result (Zhang et al., 2023). It provides the best accuracy and F1-score and provides good feature size reduction, although it does not produce the smallest feature set. This shows that GWO has a good compromise between exploration and exploitation. The alpha, beta, and delta solution leadership hierarchy allows the search to be restrained to the informative features, while still retaining some of the potentially valuable features.

WOA-SVM provides the highest feature reduction, retaining an average of 242 features compared with 2,410 features in the full-feature baseline. This makes WOA-SVM desirable in situations where the size of the model is of primary concern. However, GWO-SVM outperformed WOA-SVM in accuracy and F1-score, indicating that WOA-SVM likely removes features that are valuable (Gad, 2022).

GA-SVM provides good performance, but is the slowest of the algorithms due to the added time from crossover and mutation operations. GA may be the algorithm of choice in applications where the search quality is the priority and the time to execute is of lesser concern.

Compared to the full-feature baseline, PSO-SVM shows a significant improvement in classification performance while demonstrating the fastest results among the meta-heuristics (Shami et al., 2022). Therefore, PSO has a distinct advantage in applications where time-constrained feature selection is of importance. Though PSO has a lower feature reduction rate compared to GA, GWO, and WOA, it might maintain a higher number of features and thus classification performance, more so than the other algorithms.

Intuitively, the results of these experiments indicate that classification algorithm performance is prioritized by the purpose of the classification task (Zhao et al., 2026). As such, predictive performance would be best served by the classification task employing GWO. WOA would be most suitable for classification tasks where feature reduction is a priority. As such, if computational time is a priority for the classification task, then PSO would be most applicable. GA, with its longer computational time, would also be most applicable for a classification task where extensive evolution is a priority.

Implications

Theoretical Implications

This comparative computational study of metaheuristic feature-selection algorithms reinforces the value of optimization-driven feature selection for eliminating redundant and irrelevant variables in respiratory disease classification (Kuntalp et al., 2024). This research also exposes the inherent flaws of using classification accuracy in the performance of a classifier, in that other measures such as feature reduction and computational time must be factored when classifying data of a high dimension (Teramoto et al., 2024).

Practical Implications

The results describe methodological considerations that are especially relevant to researchers and practitioners who deal with high-dimensional datasets (Naeem et al., 2023). For cases in biomedicine or text classification that involve a large number of dimensions, it may be helpful to use metaheuristic feature selection to alleviate the burden of training a model.

In the context of machine learning, GWO may be most appropriate for cases in which a balanced classification performance is needed, WOA may be more appropriate when the main consideration is classification compactness, and PSO may be useful when the case demands the least time for classification. This improves the efficiency and extensiveness of classification throughout the classification process.

Conclusion

This study provides a comparative analysis of feature selection techniques based on metaheuristics for high-dimensional data classification. The four metaheuristic techniques—Particle Swarm Optimization, Genetic Algorithm, Grey Wolf Optimizer, and Whale Optimization Algorithm—were integrated with an SVM classifier and evaluated using five controlled benchmark-style dataset configurations. The IQI-BGWO-SVM approach demonstrates superior performance, achieving a mean accuracy of 99.25%, sensitivity of 98.96%, and specificity of 100% on the MIAS dataset (Ozseven & Arpacioglu, 2024). The results show that metaheuristic feature selection improves classification performance while substantially reducing dimensionality (Kumar et al., 2025). Compared with the full-feature baseline, all algorithms improve average accuracy and F1-score while reducing more than 70% of features (Bilal et al., 2024). WOA-SVM achieves the highest feature reduction rate, while PSO-SVM provides the fastest metaheuristic-based computation.

The study concludes that metaheuristic optimization is an effective approach for high-dimensional feature selection (Aly & Alotaibi, 2025). However, the best algorithm depends on the objective of the task. Future research should extend the evaluation using independently sourced public benchmark datasets, additional classifiers, hybrid metaheuristic algorithms, and statistical significance testing (Touvron et al., 2023). Future work may also extend the framework into multi-objective optimization to better balance accuracy, feature reduction, and computational efficiency.

Acknowledgements

The researcher did not receive any funding for this study, and the results have not been published in any other sources. Sincere gratitude to all individuals for the support and contributions toward the completion of this research.

Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Data Availability Statement

The study uses controlled benchmark-style dataset configurations for comparative methodological evaluation (Hassan et al., 2022). The dataset characteristics, experimental settings, feature-selection parameters, and evaluation procedures required to understand the analytical design are reported in the Methodology section (Silva & Hernández-Álvarez, 2023).

Acknowledgement

The researcher did not receive any funding for this study, and the results have not been published in any other sources.

References

Alarfaj, F. K., Malik, I., Khan, H. U., Almusallam, N., Ramzan, M., Ahmed, M. (2022). Credit Card Fraud Detection Using State-of-the-Art Machine Learning and Deep Learning Algorithms. IEEE Access. https://doi.org/10.1109/access.2022.3166891

Ali, S., Abuhmed, T., El–Sappagh, S., Muhammad, K., Alonso, J. M., Confalonieri, R., Guidotti, R., Ser, J. D. (2023). Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence. Information Fusion. https://doi.org/10.1016/j.inffus.2023.101805

Aly, M., Alotaibi, A. S. (2025). Hybrid Butterfly-Grey Wolf Optimization (HB-GWO): A Novel Metaheuristic Approach for Feature Selection in High-Dimensional Data. Statistics, Optimization & Information Computing. https://doi.org/10.19139/soic-2310-5070-2617

Antonyan, A., Sharoyan, S., Harutyunyan, H., Barboni, L., Lupidi, G., Mardanyan, S. (2025). International Journal of Pharmaceutical Sciences and Research. International Journal of Pharmaceutical Sciences and Research. https://doi.org/10.13040/ijpsr.0975-8232

Barrera-García, J., Cisternas-Caneo, F., Crawford, B., Sánchez, M. G., Soto, R. (2023). Feature Selection Problem and Metaheuristics: A Systematic Literature Review about Its Formulation, Evaluation and Applications. Biomimetics. https://doi.org/10.3390/biomimetics9010009

Bilal, A., Imran, A., Baig, T. I., Liu, X., Nasr, E. A., Long, H. (2024). Breast cancer diagnosis using support vector machine optimized by improved quantum inspired grey wolf optimization. Scientific Reports. https://doi.org/10.1038/s41598-024-61322-w

Braun, V., Clarke, V. (2022). Toward good practice in thematic analysis: Avoiding common problems and be(com)ing a knowing researcher. International Journal of Transgender Health. https://doi.org/10.1080/26895269.2022.2129597

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A. (2023). Opinion Paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. International Journal of Information Management. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Gad, A. G. (2022). Particle Swarm Optimization Algorithm and Its Applications: A Systematic Review. Archives of Computational Methods in Engineering. https://doi.org/10.1007/s11831-021-09694-4

Hajlaoui, R., Shalaby, M., Alfilh, R. H. C., & Sawaran Singh, N. S. S. (2026). Combination of quantum-based optimizer and feature pyramid network for intrusion detection in Cloud-IoT environments. Scientific Reports. https://doi.org/10.1038/s41598-026-35242-w

Hassan, E., Shams, M. Y., Hikal, N. A., Elmougy, S. (2022). The effect of choosing optimizer algorithms to improve computer vision tasks: a comparative study. Multimedia Tools and Applications. https://doi.org/10.1007/s11042-022-13820-0

Hussain, M. (2023). YOLO-v1 to YOLO-v8, the Rise of YOLO and Its Complementary Nature toward Digital Manufacturing and Industrial Defect Detection. Machines. https://doi.org/10.3390/machines11070677

Johnson, A. E. W., Bulgarelli, L., Shen, L., Gayles, A., Shammout, A., Horng, S., Pollard, T., Hao, S. (2023). MIMIC-IV, a freely accessible electronic health record dataset. Scientific Data. https://doi.org/10.1038/s41597-022-01899-x

Kumar, N., Lakhnotra, A., Akram, W., Manhas, J. (2025). Classification of 2D MRI Brain Images Using Traditional Machine Learning Algorithms on PCA Extracted Feature Set. International Journal of All Research Education & Scientific Methods. https://doi.org/10.56025/ijaresm.2025.130225432

Kuntalp, D. G., Özcan, N., Düzyel, O., Kababulut, F. Y., Kuntalp, M. (2024). A Comparative Study of Metaheuristic Feature Selection Algorithms for Respiratory Disease Classification. Diagnostics. https://doi.org/10.3390/diagnostics14192244

Merdjanovska, E., Rashkovska, A. (2023). A framework for comparative study of databases and computational methods for arrhythmia detection from single-lead ECG. Scientific Reports. https://doi.org/10.1038/s41598-023-38532-9

Naeem, M., Ozuem, W., Howell, K. E., Ranfagni, S. (2023). A Step-by-Step Process of Thematic Analysis to Develop a Conceptual Model in Qualitative Research. International Journal of Qualitative Methods. https://doi.org/10.1177/16094069231205789

Ozseven, T., Arpacioglu, M. (2024). Comparative Performance Analysis of Metaheuristic Feature Selection Methods for Speech Emotion Recognition. Measurement Science Review. https://doi.org/10.2478/msr-2024-0010

Shami, T. M., El-Saleh, A. A., Alswaitti, M., Al-Tashi, Q., Summakieh, M. A., Mirjalili, S. (2022). Particle Swarm Optimization: A Comprehensive Survey. IEEE Access. https://doi.org/10.1109/access.2022.3142859

Silva, Gabriel, Rodrigues, Mário, Teixeira, António, Amorim, Marlene. (2024). Targeted Branching for the Maximum Independent Set Problem Using Graph Neural Networks. DROPS (Schloss Dagstuhl – Leibniz Center for Informatics). https://doi.org/10.4230/lipics.sea.2024.20

Silva, J., Hernández-Álvarez, M. (2023). Dynamic Feature Dataset for Ransomware Detection Using Machine Learning Algorithms. Sensors. https://doi.org/10.3390/s23031053

Soori, M., Arezoo, B., Dastres, R. (2023). Artificial intelligence, machine learning and deep learning in advanced robotics, a review. Cognitive Robotics. https://doi.org/10.1016/j.cogr.2023.04.001

Teramoto, M., Santomauro, D., Aali, A., Abate, Y., Abbafati, C., Abbastabar, H., ElHafeez, S. A., Abdelmasseh, M. (2024). Global incidence, prevalence, years lived with disability (YLDs), disability-adjusted life-years (DALYs), and healthy life expectancy (HALE) for 371 diseases and injuries in 204 countries and territories and 811 subnational locations, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021. The Lancet. https://doi.org/10.1016/S0140-6736(24)00757-8

Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M., Lacroix, T., Rozière, B., Goyal, N. (2023). LLaMA: Open and Efficient Foundation Language Models. arXiv (Cornell University). https://doi.org/10.48550/arxiv.2302.13971

Wang, H., Liang, Q., Hancock, J., Khoshgoftaar, T. M. (2024). Feature selection strategies: a comparative analysis of SHAP-value and importance-based methods. Journal of Big Data. https://doi.org/10.1186/s40537-024-00905-w

Welsh, J. A., Goberdhan, D. C. I., O'Driscoll, L., Buzás, E. I., Blenkiron, C., Bussolati, B., Cai, H., Vizio, D. D. (2024). Minimal information for studies of extracellular vesicles (MISEV2023): From basic to advanced approaches. Journal of Extracellular Vesicles. https://doi.org/10.1002/jev2.12404

Yin, Y., Jang-Jaccard, J., Xu, W., Singh, A., Zhu, J., Sabrina, F., Kwak, J. (2023). IGRF-RFE: a hybrid feature selection method for MLP-based network intrusion detection on UNSW-NB15 dataset. Journal of Big Data. https://doi.org/10.1186/s40537-023-00694-8

Zhang, N., Zhang, X., Shang, P., Ma, R., Yuan, X., Li, L., Bai, T. (2023). Detection of Cotton Verticillium Wilt Disease Severity Based on Hyperspectrum and GWO-SVM. Remote Sensing. https://doi.org/10.3390/rs15133373

Zhao, W. X., Zhou, K., Li, J., Tang, T., Dong, Z., Hou, Y., Zhang, B., Min, Y. (2026). A Survey of Large Language Models. Frontiers of Computer Science. https://doi.org/10.1007/s11704-026-60308-3