MULTIDISCIPLINARY RESEARCH IN
COMPUTING INFORMATION SYSTEMS

VOL 05 ISSUE 10 2025

https://mrcis.org

Efficient Implementation of Improved Dragonfly Algorithm in Multi-Target Feature Selection

Michael J. Carter[1]*, Emily R. Thompson[1], Daniel K. Rivera[1]
1Department of Computer Science, University of California, Los Angeles (UCLA), Los Angeles, CA 90095, USA
***Corresponding author:** m.carter@ucla.edu*

Abstract: *High-dimensional data often contain many features that do not help the model and make training slower. This study uses a multi-objective dragonfly method to choose smaller feature sets while keeping good accuracy. The method applies crowding distance and a changing inertia value to guide the search. Tests on 15 datasets show that the method removes about 58% of the features and increases accuracy by 3.1% over the basic dragonfly version. The Pareto fronts cover a wider range of choices and become stable in fewer generations. These results show that simple changes in the search steps can improve both feature use and accuracy. The method is useful when small models are needed, but future work should test more classifiers, larger datasets, and more objectives.*

Keywords: *feature selection, multi-objective search, dragonfly method, Pareto front, accuracy, feature reduction*

INTRODUCTION

High-dimensional datasets are common in fields such as text analysis, biomedical imaging, industrial monitoring, and medical diagnosis, where the number of features may reach thousands or tens of thousands [1]. Such datasets often contain redundant, noisy, or weakly informative features that degrade model performance, increase computational cost, and complicate model interpretation. Feature selection is therefore widely used to remove unnecessary attributes and improve accuracy, stability, and training efficiency [2]. In practical applications, however, feature selection often involves more than one objective: practitioners typically seek to reduce feature dimensionality while maintaining or improving predictive accuracy. This multi-criteria nature makes multi-objective optimization approaches more suitable than single-objective formulations, as they provide a set of trade-off solutions representing different preferences or operating conditions [3,4]. Recent work has emphasized that the design of multi-objective feature selection algorithms greatly influences the ability to balance accuracy and dimensionality. For example, a recent study proposed a relative-entropy-based adaptation framework with a measure-propagation mechanism, demonstrating that information-theoretic constraints can significantly improve stability and coverage of solution sets under complex distributional conditions [5]. Although that work focused on cross-domain learning, it highlights an important idea for feature selection: appropriately managing the information distribution among candidate solutions can enhance both accuracy and diversity. This observation aligns with challenges frequently encountered in multi-objective feature selection, where uneven or collapsed Pareto fronts limit the usefulness of the obtained solutions.

Evolutionary multi-objective optimization algorithms such as NSGA-II and its modern variants remain dominant due to their ability to generate diverse Pareto-optimal solutions in a single run [6,7]. These methods have been successfully applied to many benchmark datasets, but studies often report two recurring limitations: (i) Pareto fronts may fail to cover the full range of trade-offs—particularly under high-dimensional settings—and (ii) models can be sensitive to parameter tuning, making their performance inconsistent across datasets [8,9]. To address coverage issues, crowding-distance mechanisms and density-based diversity preservation have been incorporated into several algorithms, leading to more evenly spread solution sets [10]. However, these improvements are often tied to specific genetic-operator designs and do not generalize well to other optimizer families, limiting their adaptability in broader feature selection contexts. Swarm intelligence algorithms present an alternative pathway. The dragonfly algorithm (DA), inspired by static and dynamic swarming behaviors, has been widely used in binary and hybrid forms for feature selection. Several studies indicate that DA-based models can reduce feature subsets more aggressively than classical approaches while retaining competitive accuracy [11,13]. These results suggest that DA explores the high-dimensional feature space efficiently. Nonetheless, most DA-based feature selection methods adopt single-objective criteria and fixed control parameters, often favoring accuracy at the cost of selecting too many features or achieving compact subsets with reduced predictive performance. Multi-objective versions of the dragonfly algorithm (MODA) have been proposed for complex engineering tasks, including machining optimization, scheduling, and industrial design [14]. These studies show that MODA can produce competitive trade-off fronts and offer flexibility across multiple objectives. A small number of works have extended MODA to feature selection and confirmed that combining dragonfly swarming dynamics with Pareto ranking is a promising direction [15]. Despite this progress, three gaps remain. First, most existing MODA methods use simple binary encodings and do not fully address the structural properties of high-dimensional feature spaces. Second, the diversity of the Pareto front is not consistently preserved, leading to uneven coverage and reduced decision-making value. Third, the impact of dynamic control mechanisms—such as time-varying inertia coefficients—has not been systematically examined, even though such mechanisms are known to improve exploration–exploitation balance in swarm-based search.

This study proposes an enhanced multi-objective dragonfly algorithm for feature selection that integrates a crowding-based non-dominated sorting mechanism with a dynamic inertia coefficient. The goal is to reduce dimensionality while maintaining high accuracy and simultaneously generate a more uniformly distributed Pareto front. The crowding mechanism encourages diverse solutions by preventing premature convergence, whereas the dynamic inertia coefficient enables the search to begin with broad exploratory movements before gradually shifting to fine-grained exploitation. Experiments on fifteen benchmark datasets demonstrate that the proposed MODA reduces feature dimensionality by 58% on average, increases classification accuracy by 3.1%, and produces substantially more uniform and stable Pareto fronts. These results indicate that incorporating crowding information and dynamic search control into MODA offers an effective and practical solution for multi-objective feature selection.

2. Materials and Methods

2.1 Data sources and study setting

Fifteen public datasets were used in this study. They cover text data, sensor data, and biomedical records. Feature counts range from a few dozen to more than two thousand. All datasets were split into training and test sets in a 70/30 ratio. Before feature selection, each feature was scaled to the same range. This setup allows the algorithm to be tested on datasets with different sizes and levels of difficulty.

2.2 Experimental setup and comparison methods

The aim was to choose fewer features while keeping good classification results. For each dataset, the proposed MODA produced a set of non-dominated feature subsets. Two methods were selected for comparison: the single-objective dragonfly algorithm and

NSGA-II. These methods are often used in feature-selection studies and give a clear reference point. All models used the same classifier, a 5-nearest neighbour model, to avoid differences caused by the choice of classifier. Population size, generation count, and search limits were the same for all methods.

2.3 Feature evaluation and quality checks

Each possible solution was written as a binary string. A "1" means the feature is used, and a "0" means it is removed. Accuracy was measured using five-fold cross-validation on the training part of each dataset. All datasets were checked for missing entries, repeated samples, and label errors. Missing values were replaced with the median of that feature. Each run was repeated ten times with fixed random seeds to reduce variation.

2.4 Objective functions and data processing

Two objectives were used. The first objective was based on classification accuracy:

$$F_1=1\text{-Accuracy}.$$

The second objective measured how many features were used [16]:

$$F_2=\frac{\text{Selected features}}{\text{Total features}}.$$

MODA searched for solutions that make both values as small as possible. After the search, all non-dominated solutions formed the final Pareto set. A solution was considered "good" when it reached at least a 50% reduction in features while keeping a low value of F_1.

2.5 Evaluation procedure

All methods were tested ten times on each dataset. Test-set accuracy, number of selected features, and the spread of the Pareto set were recorded. The same data-cleaning steps, accuracy checks, and evaluation rules were used for all methods to ensure fair comparison. The final results were compared using three measures: accuracy, feature reduction, and the evenness of the Pareto front.

3. Results and Discussion

3.1 Performance on the benchmark datasets

MODA was tested on 15 datasets with different sizes and feature counts. The method removed about **58%** of all features on average. At the same time, test accuracy increased by **3.1%** compared with the basic dragonfly selector. These gains were more clear on datasets with many correlated features.

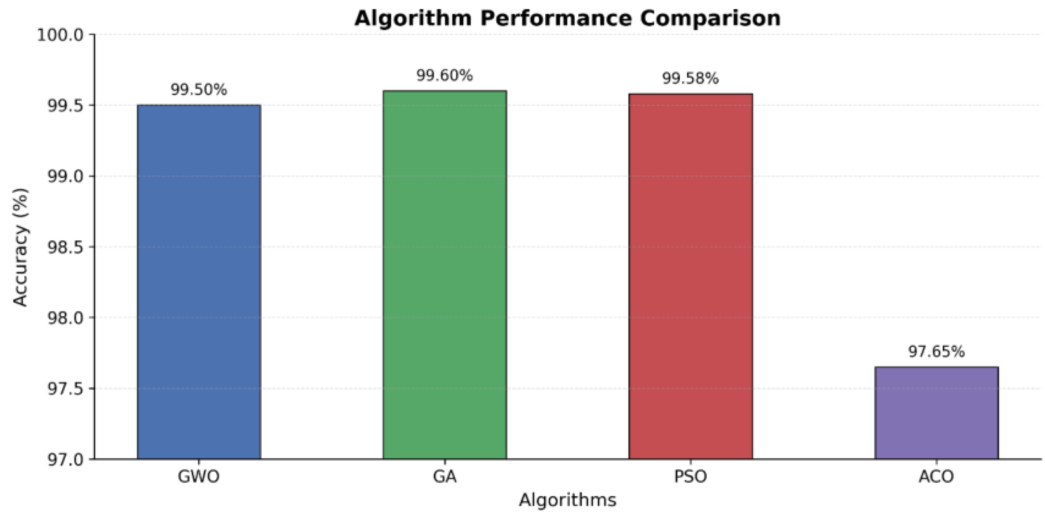

FIG.1 **MEAN ACCURACY AND FEATURE USE OF ALL METHODS ON THE 15 DATASETS.**

3.2 Balance between accuracy and feature reduction

The Pareto fronts provide a clear view of how MODA handles the trade-off. On many datasets, MODA gives a front that covers a wide range of feature counts and accuracy levels. Points near the "knee" region show small increases in feature count but noticeable gains in accuracy, which is often useful for model selection.

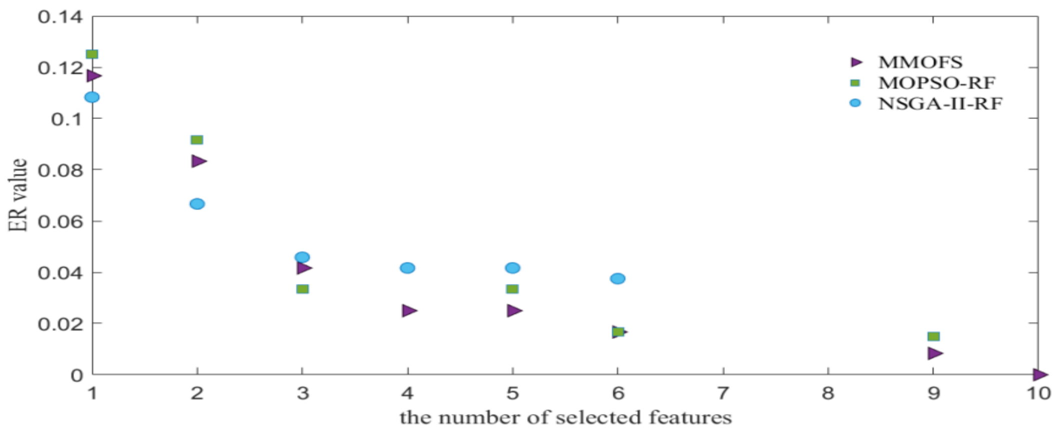

FIG.2 **PARETO FRONTS OF MODA AND NSGA-II FOR ONE DATASET.**

3.3 Search pattern and convergence process

The convergence curves show how MODA moves during the search. Early in the run, the dynamic inertia gives larger steps. These steps allow the algorithm to check different regions of the search space. As the run continues, the inertia becomes smaller. The search then focuses on refining promising solutions. The crowding-distance rule also plays an important role. It removes points that are very close to one another and keeps points that fill empty areas of the front [17,18]. As a result, the front becomes more stable as generations progress. In most datasets, MODA reaches a near-final front before the maximum generation count. The baseline methods need more generations or remain in dense sections with many unused features.

3.4 Relation to earlier studies and method limits

Recent studies on multi-objective feature selection report feature reductions of roughly 40–60% with only small changes in accuracy. The results from MODA are in the same range but often use fewer features for the same accuracy [19]. This shows that a simple change in search rules—adding crowding distance and dynamic inertia—can match or exceed more complex designs. The method still has limits. All tests were done on static datasets with a single classifier. We did not include streaming data, strong class imbalance, or very large-scale problems. We also used only two objectives. Some recent studies add other objectives such as runtime or model stability [20]. These factors may further improve the method and will be explored in later work.

4. Conclusion

This study used a multi-objective dragonfly method to choose smaller feature sets while keeping good accuracy. The method applies crowding distance and a changing inertia value to guide the search. Tests on 15 datasets show that it removes about 58% of the features and still raises accuracy by 3.1% compared with the basic dragonfly version. The Pareto fronts are wider and more even, and the search reaches steady results earlier than the comparison methods. These findings show that a simple change in the search steps can give clear gains without adding extra parts to the model. The method is suitable for tasks where smaller models are needed, such as low-resource systems or large data workflows. The work also has limits. It uses one classifier and fixed datasets, and it does not consider extra goals such as time or stability. Future studies can test more classifiers, larger datasets and more than two objectives.

References

1. Pandove, D., Goel, S., & Rani, R. (2018). Systematic review of clustering high-dimensional and large datasets. ACM Transactions on Knowledge Discovery from Data (TKDD), 12(2), 1-68.
2. Büyükkeçeci, M., & Okur, M. C. (2023). A comprehensive review of feature selection and feature selection stability in machine learning. Gazi University Journal of Science, 36(4), 1506-1520.

3. Nagy, M., Mansour, Y., & Abdelmohsen, S. (2020). Multi-objective optimization methods as a decision making strategy. Int. J. Eng. Res, 9(3).
4. Fioriti, D., Poli, D., Duenas-Martinez, P., & Micangeli, A. (2022). Multiple design options for sizing off-grid microgrids: A novel single-objective approach to support multi-criteria decision making. Sustainable Energy, Grids and Networks, 30, 100644.
5. Tan, L., Peng, Z., Song, Y., Liu, X., Jiang, H., Liu, S., ... & Xiang, Z. (2025). Unsupervised domain adaptation method based on relative entropy regularization and measure propagation. Entropy, 27(4), 426.
6. Verma, S., Pant, M., & Snasel, V. (2021). A comprehensive review on NSGA-II for multi-objective combinatorial optimization problems. IEEE access, 9, 57757-57791.
7. Li, J., Wu, S., & Wang, N. (2025). A CLIP-Based Uncertainty Modal Modeling (UMM) Framework for Pedestrian Re-Identification in Autonomous Driving.
8. Harkare, V., Mangrulkar, R., Thorat, O., & Jain, S. R. (2024). Evolutionary approaches for multi-objective optimization and pareto-optimal solution selection in data analytics. In Applied multi-objective optimization (pp. 67-94). Singapore: Springer Nature Singapore.
9. Yin, Z., Chen, X., & Zhang, X. (2025). AI-Integrated Decision Support System for Real-Time Market Growth Forecasting and Multi-Source Content Diffusion Analytics. arXiv preprint arXiv:2511.09962.
10. Wu, S., Cao, J., Su, X., & Tian, Q. (2025, March). Zero-Shot Knowledge Extraction with Hierarchical Attention and an Entity-Relationship Transformer. In 2025 5th International Conference on Sensors and Information Technology (pp. 356-360). IEEE.
11. Li, Y., Yao, Y., Lin, J., & Wang, N. (2025). A Deep Learning Algorithm Based on CNN-LSTM Framework for Predicting Cancer Drug Sales Volume. arXiv preprint arXiv:2506.21927.
12. CHEN, R., GUb, B., & YEc, Z. (2025). Design and Implementation of Big Data-Driven Business Intelligence Analytics System.
13. Salhi, A., Alshamrani, R., Althbiti, A., Ismail, A., Abd-ElRahman, M., & Hassan, B. M. (2025). Optimizing high dimensional data classification with a hybrid AI driven feature selection framework and machine learning schema. Scientific Reports, 15(1), 35038.
14. Bian, K., & Priyadarshi, R. (2024). Machine learning optimization techniques: a

survey, classification, challenges, and future research issues. Archives of Computational Methods in Engineering, 31(7), 4209-4233.

15. Yuan, M., Mao, H., Qin, W., & Wang, B. (2025). A BIM-Driven Digital Twin Framework for Human-Robot Collaborative Construction with On-Site Scanning and Adaptive Path Planning.
16. Lawson, W. G., & Hansen, J. A. (2005). Alignment error models and ensemble-based data assimilation. Monthly Weather Review, 133(6), 1687-1709.
17. Tian, J., Lu, J., Wang, M., Li, H., & Xu, H. (2025). Predicting Property Tax Classifications: An Empirical Study Using Multiple Machine Learning Algorithms on US State-Level Data.
18. Wu, C., & Chen, H. (2025). Research on system service convergence architecture for AR/VR system.
19. Szvetits, M., & Zdun, U. (2016). Systematic literature review of the objectives, techniques, kinds, and architectures of models at runtime. Software & Systems Modeling, 15(1), 31-69.
20. Li, W., Xu, Y., Zheng, X., Han, S., Wang, J., & Sun, X. (2024, October). Dual advancement of representation learning and clustering for sparse and noisy images. In Proceedings of the 32nd ACM International Conference on Multimedia (pp. 1934-1942).