Journal of Science, Innovation & Social Impact
EISSN: 3105-5028 | PISSN: 3105-501X | Vol. 1, No. 2 (2025)

Article

Application of the Improved Whale Optimization Algorithm in Random Forest Parameter Tuning

Oliver Bennett [1,*], Hannah Collins [1] and Samuel Turner [1]

[1] School of Computing, University of Leeds, Leeds, LS2 9JT, United Kingdom
* Correspondence: Oliver Bennett, School of Computing, University of Leeds, Leeds, LS2 9JT, United Kingdom

Abstract: Random Forest models rely on well-chosen hyperparameters, but common search methods often run slowly and give uneven results across datasets. This study introduces an Improved Whale Optimization Algorithm (IWOA) that updates inertia weight and step size during the search. The method was tested on twelve UCI classification datasets and compared with standard WOA-RF, PSO-RF, and DE-RF. IWOA-RF increased mean accuracy by 3.8% over WOA-RF and lowered training time by 11%. It also showed steadier results across repeated runs. The convergence curves showed that IWOA reached good solutions in fewer iterations. These results suggest that small changes in the update rules can make WOA more suitable for tuning RF models. The study also notes that only classification tasks were tested, and future work should include regression datasets and more complex RF settings.

Keywords: Improved Whale Optimization Algorithm; Random Forest; hyperparameter search; UCI datasets; inertia weight; step-size update; classification accuracy

Received: 24 November 2025
Revised: 03 January 2026
Accepted: 15 January 2026
Published: 20 January 2026

1. Introduction

Random Forest (RF) is widely applied in classification tasks because it models nonlinear patterns, handles heterogeneous data types, and remains robust under moderate noise with minimal preprocessing [1]. Despite these strengths, RF performance depends critically on its hyperparameters-such as the number of trees, maximum depth, and feature-sampling rate-which jointly affect predictive accuracy and computational cost. Classical tuning strategies including grid search and random search become inefficient when the search space grows large, as they require repeated model evaluations and often fail to capture complex interactions among parameters [2]. This limitation has motivated a growing body of research that applies metaheuristic optimization to automate RF hyperparameter selection and improve accuracy across diverse datasets [3]. Recent advances show that the design of the metaheuristic itself can substantially influence the stability and convergence of the optimization process. For example, one recent study introduced an improved dung beetle-based optimizer for RF tuning, demonstrating that enhanced update mechanisms and adaptive search strategies can yield more reliable hyperparameter selection compared with conventional approaches [4]. This result highlights a broader trend: performance gains increasingly depend not only on which metaheuristic is used, but on how its search operators, parameter control rules, and exploration-exploitation balance are designed.

Metaheuristic algorithms such as particle swarm optimization, evolutionary strategies, and nature-inspired methods have been widely used for feature selection, engineering design, and data-driven modeling [5,6]. Among these approaches, the Whale

Optimization Algorithm (WOA) has received particular attention due to its simple structure and minimal parameterization [7]. Although WOA performs well on many optimization tasks, its standard formulation may suffer from early convergence or rapid loss of diversity in complex, multimodal search spaces [8-10]. These limitations have motivated a series of WOA variants incorporating adaptive inertia weights, dynamic step-size control or Lévy-flight-based perturbations to improve global exploration and maintain stable convergence behavior [11,12]. Despite progress in applying WOA to machine learning models-including neural networks, SVMs, and ensemble methods-its improved variants remain underexplored for RF tuning. Existing work often employs basic versions of the algorithms, fixes internal parameters, or focuses mainly on final accuracy without examining convergence speed, training time, or robustness across datasets [13]. Furthermore, many studies evaluate improved optimizers only on synthetic benchmarks or small engineering problems rather than realistic classification datasets.

This study proposes an Improved Whale Optimization Algorithm (IWOA) integrating a dynamic inertia weight and an adaptive Lévy-flight-based step mechanism. The goal is to enhance global exploration during early iterations while preserving strong local refinement later, thus mitigating early convergence. The proposed IWOA is applied to RF hyperparameter tuning and evaluated on several UCI datasets. Performance is assessed in terms of classification accuracy, optimization stability, and training efficiency. The results show that the enhanced WOA design can significantly improve RF tuning while maintaining a simple and practical optimization framework for real-world applications.

2. Materials and Methods

2.1. Dataset and Study Description

This study used twelve UCI classification datasets, covering different data sizes and attribute types. The number of samples in these datasets ranged from 150 to more than 10,000, with both numerical and categorical variables. All datasets were publicly available and widely used for testing classification models. No additional preprocessing was performed apart from handling missing values and encoding categorical attributes. The goal of using several datasets was to examine the behavior of the proposed method under different data scales and feature distributions.

2.2. Experimental Design and Comparison Groups

The experiments compared the Improved Whale Optimization Algorithm applied to Random Forest (IWOA-RF) with three reference groups. The test group used RF hyperparameters tuned by IWOA. The first comparison group was the standard WOA-RF, which used the original WOA without inertia or step-size adjustments. The second group used Particle Swarm Optimization (PSO-RF), and the third group used Differential Evolution (DE-RF). These methods were selected because they represent common choices for hyperparameter tuning in machine learning. All models were trained with the same training and test splits for each dataset to ensure fair comparison. Each experiment was repeated ten times to reduce the effect of randomness in both the optimizer and the RF model.

2.3. Measurement Procedures and Quality Control

Model performance was evaluated using accuracy, precision, recall, and F1 score. For each dataset, the training and testing sets were created with a fixed random seed to ensure consistent comparison across methods. A five-fold cross-validation was carried out to check the stability of the results. To reduce noise in the evaluation, each optimizer was run with the same population size, maximum iteration number, and stopping criteria. All failed runs, such as those caused by numerical overflow or invalid parameter settings, were removed and repeated to keep the comparison fair. During analysis, we inspected the convergence curves of each method to identify early stagnation or unstable oscillations.

2.4. Data Processing and Model Formulation

Before optimization, input features were normalized to zero mean and unit variance. RF accuracy was used as the objective function for the optimizer. For a given set of hyperparameters, the fitness value was defined as:

$$f=1-\text{Accuracy},$$

So that the optimizer aimed to minimize f. To examine the error behavior, the mean absolute error (MAE) across repeated runs was computed as:

$$\text{MAE}=\frac{1}{n}\sum_{i=1}^{n}|y_i-\hat{y}_i|,$$

where y_i represents the true class and $\hat{y}_i$ denotes the predicted class. These formulas allowed the analysis of both classification accuracy and stability. All optimizers were implemented in Python, and RF was trained using the same entropy-based splitting rule for all experiments [14].

2.5. Statistical Analysis and Evaluation Criteria

Differences between IWOA-RF and the comparison groups were evaluated using paired t-tests with a significance level of $p < 0.05$. Mean accuracy, standard deviation, and convergence time were reported for each dataset. Performance consistency was examined by comparing the variance of the results across repeated runs. To assess convergence speed, the number of iterations required to reach 90% of the best accuracy was recorded. These evaluations helped determine whether the improvements came from higher accuracy, faster convergence, or more stable search behavior across datasets.

3. Results and Discussion

3.1. Overall Accuracy on UCI Datasets

Across the twelve UCI datasets, IWOA-RF achieved the highest mean accuracy in most tests (Figure 1). Its average accuracy was 3.8% higher than standard WOA-RF, and it also performed better than PSO-RF and DE-RF. The advantage was clearer on datasets with moderate dimension and noticeable class overlap.

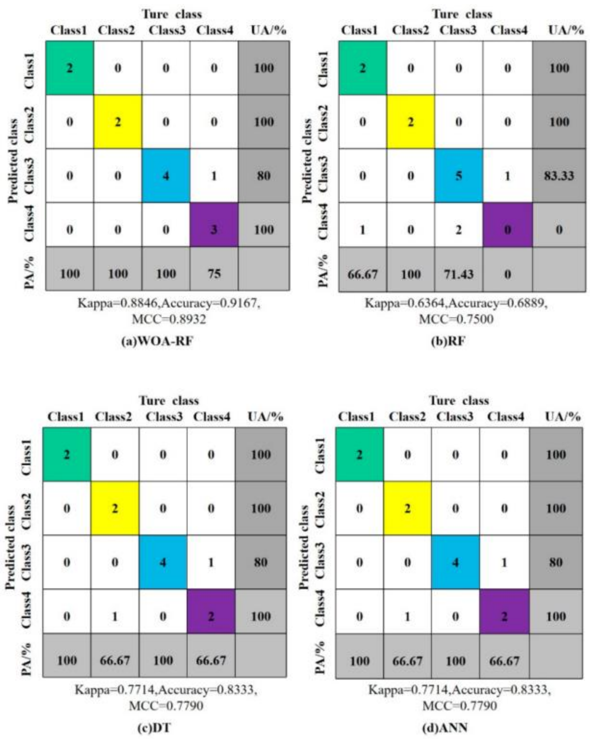

Figure 1. Mean accuracy and standard deviation for the four methods on the UCI datasets.

3.2. Convergence Speed and Computation Time

IWOA-RF also showed faster convergence (Figure 2). It reached near-optimal accuracy in fewer iterations and reduced training time by 11% compared with WOA-RF. Standard WOA-RF and DE-RF generally required more iterations to reach similar accuracy [15]. PSO-RF converged quickly on smaller datasets but showed larger variation across runs.

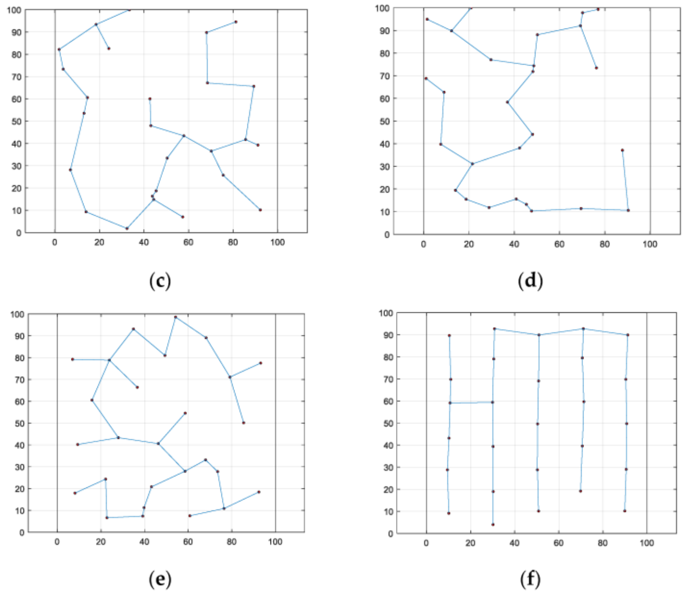

Figure 2. Convergence curves of the four methods averaged over ten runs.

3.3. Stability across Datasets and Comparison with Existing Work

IWOA-RF showed lower variance across repeated runs than the three comparison methods. This suggests that the adaptive inertia and step-size design helps keep the search stable and reduces dependence on the initial population. Previous studies on RF tuning have noted that many optimizers perform well on some datasets but become less stable when the data distribution changes [16,17]. In this study, the twelve datasets differed in size, feature type, and class balance. IWOA-RF remained stable under these changes. Compared with earlier RF-metaheuristic models developed for specific domains such as traffic analysis or blasting design, this work places more emphasis on broad testing and on analyzing both accuracy and consistency across datasets [18].

3.4. Influence of IWOA Components and Remaining Limitations

Further tests examined the role of the two main update steps. Removing the dynamic inertia term slowed down convergence and lowered accuracy. Removing the adaptive step-size update increased the number of runs that became trapped in local optima. These findings match earlier observations that multi-step WOA designs can help maintain search diversity and improve local refinement [19,20]. This study still has limits. Only twelve UCI datasets were used, and the focus was on classification tasks. High-dimensional problems and regression tasks were not included. In addition, only a small set of RF hyperparameters was tuned [21]. Future work should examine more complex RF settings, larger datasets, and comparisons with newer optimizers, while also studying how population size and stopping rules influence IWOA performance [22].

4. Conclusion

This study introduced an Improved Whale Optimization Algorithm for adjusting Random Forest hyperparameters. The method uses a dynamic inertia weight and a simple step-size update to guide the search. Across twelve UCI datasets, it reached higher accuracy than standard WOA-RF, PSO-RF, and DE-RF, and it also needed fewer iterations to approach a stable result. The repeated runs showed smaller spread in accuracy, which suggests that the update rules help keep the search steady. These findings show that a light change in the update process can make WOA more suitable for routine machine-learning tasks without adding extra tuning steps. The study still has limits, as only classification datasets and a small group of RF parameters were tested. Future work should include regression tasks, larger datasets, and more complex RF settings, and should also examine whether the same ideas work for feature-selection problems or other ensemble models.

References

1. H. A. Salman, A. Kalakech, and A. Steiti, "Random forest algorithm overview," *Babylonian Journal of Machine Learning*, vol. 2024, pp. 69-79, 2024. doi: 10.58496/bjml/2024/007
2. J. Bergstra, and Y. Bengio, "Random search for hyper-parameter optimization," *Journal of Machine Learning Research*, vol. 13, no. 1, pp. 281-305, 2012.
3. S. Kaur, Y. Kumar, A. Koul, and S. K. Kamboj, "A systematic review on metaheuristic optimization techniques for feature selections in disease diagnosis: Open issues and challenges," *Archives of Computational Methods in Engineering*, vol. 30, no. 3, p. 1863, 2022.
4. L. Tan, X. Liu, D. Liu, S. Liu, W. Wu, and H. Jiang, "An improved dung beetle optimizer for random forest optimization," In *Proceedings of the 6th International Conference on Frontier Technologies of Information and Computer*, 2024, pp. 1192-1196. doi: 10.1109/icftic64248.2024.10913252
5. M. A. Rahman, R. Sokkalingam, M. Othman, K. Biswas, L. Abdullah, and E. Abdul Kadir, "Nature-inspired metaheuristic techniques for combinatorial optimization problems: Overview and recent advances," *Mathematics*, vol. 9, no. 20, p. 2633, 2021. doi: 10.3390/math9202633
6. J. Li, S. Wu, and N. Wang, "A CLIP-based uncertainty modal modeling (UMM) framework for pedestrian re-identification in autonomous driving," 2025. doi: 10.70711/aitr.v2i10.7149
7. M. Amiriebrahimabadi, and N. Mansouri, "A comprehensive survey of feature selection techniques based on whale optimization algorithm," *Multimedia Tools and Applications*, vol. 83, no. 16, pp. 47775-47846, 2024. doi: 10.1007/s11042-023-17329-y
8. J. Tian, J. Lu, M. Wang, H. Li, and H. Xu, "Predicting property tax classifications: An empirical study using multiple machine learning algorithms on US state-level data," 2025.
9. H. M. Mohammed, S. U. Umar, and T. A. Rashid, "A systematic and meta-analysis survey of whale optimization algorithm," *Computational Intelligence and Neuroscience*, vol. 2019, no. 1, p. 8718571, 2019.
10. Y. Li, Y. Yao, J. Lin, and N. Wang, "A deep learning algorithm based on CNN-LSTM framework for predicting cancer drug sales volume," *Preprint*, 2025. doi: 10.18063/csa.v3i1.912
11. A. M. Khedr, P. R. P. V, and R. R. Mostafa, "EWOSCA: An enhanced walrus optimizer-based secure clustering approach for IoT-based WSNs under adversarial contexts," *Neural Computing and Applications*, pp. 1-34, 2025.
12. R. Chen, B. Gu, and Z. Ye, "Design and implementation of big data-driven business intelligence analytics system," 2025. doi: 10.3233/atde250877
13. Z. Yin, X. Chen, and X. Zhang, "AI-integrated decision support system for real-time market growth forecasting and multi-source content diffusion analytics," *Preprint*, 2025.
14. S. Demir, and E. K. Şahin, "Liquefaction prediction with robust machine learning algorithms (SVM, RF, and XGBoost) supported by genetic algorithm-based feature selection and parameter optimization from the perspective of data processing," *Environmental Earth Sciences*, vol. 81, no. 18, p. 459, 2022. doi: 10.1007/s12665-022-10578-4
15. M. Yuan, H. Mao, W. Qin, and B. Wang, "A BIM-driven digital twin framework for human-robot collaborative construction with on-site scanning and adaptive path planning," 2025. doi: 10.1109/cvaa66438.2025.11193176
16. K. Bian, and R. Priyadarshi, "Machine learning optimization techniques: A survey, classification, challenges, and future research issues," *Archives of Computational Methods in Engineering*, vol. 31, no. 7, pp. 4209-4233, 2024. doi: 10.1007/s11831-024-10110-w
17. C. Wu, and H. Chen, "Research on system service convergence architecture for AR/VR systems," 2025.
18. W. Sun, "Integration of Market-Oriented Development Models and Marketing Strategies in Real Estate," European Journal of Business, Economics & Management, vol. 1, no. 3, pp. 45–52, 2025
19. W. Li, Y. Xu, X. Zheng, S. Han, J. Wang, and X. Sun, "Dual advancement of representation learning and clustering for sparse and noisy images," In *Proceedings of the 32nd ACM International Conference on Multimedia*, 2024, pp. 1934-1942. doi: 10.1145/3664647.3681402
20. S. Wu, J. Cao, X. Su, and Q. Tian, "Zero-shot knowledge extraction with hierarchical attention and an entity-relationship transformer," In *Proceedings of the 5th International Conference on Sensors and Information Technology*, 2025, pp. 356-360. doi: 10.1109/icsi64877.2025.11009253
21. S. Yuan, "Data Flow Mechanisms and Model Applications in Intelligent Business Operation Platforms", Financial Economics Insights, vol. 2, no. 1, pp. 144–151, 2025, doi: 10.70088/m66tbm53.

22. F. S. Gharehchopogh, M. Namazi, L. Ebrahimi, and B. Abdollahzadeh, "Advances in sparrow search algorithm: A comprehensive survey," *Archives of Computational Methods in Engineering*, vol. 30, no. 1, pp. 427-455, 2023. doi: 10.1007/s11831-022-09804-w