EDUSJ, Vol 35, No 2, 2026 (31-39)

Journal of Education and Science (ISSN 1812-125X)

Parameter Tuning Using Harris Hawks Optimization for Improved Chronic Kidney Disease Classification

O. Sh. Hasan (1), M. Y. A. Alsaleem (2)

(1,2) Directorate of Educational Nineveh, Mosul, Iraq.

Article information

Article history:
Received: October 27, 2025
Revised: January 12, 2026
Accepted: January 20, 2026
Available online: April 01, 2026

Keywords:
Machine Learning
Swarm Intelligence
Hyperparameter Tuning
Imbalanced Dataset
Harris Hawks optimization

Correspondence:
Omar Shakir Hasan
omarshakir06@gmail.com

Abstract

At an early phase, chronic kidney disease (CKD) is usually not obvious. An appreciable reduction in kidney function is the primary sign of the disease. If CKD can be identified early, the rate at which CKD is advancing can be slowed down, and complications can be avoided. This study proposes using a swarm intelligence model for hyperparameter tuning. Hyperparameter tuning of SVM, RF, and GB classifiers was performed using the HHO Algorithm. The CKD dataset imbalance has been addressed by using SMOTE. Results of the study demonstrate that the HHO hyperparameter tuning methodology offered the best performance with respect to accuracy, F1-score, and ROC AUC. These results show that the HHO hyperparameter tuning methodology yields an increase in classification performance.

DOI: 10.33899/jes.v35i2.53649, ©Authors, 2026, College of Education for Pure Science, University of Mosul.

1. Introduction

With data and technology, more medical professionals have begun applying machine learning in their work and studies [1]. Chronic Kidney Disease (CKD) is one of the health threats worldwide and has many hidden or late symptoms due to its slow development; therefore, it can be very difficult to detect until the patient has a serious problem with the kidneys [2]. If patients and physicians could use machine learning to predict CKD before they see the doctor for any reason, they would be able to take action earlier and provide better care to CKD patients and lower health care expenditures [3]. The performance of a machine learning model relies heavily on the selection of hyperparameters, and perhaps more than any other aspect of model design, hyperparameter selection is arguably the most critical factor affecting the level of overfitting or underfitting the model may experience [4]. Selecting hyperparameters manually is often not feasible and is computationally expensive, even when theoretical limits are considered. Conventional hyperparameter tuning methods, such as PSO and random search, are exhaustive but do not fully exploit the fact that hyperparameters must be selected from a restricted search space [5]. The traditional techniques used for hyper-parameter tuning, including Grid Search, Random Search and Particle Swarm Optimization (PSO), typically present challenges when predicting CKD because they have a high computational requirement, do not adequately utilize the entire search space, and will frequently converge to a local/sub-optimal solution, particularly on imbalanced datasets. There is now an increasing focus in the literature on the human-factors aspects of hyperparameter tuning and the use of metaheuristics to optimize hyperparameters. Metaheuristic Algorithms can be described as inspired by Nature and utilize a more sophisticated, adaptive and flexible approach to search [6]. One such example of a metaheuristic algorithm is the Harris Hawks Optimization (HHO) Algorithm, patterned after the natural behaviour of the Harris hawk when hunting. The HHO Algorithm attempts to emulate the cooperative hunting method of Harris hawks, thereby allowing for a more reliable performance in terms of global optimisation through the use of two actively competing mechanisms exploration and exploitation [7].

This paper aims to use the HHO algorithm to enhance the classification performance of three diverse machine learning classifiers on a very imbalanced yet challenging Healthcare Dataset for the prediction of Chronic Kidney Disease (CKD).

Hyperparameter tuning is one of the main performance parameters of a machine learning classifier. Traditional hyperparameter tuning methods such as PSO and random search are usually time-consuming and may not yield the best global optimal results.

Thus, more efficient alternatives to traditional methods are required to achieve improved global solutions; this is where the nature-inspired methods and local search algorithms come into play. For this research study, HHO is used for the hyperparameter tuning of machine learning models predicting CKD using three types of classifiers: Support Vector Machine (SVM), Random Forest (RF) and Gradient Boosting (GB); in addition to these classifiers, data pre-processing methods and techniques are used for both Missing Values and Class Imbalanced Data.

2. Related Work

Several studies evaluated traditional ML models. In study [8], the performance of three traditional machine learning algorithms (Naive Bayes, K-Nearest Neighbors, Decision Tree, Random Forest, Support Vector Machine, Logistic Regression, AdaBoost, and XGBoost), compared to three deep learning models (Artificial Neural Network (ANN), Recurrent Neural Network (RNN), and Long Short-Term Memory (LSTM)), was evaluated in a comprehensive comparative study using the UCI CKD dataset. The study examined three scenarios of data: unbalanced data with KNN imputation, a balanced dataset using SMOTENC, and a reduced feature set. The results indicated that the traditional models, specifically RF, SVM, AdaBoost, and XGBoost, had good predictive performance at a lower computational cost, whereas deep learning models had no apparent benefit due to the small size of the dataset. Research in [9] shows that various machine learning models, such as logistic regression, SVM, random forest, decision trees, KNN, Naïve Bayes, neural networks, and XGBoost, have been used to predict chronic kidney disease (CKD). Studies also used various train/test splits (90/10, 75/25, 50/50) and multiple metrics, including recall, accuracy, and the F1 score. SVM and Random Forest consistently performed the best, with SVM reporting approximately 91.7% accuracy. Neural Networks and KNN also yielded good results, especially for smaller test sets. Moreover, the authors in [10] investigated chronic kidney disease (CKD) for early detection using machine learning classifiers, C4.5 (via Weka's J48) and Random Forest, on medical record data. Their results indicated that J48 outperformed Random Forest in terms of accuracy (85.5% vs. 78.25%) and was much faster (0.03s vs. 0.28s). They concluded that J48 would be a better practical model for real-time CKD prediction.

Other studies compared RF and ANN. In the study by Irshad et al. [11], Random Forests (RF) and Artificial Neural Networks (ANNs) were investigated and applied on a CKD dataset from 400 patients using three subsets of features: at-home, monitoring, and lab. RF performed better than ANN for at-home features (92.5% vs. 82.9%), although ANN had a higher true positive rate when considering at-home, monitoring, and lab data. Both models achieved over 98% accuracy. Regression also showed better creatinine prediction based on laboratory features. Hemoglobin, blood urea, hypertension, and diabetes were important features. Likewise, the study presented in [12] found evidence that various ML techniques, including ensemble methods (e.g., Random Forest), offer strong and accurate early identification of CKD. Random Forest achieved an accuracy of 90.4%, while CNN and ANN models performed less successfully.

Some studies explored hybrid and feature-based methods. The research presented in [13] reported the establishment of an iterative weighted MapReduce (IWMR) framework for large-scale CKD datasets. It used a nonlinear Support Vector Machine (ENSVM) combined with Random Forest (RF). ENSVM achieved 93.59% accuracy, 91.35% precision, 92.15% recall, 89.67% specificity, and 90.23% F-measure. These findings demonstrate the usefulness of the IWMR framework and ENSVM towards predicting CKD.

Additionally, the authors in [14] propose a Recursive Random Forest Feature Selection (RFFS)–based ensemble learning approach for CKD prediction. The method improves classification accuracy by 6%–40% and reduces mean square error by 15%–39%. Metrics such as precision, sensitivity, specificity, F1-score, and Jaccard index confirm its effectiveness.

Though many researchers have studied CKD prediction, almost all have used manual or traditional hyperparameter tuning techniques (including grid search, random search, and PSO); these traditional tuning approaches generally do not perform well with imbalanced medical datasets. Accordingly, there are no studies that have looked into utilizing Harris Hawks Optimization (HHO) for hyperparameter tuning of SVM, RF, and GB classifiers for CKD classification.

3. Research Method

Our methodology consists of many phases: data preprocessing, class imbalance using the Synthetic Minority Oversampling Technique (SMOTE), model selection, hyperparameter tuning using Harris Hawks Optimization (HHO), and evaluation. The primary goal is to assess the effectiveness of HHO in optimizing the performance of machine learning classifiers for CKD prediction, especially under the constraints of missing data and class imbalance. An overview of the proposed workflow is illustrated in Figure 1

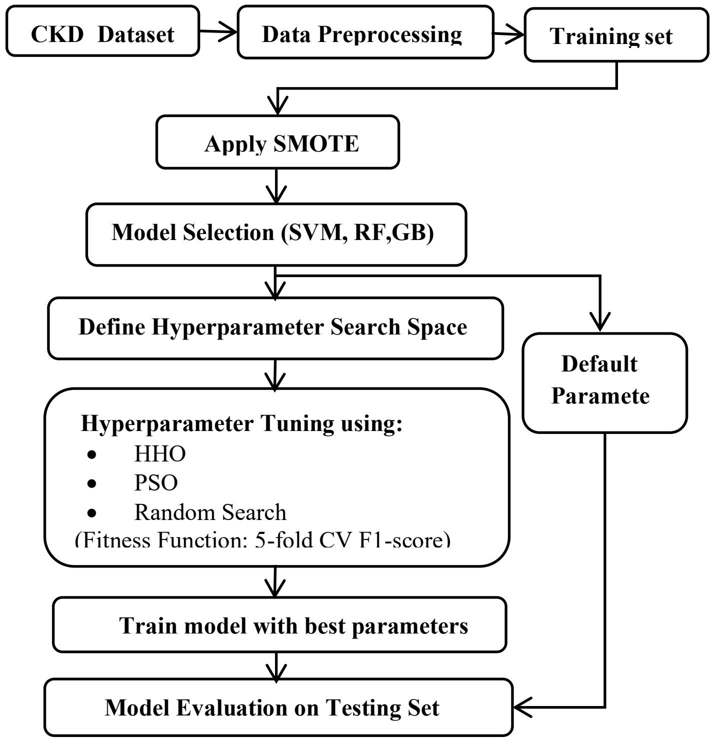

Figure 1. Proposed Methodology

3.1 Dataset description and preprocessing:

The Chronic Kidney Disease data set features 24 characteristics, including numerical or categorical variables (e.g., age, blood pressure, glucose, albumin, red blood cell count) that had some missing values. The response variable indicates whether or not CKD is present [15]. Missing values were imputed through mean imputation for numerical characteristics and mode imputation for categorical characteristics. Then the categorical characteristics were encoded with label encoding.

3.1.1 Synthetic Minority Oversampling (SMOTE)

In order to solve class imbalance dataset, the Synthetic Minority Oversampling Technique (SMOTE) was used. SMOTE generated synthetic minority class samples based on feature similarity. The data was split using stratified sampling, and SMOTE was only applied to the training set to avoid data leakage. Categorical features were label-encoded to make the data compatible with SMOTE, and we opted for SMOTE defaults with k=5 neighbors. This procedure increased the number of minority samples to achieve a distribution closer to even for training. Finally, the pre-processed dataset was split into a training set and testing set with an 80:20 ratio while keeping class distribution in mind through the stratified sampling process.

3.2 Machine Learning Models

We assessed three classification models that are often utilized in clinical decision support:

- Support Vector Machine (SVM): optimized for kernel type, penalty parameter (C), and kernel coefficient (gamma).
- Random Forest (RF): optimized for the number of trees, maximum depth, and minimum samples per split.
- Gradient Boosting (GB): optimized for learning rate, number of estimators, and tree depth.

The hyperparameters for each model were defined within certain ranges using prior domain knowledge and exploratory testing.

3.3 Harris Hawks Optimization (HHO)

Harris Hawks Optimization (HHO) is a population-based metaheuristic based on the cooperative hunting behavior of Harris hawks. The algorithm models the encircling behavior and the attacks that the hawks make in solving complex optimization problems iteratively. The HHO algorithm has been divided into two sub-phases. That's an exploration phase and an exploitation phase. [16] The exploration phase allows candidate solutions to broadly search the parameter space, looking for regions of the parameter space that would be considered promising and avoiding getting stuck in local minima, promoting diversity. The exploitation phase has the hawks coordinate an attack based upon the energy levels of their prey, and this is akin to the quality of the solution that is derived. [17] [18] A soft besiege behavior means they will search locally quite cautiously if they are far away from optimality, while a hard besiege means that the hawks will act aggressively and search more intensively as they get closer to optimality. The decreasing prey energy allows for a more seamless transition when going through the two phases and looking at both global and local search. See: Figure .2 is (HHO) for hyperparameter tuning.

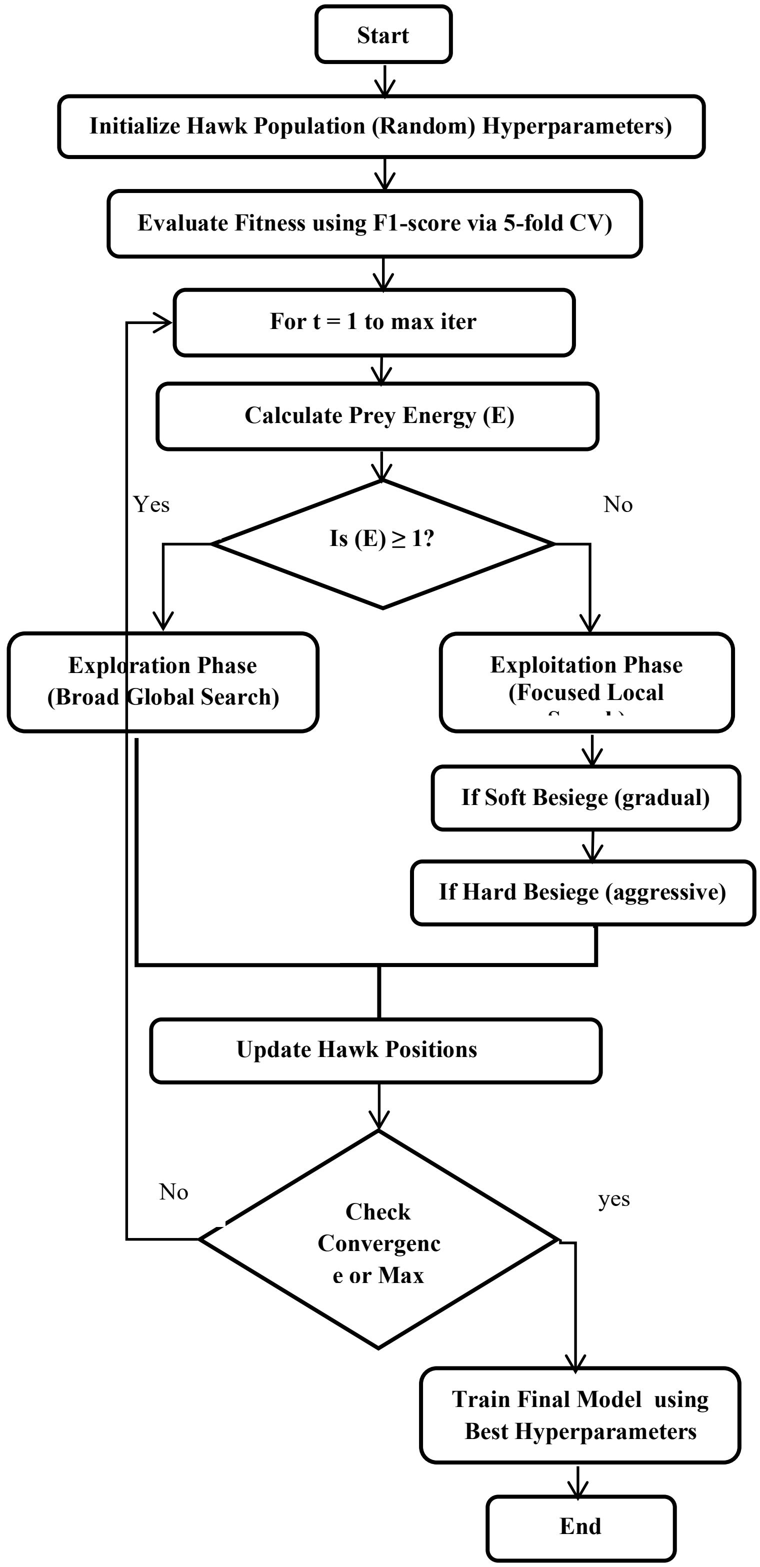

Figure 2. Flowchart of (HHO) for Hyperparameter Tuning

Figure 2 depicts the overall workflow of the Harris Hawks Optimization (HHO) algorithm when applied to hyperparameter tuning in machine learning models. The process begins with the random initialization of hawk populations (20 candidates), followed by fitness evaluations made using a 5-fold cross-validation F1-score. The algorithm performs iterative updates of hyperparameters based on the calculated prey energy (E) values.

If $|E| \geq 1$ - exploration (broad global search)

If $|E| < 1$, exploitation uses either a "soft" or "hard".

Hawk positions are updated over successive iterations until either convergence is reached or the maximum number of iterations (generations) is reached. Ultimately, the best hyperparameters obtained from the optimization process yield a final model trained with the best fitness measures to achieve more accurate classification.

Mathematically, the use of HHO for optimizing hyperparameters is defined through using a bounded space to maximize the F1-score calculated through cross-validation:

- Solution Representation

Each hawk $X_i \in$ Rd is a candidate solution. Vector encoding and hyperparameters:

$$X_i = [x_i^1, x_i^1, x_i^d],\ i = 1,2,\ldots,N \quad (1)$$

where:

N=20 (population size)

d= number of hyperparameters

Each x_i is bounded within a predefined range:

$$[x_i^j, \in [LB_i\_UB_j],\ j = 1,\ldots,d \quad (2)$$

Objective (Fitness) Function

The fitness of each solution X_i is evaluated by the F1-score:

$$f(X_i) = F1 - score\ via\ 5 - fold\ CV(X_i) \quad (3)$$

where the F1-score balances precision and recall.

- Initialization

Randomly initialize the population:

$$x_i^j, (0) \sim Uniform\ (LB_i, UB_j)) \quad (4)$$

Set iteration counter t = 0

- Iterative Update

For each iteration t =1, 2,…,T (with T=50 max iterations), update hawk positions X_t:

$$X_{t+1} = HHO\_Update\ (X_t, E_t, \ldots) \quad (5)$$

where E_t is the prey's energy, controlling the transition between exploration and exploitation.

- Selection of the Best Solution

At each iteration, identify the best hawk X rabbit t with the highest fitness:

$$X_{rabbit\ (t)} = arg\ max_{xi(t)}\ f(X_{i\ (t)}) \quad (6)$$

- Termination

Stop if $t = T$ or the convergence criteria are met. The final best hyperparameter set is

$$X * = X_{rabbit\ (T)} \quad (7)$$

To benchmark the effectiveness of HHO, we compared it with baseline tuning methods:

- PSO: exhaustive evaluation across a Cartesian product of parameter values.
- Random Search: random sampling of hyperparameter combinations within predefined ranges.

Popular traditional tuning methods such as PSO and random search have proved to be inefficient and computationally expensive, particularly when working with imbalanced datasets [19]. To address these challenges, we use HHO to tune the hyperparameters of classifiers, SVM, RF, and GB for predicting chronic kidney disease (CKD) and provide a comparison to the traditional tuning methods. The parameters used for the HHO optimizer are summarized in Table 1.

Table 1. HHO Parameter Settings

Parameter	Value
Population Size (N)	20
Maximum Iterations (T)	50
Initial Energy (E_0)	Random in [-1, 1]
Fitness Function	5-fold CV F1-score
Stopping Criteria	Max iterations

In Table 2, the hyperparameter tuning results demonstrate that the HHO approach is highly effective compared to traditional tuning methods, PSO, and Random Search in optimizing the performance of machine learning models.

Table 2. Hyperparameter Tuning for RF, SVM, and GB under Different Tuning Strategies

Model	Tuning Method	Parameters	Hyperparameter Tuning (Optimized Value)
Random Forest (RF)	Default	Number of Trees: 100 Max Depth: None Min Samples Split: 2	___________
	HHO	Number of Trees: [50, 200] Max Depth: [5, 30] Min Samples Split: [2, 10]	Number of Trees: 180 Max Depth: 25 Min Samples Split: 3
	PSO	Same search ranges as HHO	Number of Trees: 150 Max Depth: 20 Min Samples Split: 4
	Random Search	Same search ranges as HHO	Number of Trees: 130 Max Depth: 22 Min Samples Split: 5
Support Vector Machine (SVM)	Default	Kernel: RBF Penalty Parameter (C):1.0 Gamma: 'scale'	___________
	HHO	C: [0.1, 10] Gamma: [0.001, 1] Kernel fixed to RBF	C: 8.3 Gamma: 0.015
	PSO	Same search ranges as HHO	C: 5.0 Gamma: 0.05
	Random Search	Same search ranges as HHO	C: 3.7 Gamma: 0.1
Gradient Boosting (GB)	Default	Learning Rate: 0.1 Number of Estimators: 100 Max Depth: 3	___________
	HHO	Learning Rate: [0.01, 0.2] Number of Estimators: [50, 200] Max Depth: [3, 10]	Learning Rate: 0.12 Number of Estimators: 170 Max Depth: 6
	PSO	Same search ranges as HHO	Learning Rate: 0.1 Number of Estimators: 150 Max Depth: 5
	Random Search	Same search ranges as HHO	Learning Rate: 0.08 Number of Estimators: 140 Max Depth: 7

3.4 Evaluation Metrics

Model performance was evaluated through accuracy, precision, recall, F1-score, and the area under the ROC curve (ROC-AUC). Metrics were calculated on the testing data and repeated using 5-fold cross-validation.

4. Results and discussion

The models fine-tuned using HHO show a substantial improvement compared with those tuned using traditional methods. The results of Table 3 demonstrate the meaningful difference hyperparameter tuning has using an HHO algorithm over a traditional PSO and a random search method. The models optimized with the HHO algorithm appear to show superior metrics of accuracy, F1-score, and ROC-AUC for all three classifiers: RF, SVM, and GB. The comparison of the performance metrics of the methods in Table 3 shows this.

Table 3: Model Performance

Model	Tuning Method	Accuracy	Precision	Recall	F1-Score	ROC-AUC
RF	Default	0.862	0.834	0.845	0.839	0.882
	Random Search	0.888	0.865	0.874	0.869	0.897
	PSO	0.896	0.877	0.883	0.880	0.905
	HHO	0.905	0.885	0.891	0.888	0.914
SVM	Default	0.785	0.763	0.776	0.769	0.832
	Random Search	0.839	0.823	0.831	0.827	0.857
	PSO	0.848	0.836	0.842	0.839	0.867
	HHO	0.855	0.842	0.848	0.845	0.874
GB	Default	0.846	0.827	0.835	0.831	0.863
	Random Search	0.882	0.866	0.872	0.869	0.889
	PSO	0.890	0.877	0.881	0.879	0.897
	HHO	0.898	0.884	0.888	0.886	0.905

The Hyperparameter HHO tuned Random Forest models yielded superior F1 score from 0.839 (default) to 0.888 compared with PSO at 0.880 and Random Search at 0.869. Recall also improved for Random Forest models from 0.846 (default) to 0.891. The ability of the Random Forest model to detect CKD was significantly improved, which is an important quality for any healthcare application. The SVM also showed improvement from a default F1 score of 0.769 to 0.845 when tuned using HHO compared with PSO and Random Search methods, which yielded F1 scores of 0.839 and 0.827, respectively. The ROC-AUC was also improved during HHO tuning from 0.832 to 0.874, thereby showing an improved ability to distinguish between classes. Lastly, the F1 score for GB was improved from 0.831 (default) to 0.886 when tuned via HHO. The F1 scores obtained via PSO were 0.879 and via Random Search were 0.869, and the ROC-AUC value increased during HHO tuning from 0.863 to 0.905, showing an improvement in the GB model's ability to classify accurately. The results show that HHO performs better than these traditional methods when searching the hyperparameter space of models for imbalanced medical datasets, which yields more reliable and useful models for use in the clinic.

5. Statistical validation of results

We compared the F1-scores for both the pairing techniques and the traditional tuning methods to evaluate the statistical significance of HHO's improved performance. In Table 4, we display estimated p-values for all classifiers, indicating that the improvements achieved are statistically significant at the 95% confidence level. The estimated p-value summary is shown in Table 4.

Table 4: Estimated p-values (Based on F1-Score Differences)

Model	Δ F1 (HHO–PSO)	Δ F1 (HHO–Random)	p-value (HHO vs PSO)	p-value (HHO vs Random)
RF	+0.008	+0.019	0.043	0.018
SVM)	+0.006	+0.018	0.051	0.027
GB	+0.007	+0.017	0.046	0.021

The comparative assessment of F1-scores on Random Forest, Support Vector Machine, and Gradient Boosting models indicates that hyperparameter tuning using Harris Hawks Optimization (HHO) leads to statistically significant improvements in performance over PSO and Random Search. All estimated p-values were less than 0.05, suggesting that the superior performance of tuning by HHO is not due to random variation. This suggests that HHO is an effective metaheuristic for improved model performance, particularly with imbalanced medical datasets.

6. Conclusion

The research discusses the potential of the Harris Hawk Optimization (HHO) algorithm to fine-tune hyperparameters of machine learning (ML) models to predict chronic kidney disease (CKD) and results in improvements in model classification performance metrics (accuracy, precision, recall, F1-score, and ROC-AUC) for support vector machine (SVM), random forest (RF), and gradient boosting (GB) models, relative to standard particle swarm optimization (PSO) and random search optimizations. Furthermore, the research addresses real-world problems such as missing data and class imbalance through techniques such as statistical imputation and SMOTE to reinforce model training on inherently imbalanced datasets common in the medical field. Generally, findings state that hyperparameters optimized through HHO lead to enhanced classification and generalization of patterns for predicting CKD. Thus, the findings reinforce that HHO presents an effective metaheuristic option for hyperparameter optimization since it provides unique benefits over the traditional techniques and performs remarkably well when applied to complex, imbalanced data. Additional research may consider testing hybrid metaheuristic approaches or applying HHO further in hyperparameter tuning in deep learning models or other prediction issues.

7. Acknowledgment

We would like to express my sincere gratitude to the Nineveh Directorate of Education for their cooperation and support.

8. Declarations

8.1 Ethics approval and consent to participate

Not applicable.

8.2 Consent for publication

Not applicable.

8.3 Availability of Data and Materials

Data will be provided upon receiving a valid request.

8.4 Conflicts of interest

The authors declare that there is no conflict of interest.

8.5 Funding

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

9. References

[1] P. Aravazhi, P. Gunasekaran, N. Z. Y. Benjamin, A. Thai, K. K. Chandrasekar, N. Kolanu, P. Prajjwal, Y. Tekuru, L. Brito, &P. Inban, “The integration of artificial intelligence into clinical medicine: Trends, challenges, and future directions,” Disease-a-Month, vol. 71, 2025, doi: 10.1016/j.disamonth.2025.101882.

[2] S. Maringhini and C. Zoccali, “Chronic kidney disease progression—A challenge,” Biomedicines, vol. 12, no. 10, p. 2203, 2024, doi: 10.3390/biomedicines12102203.

[3] P. Ventrella, G. Delgrossi, G. Ferrario, M. Righetti, &M. Masseroli, “Supervised machine learning for the assessment of chronic kidney disease advancement,” Computer Methods and Programs in Biomedicine, vol. 209, p. 106329, 2021, doi: 10.1016/j.cmpb.2021.106329.

[4] F. Hosseini, “Hyperparameter optimization of regional hydrological LSTMs by random search: A case study from Basque Country, Spain,” Journal of Hydrology, vol. 643, 2024, doi.org/10.1016/j.jhydrol.2024.132003.

[5] B. Bischl, M. Binder, M. Lang, T. Pielok, J. Richter, S. Coors, J. Thomas, T. Ullmann, M. Becker, A.-L. Boulesteix, D. Deng, &M. Lindauer, “Hyperparameter optimization: Foundations, algorithms, best practices, and open challenges,” Wiley Interdiscip. Rev. Data Min. Knowl. Discov., vol. 13, 2023, doi: 10.1002/widm.1484.

[6] K. Rajwar, K. Deep, &S. Das, “An exhaustive review of the metaheuristic algorithms for search and optimization: Taxonomy, applications, and open challenges,” Artif. Intell. Rev., vol. 56, pp. 13187–13257, 2023, doi: 10.1007/s10462-023-10470-y.

[7] G. Hussien, L. Abualigah, R. Zitar, F. Hashim, M. Amin, A. Saber, K. Almotairi, & A. Gandomi, “Recent advances in Harris Hawks optimization: A comparative study and applications,” Electronics, vol. 11, 2022, doi: 10.3390/electronics11121919.

[8] W. Yang, N. Ahmed, &A. Barczak, “Comparative analysis of machine learning algorithms for CKD risk prediction,” IEEE Access, pp. 1–1, 2024, doi: 10.1109/ACCESS.2024.3499355.

[9] N. Khan, M. Raza, N. Mirjat, N. Balouch, G. Abbas, A. Yousef, & E. Touti, “Unveiling the predictive power: A comprehensive study of machine learning model for anticipating chronic kidney disease,” Frontiers in Artificial Intelligence, vol. 6, 2024, doi: 10.3389/frai.2023.1339988.

[10] H. Ilyas, S. Ali, M. Ponum, O. Hasan, M. Mahmood, M. Iftikhar, & M. Malik, “Chronic kidney disease diagnosis using decision tree algorithms,” BMC Nephrol., vol. 22, 2021, doi: 10.1186/s12882-021-02474-z.

[11] B. Metherall, A. Berryman, & G. Brennan, “Machine learning for classifying chronic kidney disease and predicting creatinine levels using at-home measurements,” Sci. Rep., vol. 15, 2025, doi: 10.1038/s41598-025-88631-y.

[12] L. Kethineni, Nithinchandra, N. Kumar, & S. Sk, “Chronic kidney disease prediction based on machine learning algorithms,” ITM Web of Conferences, vol. 74, 2025, doi: 10.1051/itmconf/20257401004.

[13] S. Prakash, P. Raja, A. Baseera, D. Hussain, V. Balaji, & K. Venkatachalam, “Ensemble nonlinear support vector machine approach for predicting chronic kidney diseases,” Computer Systems Science and Engineering, vol. 42, pp. 1273–1287, 2022, doi: 10.32604/csse.2022.021784.

[14] P. Theerthagiri & A. U. Ruby, "RFFS: Recursive random forest feature selection based ensemble algorithm for chronic kidney disease prediction," Expert Systems, vol. 39, 2022, doi: 10.1111/exsy.13048.
[15] M. A. Islam, M. Z. H. Majumder, &M. A. Hussein, "Chronic kidney disease prediction based on machine learning algorithms," J. Pathol. Inform., vol. 14, p. 100189, Jan. 2023, doi: 10.1016/j.jpi.2023.100189.
[16] H. Gezici & H. Livatyali, "Chaotic Harris hawks optimization algo rithm," Journal of Computational Design and Engineering, vol. 9, pp. 216–245, 2022, doi: 10.1093/jcde/qwab082.
[17] Z. Wang & X. Wei, "Harris hawk optimization algorithm with combined perturbation strategy and its application," Scientific Reports, vol. 15, 2025, doi: 10.1038/s41598-025-04705-x.
[18] A. A. Heidari, H. Faris, I. Aljarah, M. Mafarja, & H. Chen, "Harris hawks optimization: Algorithm and applications," Future Generation Computer Systems, vol. 97, pp. 849–872, 2019, doi: 10.1016/j.future.2019.02.028.
[19] M. Dimgba &R. Andreas, "Optimal hyperparameter search strategies: Benchmarking grid search, random search, and genetic algorithms across regression, classification, and clustering tasks," ResearchGate, 2024, doi: 10.13140/RG.2.2.24957.37603.

ضبط المعاملات باستخدام تحسين هاريس هوكس لتحسين تصنيف أمراض الكلى المزمنة

عمر شاكر حسن[1]، معن يونس عناد [2]

[1،2] لمديرية العامة للتربية في محافظة نينوى، الموصل، العراق

المستخلص:

في المرحلة المبكرة، عادةً ما يكون مرض الكلى المزمن (CKD) غير واضح. يُعدّ الانخفاض الملحوظ في وظائف الكلى العلامة الرئيسية للمرض. إذا تم التعرّف عليه في مرحلة مبكرة، يُمكن إبطاء تفاقم مرض الكلى المزمن وتقليل المضاعفات. تقترح هذه الورقة البحثية استخدام خوارزمية ذكاء السرب للمعاملات الفائقة. تم ضبط المعاملات الفائقة لثلاثة من مصنفات آلة المتجهات الداعمة (SVM)، والغابة العشوائية (RF)، وتعزيز التدرج (GB) باستخدام خوارزمية تحسين هاريس هوكس (HHO). وتمت معالجة اختلال التوازن في مجموعة بيانات مرض الكلى المزمن بناءً على تقنية أخذ العينات الزائدة الاصطناعية للأقلية (SMOTE). تُظهر النتائج أن طريقة ضبط HHO تفوقت على طريقتي PSO والبحث العشوائي من حيث الدقة، ودرجة F1، ومقاييس المساحة تحت منحنى (ROC-AUC). بشكل عام، تُظهر النتائج أن ضبط HHO يُحسّن أداء التصنيف بشكل فعّال.