IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 14, No. 5: October 2025, pp. 4332~4341
ISSN: 2252-8938, DOI: 10.11591/ijai.v14.i5.pp4332-4341

Heart disease prediction optimization using metaheuristic algorithms

Zaid Nouna[1], Hamid Bouyghf[1], Mohammed Nahid[1], Issa Sabiri[2]
[1]Laboratory of Engineering Sciences and Biosciences, Faculty of Sciences and Techniques of Mohammedia, Hassan II University of Casablanca, Mohammedia, Morocco
[2]Laboratory of Systems, Signals and Artificial Intelligence, ENSET Institute, Hassan II University of Casablanca, Mohammedia, Morocco

Article Info

Article history:

Received Jul 4, 2024
Revised Jul 21, 2025
Accepted Aug 6, 2025

Keywords:

Feature selection
Heart disease
Hyperparameter tuning
Machine learning
Metaheuristic
Preprocessing

ABSTRACT

This study explores metaheuristics hyperparameter tuning effectiveness in machine learning models for heart disease prediction. The optimized models are k-nearest neighbors (KNN) and support vector machines (SVM) using metaheuristics to identify configurations that minimize prediction error. Even though the main focus is utilizing metaheuristics to efficiently navigate the hyperparameter search space and determine optimal setting, a pre-processing and feature selection phase precedes the training phase to ensure data quality. Convergence curves and boxplots visualize the optimization process and the impact of tuning on model performance using three different metaheuristics, where an error of 0.1188 is reached. This research contributes to the field by demonstrating the potential of metaheuristics for improving heart disease prediction performance through optimized machine learning models.

Corresponding Author:

Zaid Nouna
Laboratory of Engineering Sciences and Biosciences, Faculty of Sciences and Techniques of Mohammedia
Hassan II University of Casablanca
Casablanca, Morocco
Email: zaid.nouna-etu@etu.univh2c.ma

1. INTRODUCTION

Responsible for over 70% of all fatalities, cardiovascular diseases (CVDs) are the leading global cause of illness and death [1], unhealthy behavior that result in obesity and overweight, high cholesterol, hypertension, and hyperglycemia raises the risk of heart disease [2]. Between 2010 and 2022, millions of adults aged 35 and older died from CVD, with heart disease accounting for the vast majority (75.6%) [3]. Early detection of heart disease is essential for minimizing health risks and averting cardiac arrests [4].

Machine learning is revolutionizing the field of disease prediction and has the potential to significantly improve healthcare. Machine learning enables machines to make predictions, group data (clustering), or automating decision-making [5]. These algorithms are highly influenced by their hyperparameters, hence the importance of finding the optimum setting. The process of creating the ideal model architecture with the best hyperparameter configuration is known as hyperparameter optimization (HPO) [6].

Optimization techniques as the grid search is an exhaustive search which can exercise to compute the optimal values of hyperparameters [7], and becomes computationally expensive for numerous hyperparameters and values. Random search is a hyperparameter tuning method explores a portion of the search space, resulting in lower computational cost and a level of accuracy improvement comparable to grid search [8]. Both methods offer robust starting point, however, for complex models with a multitude of hyperparameters, more sophisticated techniques are often required. Metaheuristics are problem-independent strategies, which can be applied to a broad range of problems [9], aiming to discover a solution that is near to the best answer at a lower cost, while exact approaches that explore the entire search space result in a very complex and expensive process [10]. They were developed to address the growing complexity of the problem, especially with the inclusion of uncertainties into the system, which may exceed the capabilities of conventional algorithms [11]. These algorithms are highly used in the medical field, for example, Sabiri *et al.* in [12] a modified particle swarm optimization (PSO) algorithm is used to maximize electrode's sensitivity and minimize cut-off frequency, Sabiri *et al.* in [13] used artificial bee colony (ABC) algorithm to optimize a complementary metal-oxide-semiconductor (CMOS) current mode instrumentation amplifier for biomedical applications.

Numerous research has been done specifically in the field of heart disease prediction. According to Al Bataineh and Manacek [14], an multi-layer perceptron (MLP)-PSO algorithm is proposed to predict heart disease using the Cleveland heart disease dataset (CHDD) and compared it to other machine learning algorithms. PSO is used in the training phase to find weights that minimize the error function as the optimization objective of the MLP network, and this technique outperformed other tested machine learning algorithms. Chandrasekhar and Peddakrishna [15] used GridSearchCV with five-fold cross-validation for six machine learning algorithms hyperparameters optimization. The soft voting ensemble classifier combining the six algorithms outperformed logistic regression and AdaBoost classifier on Cleveland and IEEE Dataport datasets. Ozcan and Peker [16] employed classification and regression tree (CART) supervised machine learning method to predict heart disease which has shown great results, the decision rules were extracted to rank the features based on importance in order to simplify the use for clinical purposes. Ogundepo and Yahya [17] considered both Cleveland dataset for building classification models and the Statlog data for results validation. Some of the bio-clinical categorical variables are found to be strongly associated with the heart disease conditions of the patients, and the support vector machines (SVM) achieved best predictive performances compared to the other tested algorithms. Research by Gupta and Sedamkar [18], genetic algorithm is used for feature selection and hyperparameter tuning on both SVM and neural network (NN) for heart disease prediction. C and γ are the optimized parameters in radial basis function (RBF) kernel for SVM, while no. of hidden layers, no. hidden nodes, learning rate momentum, and optimizer are the tuned parameters in MLP NN classifier. The results were better than utilizing Greadsearch for the same. This work aims to evaluate heart disease prediction performances by analyzing the impact of each step of the approach: pre-processing, features selection, metaheuristics validation and hyperparameters tuning comparing three different algorithms (PSO, grey wolf optimizer (GWO), differential evolution (DE)) on both k-nearest neighbors (KNN) and SVM classifiers.

The paper is structured as follows: section 2 highlights the different steps of the approach. Section 3 exposes the results and discussions of the used method in addition to the metaheuristics validation on known functions. Finally, a conclusion of this study is given in section 4 outlining promising directions for future research.

2. METHOD

In this section, the different steps from data collection to hyperparameters tuning are explored to expose how the study was conducted. First, the CHDD and the used preprocessing techniques are detailed. Then the machine learning models and the metaheuristic algorithms are presented, to finally explain how the hyperparameter tuning is performed.

2.1. Dataset collection

For analyzing coronary heart disease, the CHDD is widely considered the standard reference [19], 13 risk factors among 76 presented by the dataset as being associated to heart disease, are identified to be as significant contributors [20]. According to the UC Irvine Machine Learning Repository, the dataset comprises records for 303 patients including 13 input features for each patient : age, sex, chest pain type (cp), resting blood pressure (trestbps), serum cholesterol (chol), fasting blood sugar (fbs), resting electrocardiographic results (restecg), maximum heart rate achieved (thalach), exercise induced angina (exang), ST depression induced by exercise relative to rest (oldpeak), slope of the peak exercise ST segment (slope), number of major vessels colored by fluoroscopy (ca), thallium stress test result (thal), and an output/target which represents heart disease presence (Yes or No). The dataset contains both categorical and numerical data.

2.2. Preprocessing and features selection

Data will contain errors, out of range values, impossible data combinations, missing values or most substantially, data is not suitable to start a data mining process [21], the fundamental objective of data preparation is to increase the data's correctness and quality so that it is more suited for analysis [22]. In our case study, the dataset is divided into numerical and categorical features, to employ standardization and one-hot encoding techniques on the two types of features successively, then the features selection technique is performed. Standardization part aims to ensure all features are on a common scale, the one-hot encoding technique transforms categorical variables into a format (binary vector) that is suitable for machine learning models.

Features selection aims to identify the most relevant features for machine learning model and allows model performance improvement, model complexity and training time reduction. The performance of a wrapper method directly depends on the used classifier, since the variables selection is done based on machine learning algorithm performances with different subsets of features [23]. In terms of accuracy, the wrapper strategy can outperform the filter approach [24]. Figure 1 presents a conceptual diagram of the wrapper method for feature selection. In our study, we specifically used backward elimination wrapper method, which consists in starting with all the features, and removes the least significant one at each iteration until a stopping criterion is met. The main steps are:
- Forming a pool of candidate subsets.
- Model training for each subset and evaluation based on the chosen metric.
- Subset selection based on the performance on validation set.

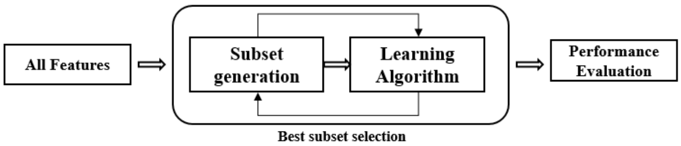


Figure 1. Wrapper method process

2.3. Machine learning models

Machine learning is the field dedicated to developing computer algorithms that gives computers the capacity to make predictions by learning from data and past experiences without human involvement [25]. In this study we have chosen KNN and SVM. The first algorithm classifies new data points by finding the 'k' closest training data points (neighbors) to the query being tested, then uses these neighbors to determine the classification [26]. On the other hand, in SVM, data points which are closest to the hyperplane are known as support vectors, and the margin represents the distance from the hyperplane to the closest data point from either of the classes it's separating. The primary goal is to identify the hyperplane that creates the largest possible margin to increase the probability that new, unseen data points will be classified correctly [27].

2.4. Metaheuristics

PSO, GWO, and DE algorithm are the selected algorithms for tuning in this study. Their main objective is to explore the search space, generate solutions and evaluate their performance. Figure 2 describes the metaheuristics during the hyperparameter tuning process, beginning from data preprocessing, to the best solution selection, with the number of iterations as the stopping criterion.

Velocity update is a crucial step in the PSO algorithm's search process. Each particle's new position is dynamically determined by combining its current velocity, its individual best-found position, and the best position discovered by the entire swarm. This integration of local and global information enables effective exploration and exploitation of the search space.

The GWO algorithm simulate the leadership structure and hunting behavior of grey wolves observed in nature. Four distinct classes of grey wolves alpha, beta, delta, and omega are utilized to represent the hierarchy of leadership. In the update process, the algorithm allows its search agents to update their position based on the location of the alpha, beta, and delta wolves; and attack towards the prey [28].

In the DE method, the initial population is created by randomly selecting values for each variable, the lower and upper bounds are defined by the user based on the specific addressed problem [29]. The update is made through succession of the mutation, crossover and selection operations. The mutation process creates new elements by randomly altering existing ones, then the crossing between the parents and the created elements and a selection operation to keep only the most suitable elements for the subsequent generation [30].

2.5. Hyperparameters tuning

HPO aims to find the optimal configuration of an unknown objective function, where search space is a hyperparameter space comprising categorical, discrete, and continuous variables [31]. Optimizing hyperparameters is a crucial step that ensures the appropriate selection of parameter values, for an improved classification performance [32]. Figure 3 shows the different steps for performance evaluation of the machine learning model, the step that is following the metaheuristic solution generation.

During the evaluation, each hyperparameter is assigned with its value determined by metaheuristic algorithm, then the model train/test phase is performed through cross-validation and the average error is calculated to evaluate the model's performance. While accuracy represents the ratio of correct predictions to the total number of predictions, the error indicates the ratio of incorrect predictions to the total number of predictions. More precisely, by minimizing the error we do increase the accuracy.

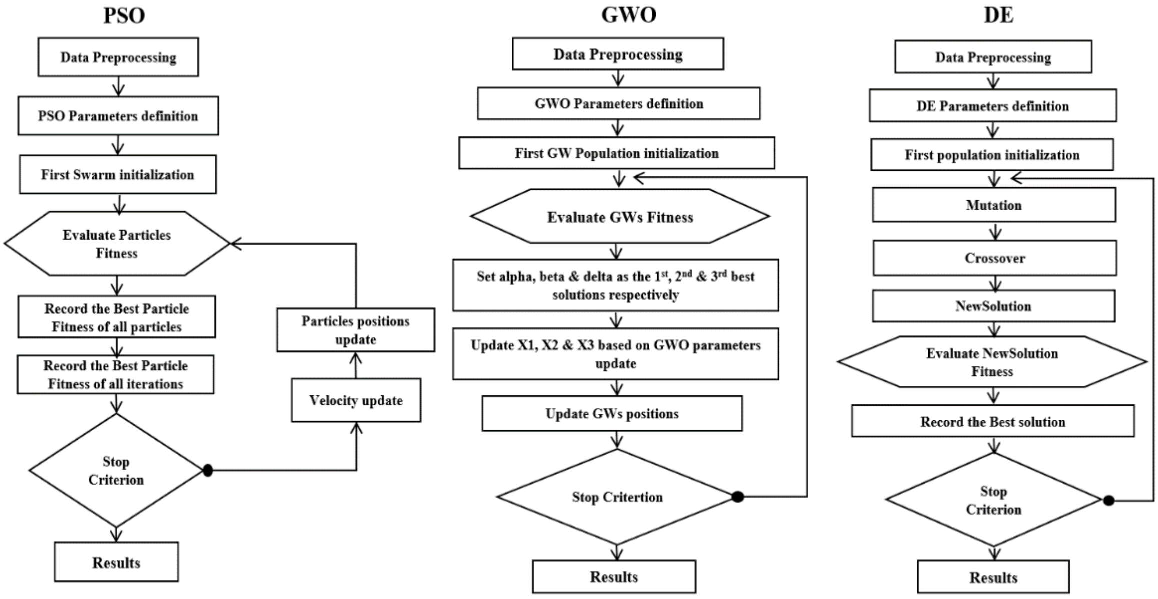


Figure 2. Flowcharts of PSO, GWO and DE algorithms

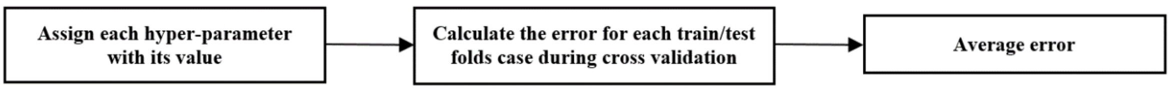


Figure 3. Evaluation phase

3. RESULTS AND DISCUSSION

In this section, the results of the different sections are exposed and discussed. First the pre-processing stage impact is highlighted, then the used metaheuristics are validated on known function, to finally use them on our case study, hyperparameters tuning using those algorithms. The results are presented in form of convergence curves and boxplots.

3.1. Pre-processing and features selection impact

Table 1 summarizes the error results obtained from our experiments. These results specifically highlight the impact of data pre-processing, comparing performance both before and after applying standardization and one-hot encoding techniques. The table provides a direct insight into the effectiveness of these pre-processing steps.

Table 1. Pre-processing impact

Algorithms	Before	After
KNN	0.4121	0.2268
SVM	0.1729	0.1689

Our findings reveal distinct impacts of pre-processing across different machine learning models. For the KNN classifier, pre-processing has reduced the error by approximately 0.21, while in SVM, it has reduced it by 0.004. It clearly shows that some machine learning models are more sensitive to specific pre-processing techniques than others, which can be very decisive in model's performances.

The wrapper method was applied for feature selection in a sequential manner as shown in Table 2. In the first step, aiming for the removal of just one feature, 'Chol' was identified as the optimal choice. Then, the method was re-applied to remove two features, and it converged on 'Chol' and 'Trestbps' as the selected pair for removal.

Table 2. Features selection impact

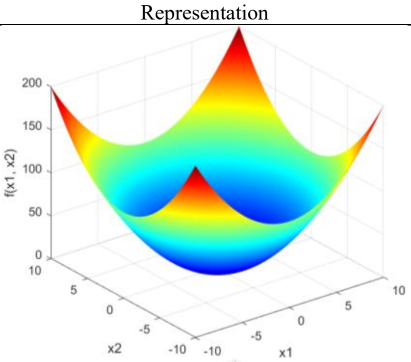
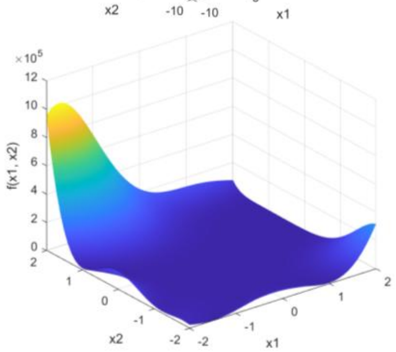
Algorithms	All features	1 removed feature	2 removed features
KNN	0.2257	0.2187	0.2099
SVM	0.1658	0.1634	0.1610

In addition to the slight improvement on the error as some features are removed, a significant benefit is the resulting decrease in model complexity. This reduction in complexity directly translates into faster training times, making the model more computationally efficient. Furthermore, a less complex model often exhibits better generalization capabilities, improving its performance on unseen data.

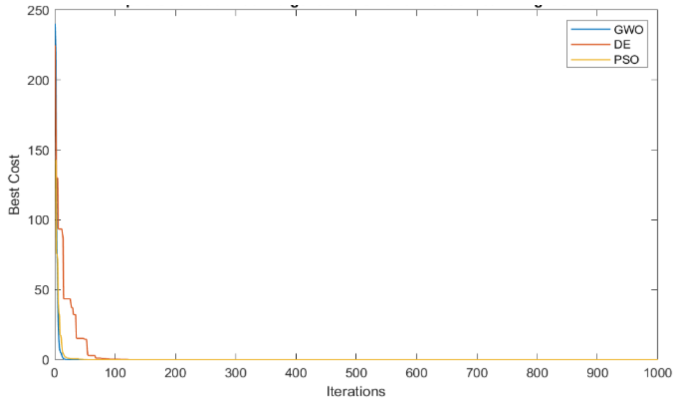
3.2. Metaheuristic algorithms validation

Metaheuristic algorithms validation is a crucial step that precedes their application to the actual problem. Sphere, Goldenstein-Price, and the personalized FindValues functions are the chosen one for validation. Table 3 exposes the range and global minimum of each function.

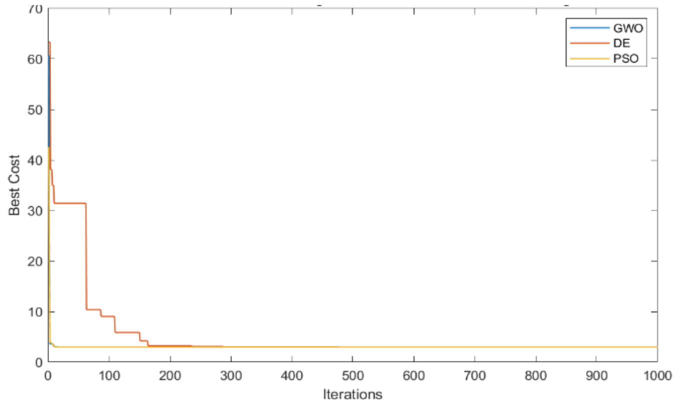
Table 3. Validation functions details

Functions	Representation	Range	Global minimum
Sphere		$Ub = 10$ $Lb = -10$	0
Goldenstein-Price		$Ub = 2$ $Lb = -2$	3
FindValues	-	$Ub = [100\ 700\ 50\ 200\ -400]$ $Lb = [0\ 200\ -200\ -100\ -900]$	0 (Error function)

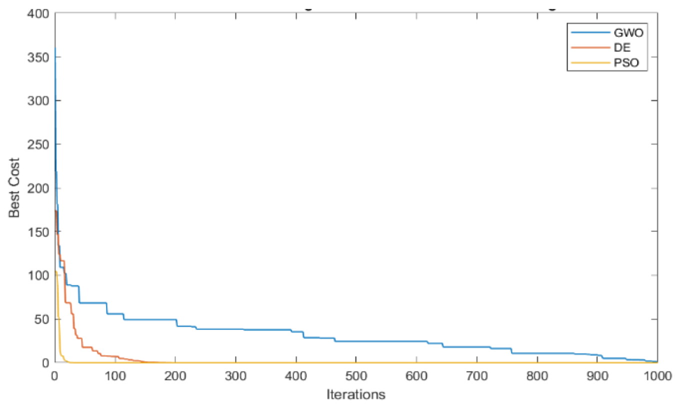
While Sphere and Goldenstein-price functions are known in the literature, the 'FindValues' function is personalized function as shown in Figure 4. It aims to find a pre-defined vector by reducing the errors between those values and the values selected by the algorithms, and the vector to be found in this case is [21 555 - 106 74 - 701]. The convergence curve of the three algorithms, GWO, DE, and PSO for the Sphere function is presented in Figure 4(a), where both DE and PSO have converged close to the global minima during the first 20 iterations and DE algorithm at about 90. Figure 4(b) displays the convergence using Goldstein-price function with global minima equal to three. Results are showing an earlier convergence compared to the previous function, regarding the same algorithms (GWO and PSO), at about 10 iterations. While it took about 180 iterations to reach good results using DE algorithm. The final function, FindValues in Figure 4(c) shows the convergence curve for an error function example, with a global minimum of zero. PSO converged in the first 10 iterations, then DE in about 150 iterations, while GWO started to reach good results at the end of the simulation near to 1,000 iterations.



(a)



(b)



(c)

Figure 4. Convergence behavior of metaheuristic algorithms on different validation functions of (a) Sphere, (b) Goldstein-Price, and (c) FindValues

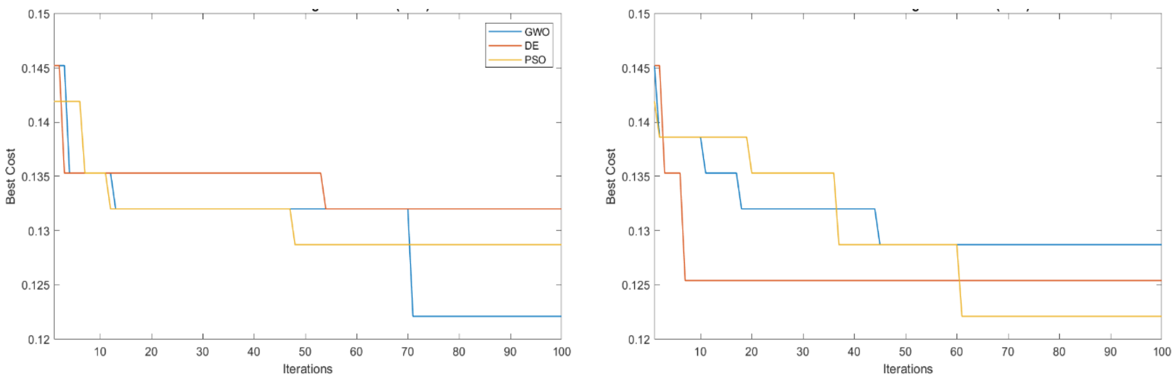
The three algorithms converge with different performances. DE algorithm is the algorithm with the higher number of iterations needed to final convergence in the Sphere and Goldstein-price functions, while it presented better performances compared to GWO in the FindValues function. On the other hand, PSO algorithm kept good performances for all three functions. All the three algorithms succeeded in the validation, and it is now possible to apply the metaheuristic algorithms in our case study, hyperparameter tuning, in which the function to evaluate would be the error of the machine learning models.

3.3. Hyperparameters tuning

PSO, GWO, and DE algorithms are validated on functions known in the literature. Table 4 presents both machine learning hyperparameters and metaheuristics parameters used during the search phase. The results are shown in Figure 5. Figure 5(a) presents the convergence curve during the error minimization by tuning KNN using metaheuristics. GWO has reached the best solution during the 100 iterations, at about 0.122, followed by PSO then DE algorithm in the final position. Figure 5(b) displays the tuning of SVM hyperparameters, where identical best solution has been found, using PSO this time, followed by DE algorithm with a 2nd best solution, then GWO by the end.

Table 4. Hyperparameters ranges and metaheuristics parameters

Algorithms	Hyperparameters	PSO	GWO	DE
KNN	Neighbors=[1:20]	nPop= 0; w=1;	a=2; wolvesNo=20	nPop=20; beta_min=0.2;
SVM	Kernel=RBF C= [0.1:50] gamma=[0.01:15]	wdamp=0.7; c1= 2; c2= 2		beta_max=0.9; pCR=0.1



(a) (b)

Figure 5. Errors convergence curves using metaheuristics algorithms for (a) KNN and (b) SVM

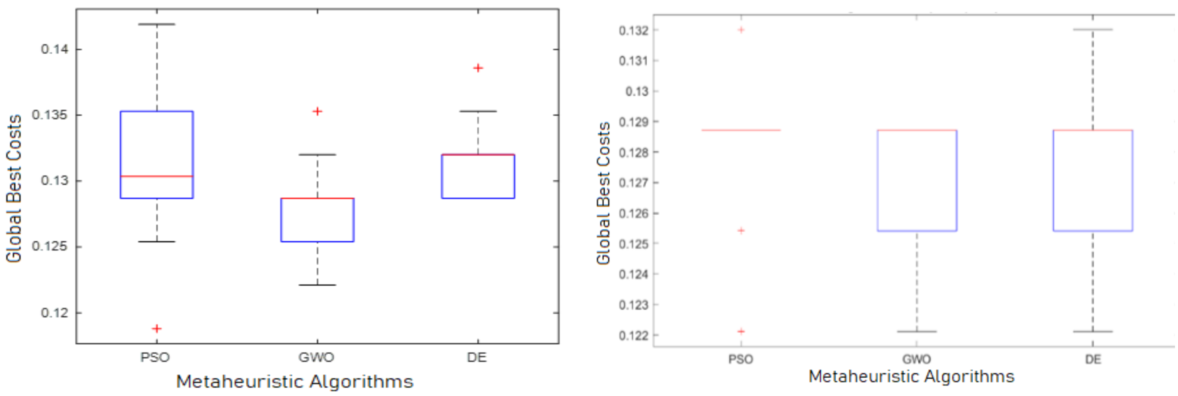
Table 5 resume the final results of metaheuristics algorithms for each machine learning model. They expose the best found solution using cross validation with its corresponding hyperparameters combination. The cross-validation technique splits to training and testing data folds randomly, it is used to give more realistic idea of how well the model generalizes to new information.

Table 5. KNN and SVM results by optimization algorithms

Algorithms	KNN results		SVM results		
	Error	N_Neighbors	Error	C	Gamma
PSO	0.1287	7	0.1221	22.96	9.16
GWO	0.1221	5	0.1287	20.28	3.58
DE	0.1320	3	0.1254	8.006	8.19

Figure 6 is the boxplots for the solutions found by each metaheuristic algorithm in each of the machine learning algorithm. The Boxplot is a robustness test that provides an information about solutions dispersion by running the algorithms for 10 runtimes (100 iterations each). A larger box indicates a wider spread of solutions, while a shorter box indicates a more concentrated set of solutions. Figure 6(a) is the dispersion of solutions for the KNN model, it shows a good spread over the 10 runs, with only one outlier per algorithm, with a minimum error found of 0.1188 and a maximum of about 0.142. Regarding SVM in Figure 6(b), 3 outliers are detected on the PSO algorithm, with a consistency of found solutions for many runs. The maximum error over the three algorithms is 0.132, and a minimum of 0.122, with reasonable spreads.

From the convergence curves in Figure 5, we observe that we could reach errors of about 0.12, which is great result for 100 iterations and 20 number of populations. On the other hand, the boxplots show the consistence of providing good results since the majority of the found errors are under 0.13 and the other minority rarely reach an error of 0.14, with a best solution of 0.1188 reached by the PSO-KNN. The non-visibility of all the quartiles boxplot is due to identical error result in different runs.



(a) (b)

Figure 6. Errors boxplots using metaheuristics algorithms for (a) KNN and (b) SVM

4. CONCLUSION

In this work aiming to improve heart disease prediction, we conclude that pre-processing and features selection are an essential step in machine learning algorithms and has a great impact on the prediction performances, whether to minimize complexity, computation time or for better results. On the other hand, the study confirms that metaheuristics for hyperparameters tuning is a promising approach. The used metaheuristics parameters are chosen in order to keep balance between exploration and exploitation phases, they have been used on the main hyperparameters of the machine learning algorithms, and the error minimization results are impressive. Exploring further metaheuristics parameters combinations, hybrid metaheuristics, and specific machine learning models hyperparameters (degree, distance), will improve performances to another level.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Zaid Nouna	✓	✓	✓	✓	✓	✓		✓	✓			✓		
Hamid Bouyghf	✓						✓			✓		✓	✓	
Mohammed Nahid	✓						✓			✓		✓	✓	
Issa Sabiri	✓			✓						✓		✓	✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are openly available in UCI Machine Learning Repository at https://archive.ics.uci.edu/dataset/45/heart+disease.

REFERENCES

[1] C. M. Bhatt, P. Patel, T. Ghetia, and P. L. Mazzeo, "Effective heart disease prediction using machine learning techniques," *Algorithms*, vol. 16, no. 2, 2023, doi: 10.3390/a16020088.
[2] P. Pitchal, S. Ponnusamy, and V. Soundararajan, "Heart disease prediction: improved quantum convolutional neural network and enhanced features," *Expert Systems with Applications*, vol. 249, 2024, doi: 10.1016/j.eswa.2024.123534.
[3] R. C. Woodruff *et al.*, "Trends in cardiovascular disease mortality rates and excess deaths, 2010–2022," *American Journal of Preventive Medicine*, vol. 66, no. 4, pp. 582–589, 2024, doi: 10.1016/j.amepre.2023.11.009.
[4] A. Dutta, T. Batabyal, M. Basu, and S. T. Acton, "An efficient convolutional neural network for coronary heart disease prediction," *Expert Systems with Applications*, vol. 159, 2020, doi: 10.1016/j.eswa.2020.113408.
[5] M. Mohammed, M. B. Khan, and E. B. M. Bashie, *Machine learning: algorithms and applications*, Boca Raton, Florida: CRC Press, 2016, doi: 10.1201/9781315371658.
[6] F. Abbas *et al.*, "Optimizing machine learning algorithms for landslide susceptibility mapping along the Karakoram highway, Gilgit Baltistan, Pakistan: a comparative study of baseline, Bayesian, and metaheuristic hyperparameter optimization techniques," *Sensors*, vol. 23, no. 15, 2023, doi: 10.3390/s23156843.
[7] E. K. Hashi and Md. S. U. Zaman, "Developing a hyperparameter tuning based machine learning approach of heart disease prediction," *Journal of Applied Science & Process Engineering*, vol. 7, no. 2, pp. 631–647, 2020, doi: 10.33736/jaspe.2639.2020.
[8] L. V.-Arias, C. Q.-López, J. G.-Coto, A. Martínez, and M. Jenkins, "Evaluating hyperparameter tuning using random search in support vector machines for software effort estimation," *PROMISE 2020 - Proceedings of the 16th ACM International Conference on Predictive Models and Data Analytics in Software Engineering, Co-located with ESEC/FSE 2020*, pp. 31–40, 2020, doi: 10.1145/3416508.3417121.
[9] B. Toaza and D. E.-Kiss, "A review of metaheuristic algorithms for solving TSP-based scheduling optimization problems," *Applied Soft Computing*, vol. 148, 2023, doi: 10.1016/j.asoc.2023.110908.
[10] S. Nematzadeh, F. Kiani, M. T.-Afshar, and N. Aydin, "Tuning hyperparameters of machine learning algorithms and deep neural networks using metaheuristics: a bioinformatics study on biomedical and biological cases," *Computational Biology and Chemistry*, vol. 97, 2022, doi: 10.1016/j.compbiolchem.2021.107619.
[11] A. M. Nassef, M. A. Abdelkareem, H. M. Maghrabie, and A. Baroutaji, "Review of metaheuristic optimization algorithms for power systems problems," *Sustainability*, vol. 15, no. 12, 2023, doi: 10.3390/su15129434.
[12] I. Sabiri, H. Bouyghf, and A. Raihani, "Optimal interdigitated electrode sensor design for biosensors using multi-objective particle-swarm optimization," *International Journal of Electrical and Computer Engineering*, vol. 13, no. 3, pp. 2608–2617, 2023, doi: 10.11591/ijece.v13i3.pp2608-2617.
[13] I. Sabiri, H. Bouyghf, A. Raihani, and B. Ouacha, "Optimal design of CMOS current mode instrumentation amplifier using bio-inspired method for biomedical applications," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 25, no. 1, pp. 120–129, 2022, doi: 10.11591/ijeecs.v25.i1.pp120-129.
[14] A. Al Bataineh and S. Manacek, "MLP-PSO hybrid algorithm for heart disease prediction," *Journal of Personalized Medicine*, vol. 12, no. 8, 2022, doi: 10.3390/jpm12081208.
[15] N. Chandrasekhar and S. Peddakrishna, "Enhancing heart disease prediction accuracy through machine learning techniques and optimization," *Processes*, vol. 11, no. 4, 2023, doi: 10.3390/pr11041210.
[16] M. Ozcan and S. Peker, "A classification and regression tree algorithm for heart disease modeling and prediction," *Healthcare Analytics*, vol. 3, 2023, doi: 10.1016/j.health.2022.100130.
[17] E. A. Ogundepo and W. B. Yahya, "Performance analysis of supervised classification models on heart disease prediction," *Innovations in Systems and Software Engineering*, vol. 19, no. 1, pp. 129–144, 2023, doi: 10.1007/s11334-022-00524-9.
[18] S. Gupta and R. R. Sedamkar, "Genetic algorithm for feature selection and parameter optimization to enhance learning on framingham heart disease dataset," *Intelligent Computing and Networking*, Singapore: Springer, pp. 11–25, 2021, doi: 10.1007/978-981-15-7421-4_2.
[19] P. R. Kumar, S. Ravichandran, and S. Narayana, "Ensemble classification technique for heart disease prediction with meta-heuristic-enabled training system," *Bio-Algorithms and Med-Systems*, vol. 17, no. 2, pp. 119–136, 2021, doi: 10.1515/bams-2020-0033.
[20] R. Perumal and K. Ac, "Early prediction of coronary heart disease from cleveland dataset using machine learning techniques," *International Journal of Advanced Science and Technology*, vol. 29, no. 6, pp. 4225–4234, 2020.
[21] S. García, J. Luengo, and F. Herrera, "Tutorial on practical tips of the most influential data preprocessing algorithms in data mining," *Knowledge-Based Systems*, vol. 98, pp. 1–29, 2016, doi: 10.1016/j.knosys.2015.12.006.
[22] N. Sirisha, M. Gopikrishna, P. Ramadevi, R. Bokka, K. V. B. Ganesh, and M. K. Chakravarthi, "IoT-based data quality and data preprocessing of multinational corporations," *Journal of High Technology Management Research*, vol. 34, no. 2, 2023, doi: 10.1016/j.hitech.2023.100477.
[23] T. M. Le, T. M. Vo, T. N. Pham, and S. V. T. Dao, "A novel wrapper-based feature selection for early diabetes prediction enhanced with a metaheuristic," *IEEE Access*, vol. 9, pp. 7869–7884, 2021, doi: 10.1109/ACCESS.2020.3047942.
[24] H. Das, B. Naik, and H. S. Behera, "A Jaya algorithm based wrapper method for optimal feature selection in supervised classification," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 6, pp. 3851–3863, 2022, doi: 10.1016/j.jksuci.2020.05.002.
[25] M. Soori, B. Arezoo, and R. Dastres, "Machine learning and artificial intelligence in CNC machine tools, A review," *Sustainable Manufacturing and Service Economics*, vol. 2, 2023, doi: 10.1016/j.smse.2023.100009.
[26] S. Uddin, I. Haque, H. Lu, M. A. Moni, and E. Gide, "Comparative performance analysis of K-nearest neighbour (KNN) algorithm and its different variants for disease prediction," *Scientific Reports*, vol. 12, no. 1, 2022, doi: 10.1038/s41598-022-10358-x.
[27] N. Bora, S. Gutta, and A. Hadaegh, "Using machine learning to predict heart disease," *Wseas Transactions on Biology and Biomedicine*, vol. 19, pp. 1–9, 2022, doi: 10.37394/23208.2022.19.1.
[28] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Advances in Engineering Software*, vol. 69, pp. 46–61, 2014, doi: 10.1016/j.advengsoft.2013.12.007.

[29] I. Sabiri, H. Bouyghf, and A. Raihani, "Optimal interdigitated electrode sensor design for biosensors using differential evolution algorithm," *E3S Web of Conferences*, vol. 351, 2022, doi: 10.1051/e3sconf/202235101031.
[30] I. Sabiri, H. Bouyghf, and A. Raihani, "Optimal sizing of RF integrated inductors for power transfer of implantable biosensors," *Proceedings*, vol. 60, no. 1, 2020, doi: 10.3390/iecb2020-07053.
[31] H. Cho, Y. Kim, E. Lee, D. Choi, Y. Lee, and W. Rhee, "Basic enhancement strategies when using Bayesian optimization for hyperparameter tuning of deep neural networks," *IEEE Access*, vol. 8, pp. 52588–52608, 2020, doi: 10.1109/ACCESS.2020.2981072.
[32] K. Shankar, Y. Zhang, Y. Liu, L. Wu, and C. H. Chen, "Hyperparameter tuning deep learning for diabetic retinopathy fundus image classification," *IEEE Access*, vol. 8, pp. 118164–118173, 2020, doi: 10.1109/ACCESS.2020.3005152.

BIOGRAPHIES OF AUTHORS

Zaid Nouna was born in Casablanca, Morocco, in 1999. He got his bachelor's degree in Electrical Engineering and Industrial Computing in 2020, and then an engineering degree in Electrical and Telecommunications in 2023, both at the University Hassan II Mohammedia-Morocco. He is currently a Ph.D. student at Mohammedia-Faculty of Science and Technique Hassan II University of Casablanca, in the Laboratory of Engineering Sciences and Biosciences (LSIB). His research and interests are centered on the development and optimization of electronic systems for biomedical engineering and data science techniques for medical applications. He can be contacted at email: zaid.nouna-etu@etu.univh2c.ma.

Hamid Bouyghf was born in Errachidia, Morocco, in 1982. He got his B.S. and M.S. degrees in Electrical Engineering and Telecom from the University of Science and Technology in Fez, Morocco, in 2007, and his Ph.D. in Electrical Engineering and Telecom from Hassan II University of Casablanca, Morocco, in 2019. From 2015 until 2019, he worked as a Research Assistant at the Princeton Plasma Physics Laboratory. Since 2019, he has worked as an Assistant Professor at the Electrical Engineering Department of Hassan II University's FST Mohammedia in Casablanca, Morocco, and has received his habilitation in Electrical Engineering and Artificial Intelligence in 2023. In the topic of IC optimization, he has written numerous articles. His research interests include biomedical electronics, analog IC design, electromagnetic fields, low power design, and BLE applications. He can be contacted at email: hamid.bouyghf@gmail.com.

Mohammed Nahid studied electronics and received the B.S diploma in Electronics at ENSET Institute, Mohammedia, Morocco in 1994, received the Aggregation diploma of electronics and electrical engineering, from ENSET institute of Rabat, Morocco in 2000. He received the DESA diploma in computer engineering, telecoms and multimedia in 2004. He conducted doctoral research at images, videos coding and quality assessment under psychovisual quality criteria since 2004 at Mohamed V university of Rabat, Morocco, where he received his Ph.D. in 2010. Dr. Nahid, which is research professor in the Faculty of Science and Technology of Mohammedia since 2011, conducts research on visual perception and its application to coding, understanding, and display of visual information and to human vision, image quality, and digital imaging. Research and projects are also focused on artificial intelligency, embedded systems, biomedical systems, smart grids, digital and analog IC design. He can be contacted at email: mohammed.nahid@gmail.com.

Issa Sabiri was born in Msemrir, Morocco, 1995. In 2023, he got a Ph.D. degree in Electrical Engineering and Artificial Intelligence from ENSET Mohammedia-Faculty of Science and Technique Hassan II University of Casablanca, Morocco. His license degree in Bio Medical Instrumentation and Maintenance, at the Institute of Health Sciences Settat-Morocco (ISSS) at the University Hassan I, Settat-Morocco. He also got a Master's degree in biomedical engineering from FST Settat in 2018. His work, studies, and interests are focused on the development, design, and optimization of electronic systems for biomedical engineering and health sciences. He can be contacted at email: issa.sabiri@etu.fstm.ac.ma.