Journal of Integrated Sciences
Special Issue, February 2026
ISSN: 2806-4801

APPLY OPTIMIZATION TECHNIQUES IN MACHINE LEARNING MODEL FOR BREAST CANCER DETECTION

[1]Ramen Kumar Das, [2]Shatabdi Chatterjee, [3]Sajnin Sampa

[1]Bangladesh University of Engineering and Technology (BUET)
[2,3]Daffodil Institute of IT

ABSTRACT

Breast cancer is a serious issue for humans and one of the most common disorders for women. In the era of technology, researchers try to address any complication with ML, AI, Optimization and DL. This study aims to address the breast cancer problem with optimization and ML algorithm. The goal of this work is to identify the best suite ML models and optimization techniques for breast cancer detection. This study employs six optimization techniques and three ML models for breast cancer analysis. Random search, Grid search, Bayesian, Ant Bee Colony, Blue Whale, particle swarm optimization with SVM, XG-Boost, and ANN architecture impose this study to investigate breast cancer. This study utilizes a quantitative correlational approach based on the Mammographic Mass dataset. In SVM and XG Boost, Grid search shows the highest accuracy of 86.46% and 85.93%. However, ANN outperforms the accuracy of 83.13% for the Random search approach. The findings of this work are that Random search and Grid search contain maximum accuracy, but concerning time, Bayesian optimization is superior.

Keywords: *Hyperparameter, Machine Learning, Tuning, Breast Cancer.*

Corresponding author: Ramen Kumar Das can be contacted at rkdas.ict@gmail.com

1. INTRODUCTION

Breast cancer is a serious concern for women of all ages. The study proved that women aged 50 and older are highly affected by this disease. Mainly, Cancer originates in breast tumors. Tumors lead to breast cancer, and it is a significant threat not only to older women but also to females of all ages. Technology is everywhere, so addressing an issue like breast cancer through ML and optimization paves the way to get a solution against this serious issue affecting half of human beings. As it is a medical science scenario that's why it will be significantly practical not only to apply a Machine Learning algorithm to find this but also to combine the optimization scheme to develop the solution path.

Machine Learning is an emerging technology for the recent and upcoming world. It has gone to the point where it seems a standard life tool in our lives. Optimization means, from a set of solutions, considering the environment and parameters, which solution is the best according to the aspect of an application. Many optimization methods were introduced recently in the previous year. In this study, we used six optimization techniques: random search, Grid search, Bayesian optimization, Ant Bee colony optimization, Particle swarm optimization, blue whale optimization, and three machine learning models, namely support vector machine, XG boost, and ANN.

Breast Cancer is a serious issue for women. It is a significant concern for the whole world. It mainly refers to unusual cell growth in the breast. 2.3 million people were infected with breast cancer between 2022 and 670000 died from this serious issue (*Breast Cancer*, 2025). Females are at a serious risk of this illness. Scientists have revealed some factors that contribute to diseases. Age, Obesity, family history, alcohol consumption, radiation exposure, and reproductive history are the common factors that increase the risk of breast cancer. As women are half of humankind, so it is a burning topic to address breast cancer in research.

2. REVIEW OF LITERATURE

A previous study review is a strategic approach to any investigation. The literature involved here is mainly related to optimization and Machine Learning methods.

In Vaiyapuri (2025), six ensemble techniques and Bayesian optimization were utilized for human resource analytics in a corporate scenario. This study employed the IBM HR dataset, where the categorical boosting structure showed superior performance accuracy, like 95.8 %, and the AUC was 0.98. This study worked on Random Forest (RF), gradient boosting (GB), adaptive boosting (ADA), light gradient boosting machine (LGBM), extreme GB (XGB), and categorical boosting (CB). The limitations of this study were limited with one optimization method and not any deep model, except those ensembles operated solely here.

In Jiang (2025), computational fluid dynamics, ML and GA were imposed for finding the optimal contraction angle in the inlet passage. Three ML regression architecture showed their performance, for instance, Gaussian process regression, Feedforward neural network and Support vector regression. GA involved here to get the objective function value for the GPR model, where the value was 0.963, and the corresponding three contraction angles were θ1 = 13.34◦, θ2 = 28.36◦ and θ3 = 3.64◦. On the test set, the R^2 scores for GPR, FNN, and SVR were 0.958, 0.948, and 0.948, respectively. The limitation of this study was the potential of overfitting as the R-squared value is high, a limited dataset size of only 125 CFD simulations, and a narrow area of geometric pattern.

In Choi (2025), nine-regression model, conventional ensemble technique and deep model were utilized for port congestion and reduction of waiting time. SHAP were employed here for feature selection. Additionally, Genetic-algorithm-based hyperparameter modifying is used to optimize the XGBoost Regressor (XGBR), enhancing the coefficient of determination (R2) from 0.2791 to 0.2949 and decreasing the mean squared error (RMSE) from 20.9531 to 19.6387 and mean absolute error (MAE) from 13.6821 to 12.6753.

In Al-Fakih (2025), Bayesian optimization and ten different ML algorithms were utilized to predict geothermal temperature. The ten ML model showed their performance before and after optimization. MLP was observed to hold the highest R^2 value as well as the lower-level value of MAE and MSE. MLP, SMOR, and RF presented significant upper-level performance. After applying optimization (BO), the accuracy and performance of the investigation improved. For MLP the R^2 Value was 0.999, and the MAE was 0.15.

Table 1. Literature review of different studies based on HPO and ML

Paper	Model Used	Dataset Used	Accuracy/ Performance	Limitations
André Pfob (2022)	Logistic regression with elastic net penalty, Extreme Gradient Boosting, Multivariate Adaptive Regression Spline algorithm, Support Vector Machine, neural network	"Mammographic Mass	AUROC value LR-0.89 XG-Boost-0.88 MARS-0.88 SVM-0.89 NN-0.89	limited technical depth, sophisticated ML were not utilized, Advanced HPO were not employed, Lack of discussion on medical data challenges
Bernd Bischl [2022]	grid or random search, evolution strategies, Bayesian optimization, Hyperband, and racing.	No dataset mentioned as review or summarization	No numerical value, This showed HPO had a significant contribution to the analysis	no quantitative value for the result, broad conceptual, less area guidance regarding deep learning HPO
Yasser A. Ali (2023)	ML Model- SVM Opt. methods The Ant Bee Colony Algorithm, the Genetic Algorithm, the Whale Optimization and the Particle Swarm Optimization	Pima Indians Diabetes Dataset and Heart Disease datasets	ACO- 98.1 and 97, GA-98.9 and 97.8, WOA-72.3 and 71, PSO 80 and 79	only depends Classical ML model, limited dataset, no statistical test, Complexity was not explained
Warut Pannakkong (2022)	Fine Tuning- response surface methodology (RSM) ML model- artificial neural network (ANN), support vector machine (SVM), and deep belief network (DBN).	Pima Indians Diabetes Dataset	ANN-MAE was 1.09 00. SVM-MAE was 1.0 53, DBN-MAE was 1.0528	Proposed RSM was not applicable in high dimensional search space, RSM did not extensively compare, dedicated for single objective opt., RSM was manual and experimentally costly for heavy weight
Alejandro Morales-Hernánde z (2023)	evolutionary algorithm, Bayesian opt., surrogate, hybrid opt., multi-fidelity	Review study	N/A Finding: Lack of multi objective HPO work, guide for search space, early stopping criteria	mainly focused on neural architecture, Multi-objective HPO was not abundant, generalization findings, insufficient detail in methodology

Source: The authors' own work.

3. RESEARCH METHODOLOGY

This study follows the quantitative correlational approach. The dataset used here is a secondary dataset collected from the UCI machine learning repository, the Mammographic Mass dataset [elter,2007]. Three machine learning models were used for data analysis: SVM, XG-Boost and ANN. Six optimization methods were used to optimize the previous model: random search optimization, Grid search, Bayesian optimization, Ant bee colony optimization, blue whale optimization and particle swarm optimization. Confusion matrix and ROC curve were utilized to evaluate the model's performance. For the selection of an essential feature out of five features, three types of schemes were imposed: one is the ANOVA test, the second one is model- based feature selection, and the third is iterative feature selection.

3.1 Data Set Description:

The dataset has six columns. Five are feature columns, and the remaining one is the severity column, which is our output column. There are 961 samples and six columns: BI-RADS, Age, Shape, Margin, Density, and Severity.

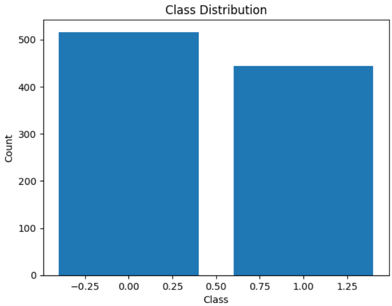

Figure 1. Class Differentiation Count in the dataset

It is a binary classification statement. 516 is in class 0, and 444 is in class 1. The class proportion is 53 percent (0 class) and 46 percent (1 class). The class distribution is closer to one another, which illustrates that this is not an imbalanced dataset. This dataset sampling area was not clearly disclosed, as it is a publicly available dataset; however, it is a possible area of US-based healthcare settings for female patients.

Optimization: Hyperparameters are the parameters that are used in the learning process. Mostly, learning rate, number of epochs, and n_estimators are called hyperparameters. The objective of hyperparameters is to obtain a value for the model parameter that gives the best or most suitable solution.

Data analysis Procedure is the Crucial section of a study. This study follows the general ML and Optimization working procedure. We explain the customized data analysis scheme imposed in this study. The step-by-step data analysis flow is as follows.

Data collection, data preprocessing, import the python libraries, divide the Independent and Dependent variable, split the dataset into training and test set, define the parameter space, Initialize the classifier and optimization scheme, start count CPU time, train the model, find best score and parameter value, get the accuracy of model prediction, develop ROC curve, Develop Confusion matrix, comparison of six optimization method.

4. RESULTS

Feature importance means which attribute affects the output column the most. Here 3 techniques utilized to do these jobs. Anova test, model-based feature selection and iterative feature selection.

Table 2. ANOVA test results for all features of the dataset

Feature name	sum of squares	degrees of freedom (df)	F-statistics	PR(>F) p-value
BI-RADS	22.235884	7.0	25.823733	3.384926e-32
Age	10.314111	73.0	1.148607	1.928657e-01
Shape	3.400564	4.0	6.911202	1.774335e-05
Margin	2.586660	5.0	4.205640	8.787262e-04
Density	0.396053	4.0	0.804926	5.221188e-01
Residual	106.525902	866.0		

Source: The authors' own work.

For Column BI-RADS, a high F-statistic and low p-value indicate this feature has high significance for the dependent (severity) column. Shape and Margin column F statistics values are 6.911 and 4.205, which are relatively higher than other columns except BI-RADS. Their P value is low, so they have moderate significance of the dependent variable but not as much as the first column. Considering the age and density column, they have closer F-statistics and p-value, but in the age column, the degree of freedom is very high, so age cannot be a significant parameter for determining the dependent variable.

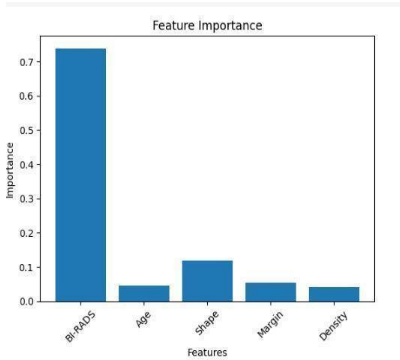

Figure 3. Model-Based feature selection for feature analysis

Figure 2 presents the features' names and relative importance. The BI-RADS feature has the highest relative importance score, 0.7; the shape feature has a score of 0.1 or more, and the margin feature has a score of 0.5 or slightly more. The age and density columns have the least effect on the severity output column.

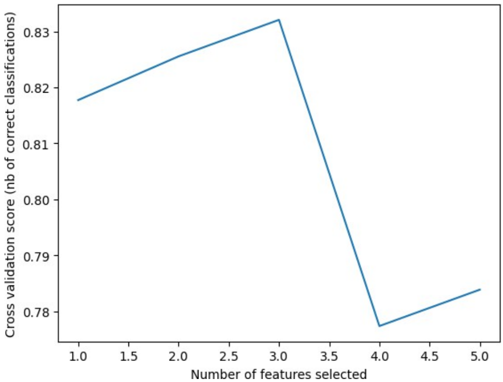

Figure 3. Iterative feature selection for feature analysis

This figure-3 illustrates the iterative feature selection method and how each feature affects the cross- validation score. IFS works by adding features one by one. The BI-RADS feature gives a good CV score and sharply raises the curve at the peak point to CV 0.83 or more. After adding a new feature, the CV decreases very sharply and gives the lowest CV value of 0.78. After that, when a new feature is added, it increases the curve again. In this analysis, three features affect the output the most: BI- RADS, Shape and Margin, which aligns with the model-based feature selection procedure.

Journal of Integrated Sciences
Special Issue, February 2026
ISSN: 2806-4801

Table 3. Best Score and Accuracy of all Optimizations with three ML models

ML model		Random Search	Grid Search	Bayesian Optimization	Ant Bee Colony	Particle Swarm	Blue Whale optimization
SVM	best score	82.42%	82.42%	0.83 %	83.54%	53.75%	83.54%
	Accuracy	86.45%	86.46%	0.81 %	84.37%	53.12%	85.41%
XG-Boost	best score	83.47%	84.24%	83.60%	83.59%	83.59%	83.20%
	Accuracy	84.89%	85.93%	84.89%	84.57%	84.90%	85.42%
ANN	best score	85.24%	85.39%	83.13%	-0.8614 or 86.14%	-0.8599 Or 85.99%	-0.8524 or 85.24%
	Accuracy	83.13%	81.93%	81.93%	79.72%	81.93%	80.72%

Source: The authors' own work.

Table 3 conveys the information about best score and accuracy of six optimizations techniques with three ML architecture. For SVM RS and GS presents the highest accuracy. Considering the XG Boost all six methods has shown near accuracy, like 84% or 85%. ANN model holds the superior accuracy for Random search scheme.

Table 4. Parameter and Optimized value of all Optimizations with SVM models

Optimization → Parameters ↓	Random search	Grid search	Bayesian optimization	Ant bee colony optimization	Blue Whale Optimization	Particle Swarm Optimization
Kernel	linear	linear	rbf	rbf fixed	Rbf fixed	rbf fixed
gamma	1000.0	0.1	0.01	0.001	0.4175	0.0190
C	0.1	0.1	8.14	1000	0.877	49.62

Source: The authors' own work.

Journal of Integrated Sciences
Special Issue, February 2026
ISSN: 2806-4801

Table 4 presents six optimization methods with their parameter values concerning the SVM classifier. In RS, GS and BO, the optimized kernel was linear, linear and RBF; however, in ABC, BWO, and PSO, the RBF kernel was fixed. RS showed the highest gamma value, and GS, BO, ABC, and PSO held smaller gamma values. Highest gamma indicates a sharp and complex decision boundary, high risk of encountering overfitting, low bias, and high variance. The smaller gamma indicates the opposite. Considering the C value, ABC holds the highest value, and RS and GS contain the same smallest value. If the C value is low, it means more regulation and a wider margin. And for a high C value, it reverses from the previous.

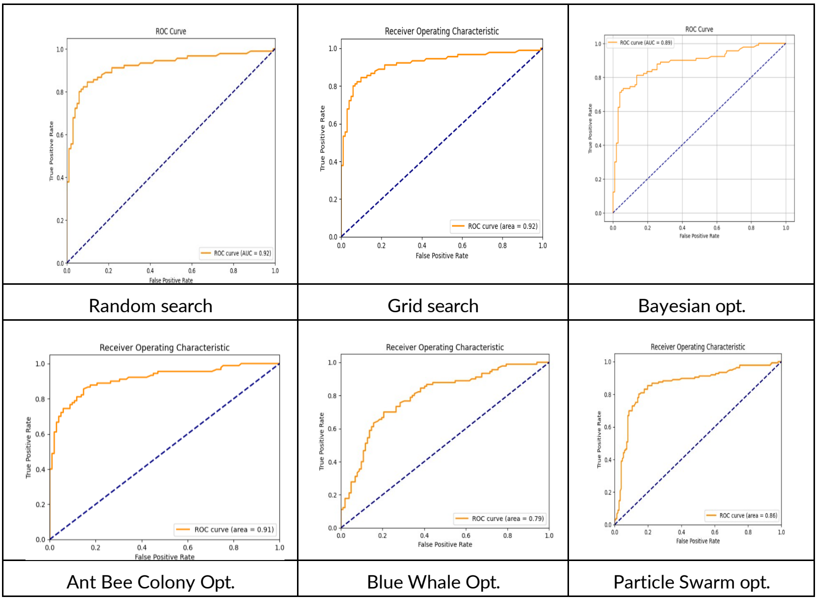

Figure 4. ROC Curve of six Optimizations for the SVM model

Figure 4 Conveys the information ROC curve for the SVM classifier with six different optimization schemes. X-axes indicate a false positive rate, and Y-axes present the actual positive rate. For Random and Grid search, the AUC value is 0.92, indicating a high-performance classifier. Additionally, this model has a 92 per cent chance of distinguishing positive and negative classes. Moreover, the orange mark in the slope's upper left corner is high, indicating high sensitivity and specificity. AUC value was lowest in the Blue whale opt. Yet, random and grid schemes held the highest AUC value. Blue whale opt. and particle swarm opt. Contained a TPR value lower than the threshold. This indicated that these two optimizations performed poorly when identifying the positive class. Regarding random search, Grid search, and ant bee colony strategies, the TPR value is nearly .4 or exactly 0.4. This conveyed that these methods could classify 40% of the positive class. BO has an 89% chance of distinguishing between positive and negative classes, and it also holds a high TPR.

ISSN: 2806-4801

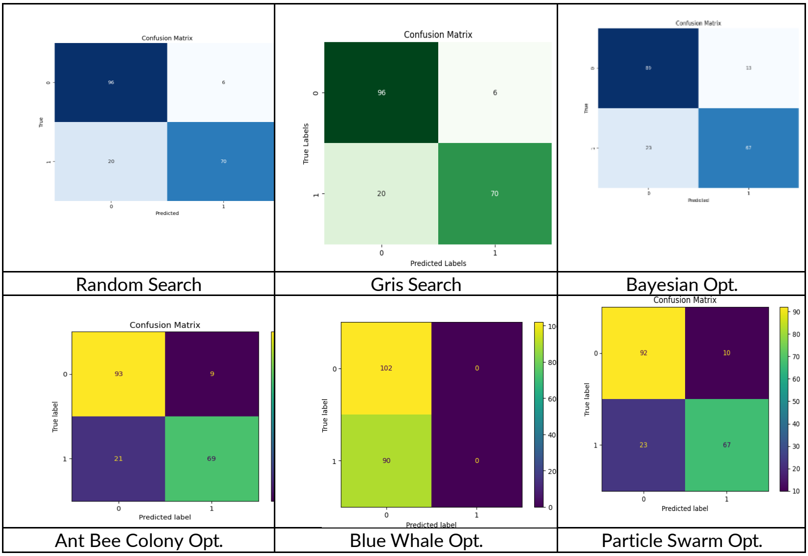

Figure 5. Confusion matrix for Six Optimizations with SVM

Figure 5 provides information regarding six confusion matrices for different optimization methods with a support vector machine. The dataset utilized in this study included 960 samples. we maintained 20 % for testing, so the total value after summation of 4 values from the confusion matrix was 192, that is 20 % of 960. For random and grid search, the TP+TN value was the same, and their accuracy from the confusion matrix was 86.45 %. In Ant Bee Colony and Particle Swarm Optimization, accuracy values were near to each other, like 84.37 % and 82.81%. 53.13% was the lowest accuracy obtained from Blue Whale Optimization, as the TP+TN value is only 102. In BO, the accuracy from the confusion matrix is 81.25%.

Table 5. Execution Time of Six Optimizations for SVM Model

Optimization Techniques	CPU time	System Time	Total	Wall Time
Random search	1327.2271 S	0.4260 S	1327.6532 S	1324.0800 S
Grid Search	1920.3228 S	0.0000 S	1920.3228 S	1920.3228 S
Bayesian	34.67 S	0.73 S	35.4 S	87.64 S
Particle Swarm	38.6 S	0.214 S	38.8 S	45.8 S
Ant Bee Colony	3min 29s	0.549 s	3min 29.549s	3 min 28s
Blue Whale	2min 54s	0.394 s	2min 54.394s	2 min 58s

Source: The authors' own work.

Table 5 presents the required time for six optimizations with SVM model. CPU, System, Wall and total time are illustrated here. BO needs the least CPU time, but RS and GS took the highest amount of CPU time.

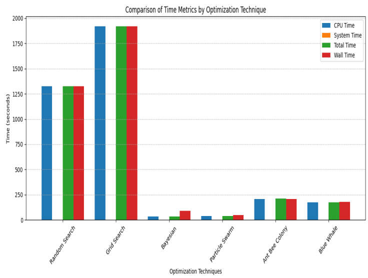

Figure 6. Time required for all optimizations with SVM

This Figure 6 demonstrates the execution time comparison for six optimization techniques involving- Random search, Grid search, Particle swarm, Ant Bee Colony, Blue whale with SVM (Support Vector Machine) Model. Four-time evaluation matrixes have been used naming- CPU time, System Time, total and wall time. Grid search displayed the highest amount of computational time with- CPU time, total and wall time; With random search showing comparable results, succeeding closely. In contrast to that, Metaheuristic algorithms- Particle swarm, Ant Bee Colony and blue whale illustrated significantly lower levels of execution time. That makes it the most efficient among the six optimization techniques.

Table 6. Parameters and Optimized value for six optimizations with XG Boost model

	Random Search	Grid search	Bayesian opt.	ABC	BWO	PSO
subsample	1	1	0.91	1	0.51	0.69
n_estimator s	100	100	968	200	51	65
min_child_ weight	2	2	10	3	1.021	4.89
max_depth	5	5	4	3	3	6
learning_rate	0.01	0.01	0.016	0.01	0.01	0.17
gamma	0	0.1	1	0.0	0.013	0.20
colsample_ bytree	0.9	0.8	0.81	1	0.51	0.50

Source: The authors' own work.

Table 6 illustrates a comparative analysis of seven hyper parameters optimized value for six optimization techniques with the XGBoost ML structure. Regarding Subsample, it is a hyperparameter used in tree-based structures here -RS, GS and ABC optimization techniques used the space ranges from [0.8,0.9,1] and showed the optimized value 1 which is the upper bound value that interprets it performs best in higher complexity also has the opportunity to encounter overfitting. However, Bayesian opt and BWO used the space ranges from [0.5, 1.0]. Bayesian opt displayed the optimized value 0.91 that is adjacent to the upper bound value with high data utilization.

Conversely, BWO showed the 0.51 as the optimized value adjacent to the lower bound value, reducing overfitting risk. PSO optimization technique shows the optimized value as 0.69. For n_estimators parameter - RS, GS and ABC examined a discrete set of values [100,200,300]. Both RS and GS showed 100 as the optimized value indicates that lower number of trees is sufficient to gain optimal performance and overfitting risk is lower. On the contrary, ABC identified 200 as the optimized value illustrates that a larger ensemble improves the model but further increasing the count (i.e. to 300) may occur diminishing returns. This represents a balance between the model complexity and generalization capability. For Bayesian optimization analyzed a wide continuous range between 100 to 1000. It represents 968 as the optimized value which is close to the upper bound. BWO depicts 51 as the optimized value where the rages were between 50 to 300 positioning it adjacent to the lower bound. This implies that the model performs best in smaller ensemble size. Moreover, PSO shows the optimized value as 65. In optimizing the 'min_child_weight' parameter with a range of [1, 2, 3], both RS and GS selected 2 as the optimal value. This implies a balanced approach to controlling model complexity and avoiding Excessive modeling. However, the ABC algorithm identified 3 as the best value, favoring slightly more conservative splits and potentially enhancing generalization. This difference highlights the capability of the ABC method to explore parameter configurations that emphasize robustness. It also demonstrates how bio- inspired optimization can lead to choices that prioritize stability over immediate performance gains. Bayesian optimization analyzed a wide continuous range between 1 to 10. It represents 10 as the optimized value, which is the upper bound value. For BWO optimization, the optimized value is 1.021, and PSO optimized value is 4.89. And max depth parameter optimized within the range [3, 4, 5], both RS and GS identified 5 as the optimal value. This indicates that deeper trees performed better under these traditional search methods. However, the ABC algorithm selected 3 as the optimal value, suggesting a preference for simpler models that may generalize better and reduce Excessive modeling. This contrast highlights the ABC algorithm's ability to balance model complexity and performance effectively. It also demonstrates the advantage of bio-inspired methods in exploring solutions that might be overlooked by conventional search techniques. Bayesian optimization analyzed a wide range of 3 to 10. It represented 4 as the optimized value. BWO showed 3 as the optimized value adjacent to the lower bound value reducing overfitting risk. PSO optimization technique shows the optimized value as 6. In the tuning process for the 'learning_rate' parameter with a range of [0.1, 0.01, 0.05], all three methods—RS, GS, ABC and BWO identified 0.01 as the optimal value. This consistency suggests that a lower learning rate led to better model performance across different optimization techniques. A learning rate of 0.01 allows the model to learn more gradually, reducing the risk of overshooting optimal solutions and improving overall stability. The agreement among all methods reinforces the reliability of this parameter choice. It also indicates that, in this case, traditional and bio-inspired methods converged on the same effective solution. For Bayesian optimization techniques optimized value occurs 0.016. PSO optimization technique shows the optimized value as 0.17. Regarding the 'gamma' parameter within the range of [0, 0.1, 0.2] for the optimization techniques RS, GS and ABC; RS selected 0 while GS chose 0.1 as the optimal value. This variation indicates differing preferences in how aggressively each method prunes the tree to control overfitting. Interestingly, the ABC algorithm also identified 0.0 as the best value, aligning with RS and suggesting that no regularization was more effective in this case. This outcome highlights how different optimization methods can converge or diverge depending on the parameter landscape. The agreement between RS and ABC shows a possible tendency toward simpler models without the need for additional penalty. Bayesian presented an optimized value as 1 within the range [0,1] and BWO showed the 0.013 as the optimized value within the range [0,5] adjacent to the lower bound. PSO optimization technique shows the optimized value as 0.20. For the 'colsample_bytree' parameter with a range of [0.8, 0.9, 1] for the optimization techniques- RS, GS and ABC; RS selected 0.9 while GS chose 0.8 as the optimal value. These values suggest that using a subset of features for each tree helped improve model performance in traditional search methods. However, the ABC algorithm identified 1 as the optimal value, indicating that using all available features led to better results in its exploration. This difference highlights the ABC ability to effectively evaluate the full feature set, potentially capturing more complex patterns. It also shows how bio-inspired methods can sometimes favor more comprehensive configurations that conventional techniques might overlook. For Bayesian optimization techniques range- [0.5, 1] optimized value occurs as 0.81 and BWO with the range of [0.5, 1.0] shows the 0.51 as the optimized value adjacent to the lower bound value reducing

Journal of Integrated Sciences
Special Issue, February 2026
ISSN: 2806-4801

overfitting risk. PSO optimization technique shows the optimized value as 0.50.

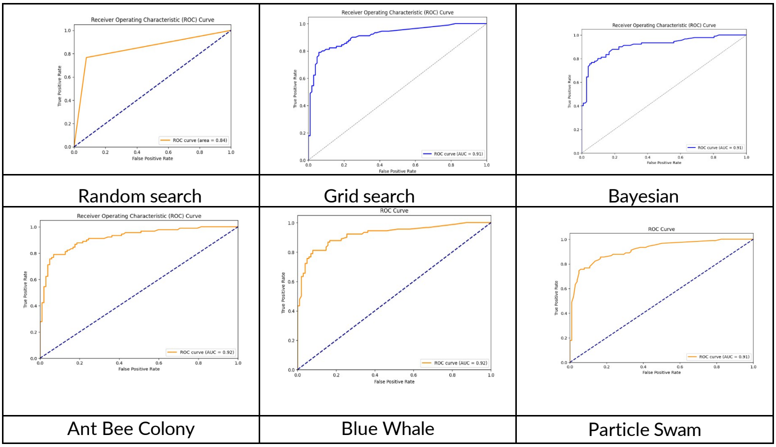

Figure 7. ROC Curve for Six Optimizations with XG Boost

The above figure illustrates six different optimization results with the XG Boost classification algorithm. The upper row three traditional showed AUC 0.84,0.91, and 0.91, but bio-inspired optimization was mainly better than the traditional method in here 1st three schemes. ABC, BWO, and PSO AUC values were higher, like 0.91 or 0.92. In RS true positive value started from less than 0.4, which implied the model threshold was strict, and it faced trouble early stages to identify true positives. However, metaheuristic optimization showed higher TPR in lower FPR illustrated that the model was more effective in the early detection phase.

Journal of Integrated Sciences
Special Issue, February 2026
ISSN: 2806-4801

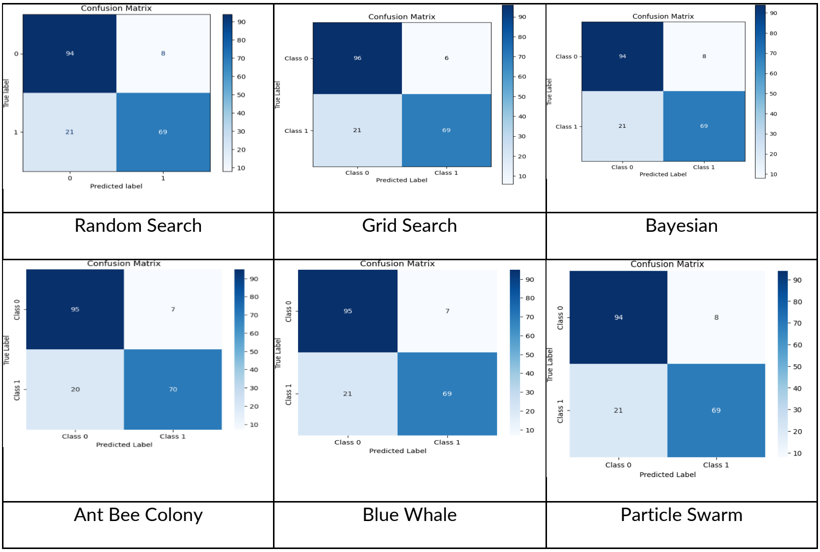

Figure 8. Confusion Matrix for six Optimizations with XG Boost

This figure conveys the scenario of six different optimization confusion matrices for the XG Boost model. The upper three matrices are for traditional optimization, and the lower rest are bio meta- heuristic. The TP+ TN values were 163,165, and 163 for random search, grid search and Bayesian method, respectively. Moreover, the accuracy was 84.89 %, 85.93%, and 84.89 % for the first three traditional methods. The other three were showed better accuracy, like 85.93%, 85.41%, and 84.89% for ABC, BWO, and PSO, respectively. In the context of XG boost, the typical and advance bio- inspired optimization techniques conveyed the accuracy value nearly.

Journal of Integrated Sciences
Special Issue, February 2026
ISSN: 2806-4801

Optimization Techniques	CPU time	System Time	Total	Wall Time
Random search	1.57 S	8.25 S	9.82 S	9.83S
Grid Search	25.84 S	0.96 S	26.8 S	213.86 S
Bayesian	377.50 S	3.43S	380.93 S	81.89 S
Ant Bee Colony	1210.53 S	1059.19 S	2269.72 S	151.33 S
Blue Whale	158.5969 S	11.8300 S	170.42 S	92.5722 S
Particle Swarm	88.58 S	7.14 S	95.72 S	54.89 S

Source: The authors' own work.

Table 7 presents CPU, System, Wall and total time for the XG Boost model with six optimization techniques. ABC took the highest CPU time. However, RS utilized only 1.57 seconds, the lowest CPU period among others.

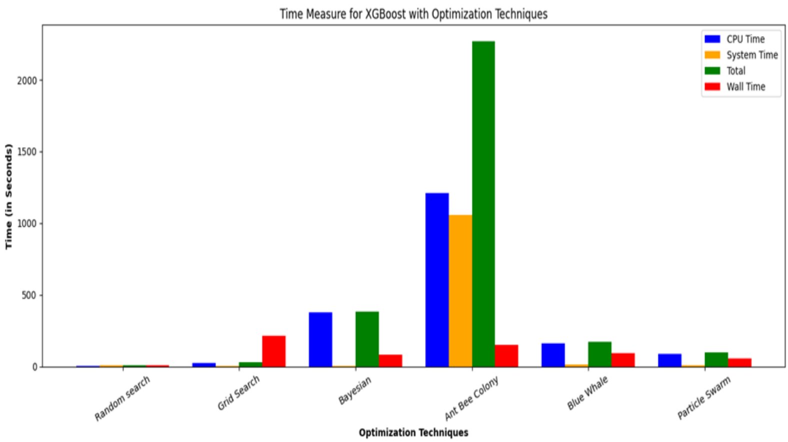

Figure 9. Time measure for all optimizations with XG-Boost

Journal of Integrated Sciences
Special Issue, February 2026
ISSN: 2806-4801

Figure 9 illustrates the comparative time performance of six optimization techniques applied to the XG-Boost algorithm. The optimization methods evaluated include Random Search, Grid Search, Bayesian Optimization, Ant Bee Colony, Blue Whale, and Particle Swarm Optimization. The time metrics considered are CPU Time, System Time, Total Time and Wall Time, measured in seconds. The Ant Bee Colony method exhibits the highest total computational time, with notably elevated CPU and system times, indicating its resource-intensive nature. Bayesian Optimization also shows a relatively high CPU and total time compared to other techniques, though significantly less than Ant Bee Colony. On the contrary, Lower time requirement making This analysis highlights the trade-off between optimization efficiency and computational cost, where advanced metaheuristic methods may provide superior optimization performance but at the expense of significantly higher time consumption. This comparison highlights the trade-off between optimization complexity and computational cost in hyperparameter tuning for XG-Boost.

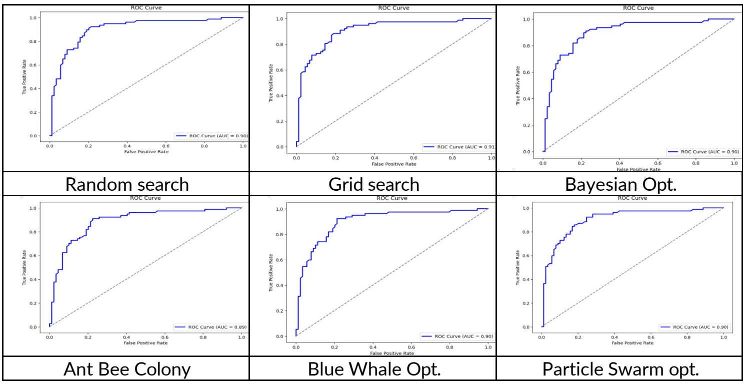

Figure 10. ROC Curve for six Optimizations with ANN

In this figure 10, six ROCs for different optimization methods are presented. All AUC values are about 0.90, which clarifies that an Artificial Neural Network is definitely a superior tool for classifying tasks rather than SVM and XG Boost when these models are applied with optimization.

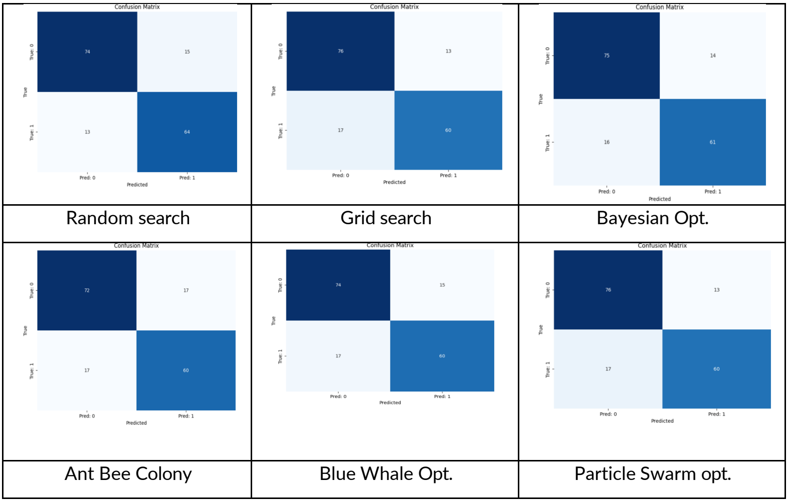

Figure 11. Confusion matrix for six optimizations with ANN

In Figure 11 confusion matrix is shown for six optimization methods with an ANN architecture. For the typical method, TP+TN values were 138, 136, and 136, which indicate 71.87% and 70.83% accuracy. For the bio-inspired meta-heuristic optimization method, the accuracy from the confusion matrix was 68.75%, 69.79%, and 70.83% for ABC, BHO and PSO, respectively. This indicates optimization with ANN yields near 70% accuracy from the confusion matrix calculation.

Table 8. Time measures for all optimizations with ANN

Optimization Techniques	CPU time	System Time	Total	Wall Time
Random search	39.96 S	7.13 S	47.09 S	38.14 S
Grid Search	265.46 S	40.99 S	306.45 S	238.19 S
Bayesian	7.30 S	17.27 S	24.3 S	30.79 S
Ant Bee Colony	28.93 S	4.13 S	33.06 S	27.23 S
Blue Whale	22.63 S	3.60 S	26.23 S	21.58 S
Particle Swarm	238.11 S	35.11 S	273.22 S	228.11 S

Source: The authors' own work.

Table 8 conveys the information of ANN analysis of the period that encounters six optimization methods. GS and PSO grabbed the highest CPU time; on the other hand, BO utilized the lowest CPU time for the ANN network.

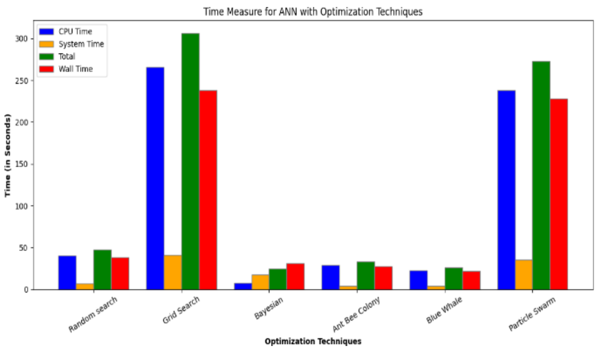

Figure 12. Time measure for six optimizations with ANN

In Figure 12 we have illustrated the Time Measure for ANN (Artificial Neural Network) with six different optimization techniques. X- Axis conveys the optimization techniques –

Random search, Grid Search, Bayesian, Ant Bee colony, Blue whale and particle swarm. And Y-Axis depicts Time (in seconds). The figure demonstrates that Grid search and Particle swarm optimization technique has the highest CPU and total times. That implies substantial computational demand. Conversely, Bayesian, Ant Bee Colony and blue whale methods showed significantly lower levels of time consumption which indicates their efficiency.

5. DISCUSSION

SVM Random and Grid search showed the highest accuracy, and PSO took the least CPU, Wall and system time. For XG Boost, all six schemes conveyed a near accuracy of about 84-85%. Moreover, ABC consumed the most time, but RS and GS needed the shortest period for execution. RS presented superior accuracy for the ANN structure, but ABC encountered the lowest accuracy. Additionally, GS and PSO required more time than the other four methods. SVM method showed ROC Curve value of about decimal 9 or more, but PSO and BWO presented a reduced value for the Curve. ROC Curve value in XG boost method held a decimal 9 or more; in contrast, RS AUC value was less than the other five optimizations. In neural network architecture, all six AUC values were near 0.90. From the confusion matrix of six optimizations for SVM, the TP+TN in RS and GS was more than 85 percent. In XG boost, the confusion matrix for every six techniques, the TP+TN rate lies between 84% and 85 %. Besides the ANN architecture, the six confusion matrices presented a TP+TN value of nearly 70%.

6. CONCLUSION

Random search and Grid search convey the highest accuracy for the SVM model. In XG-Boost, all six optimizations present near accuracy. In ANN, RS holds the best accuracy. Considering the period, BO and PSO take less time when investigating with the SVM model. In the context of XG-Boost, RS needs minimal time. For ANN BO and BWO encounters, the tiniest time. For accuracy, RS and GS can be good paths. Moreover, concerning time, BO will be a fruitful tool.

REFERENCES

Ali, Y., Awwad, E., Al-Razgan, M., & Maarouf, A. (2023). Hyperparameter Search for Machine Learning Algorithms for Optimizing the Computational Complexity. *Processes*, *11*(2), 349. https://doi.org/10.3390/pr11020349

Bischl, B., Binder, M., Lang, M., Pielok, T., Richter, J., Coors, S., Thomas, J., Ullmann, T., Becker, M., Boulesteix, A., Deng, D., & Lindauer, M. (2023). Hyperparameter optimization: Foundations, algorithms, best practices, and open challenges. *WIREs Data Mining and Knowledge Discovery*, *13*(2), e1484. https://doi.org/10.1002/widm.1484

ChatGPT. (n.d.). Retrieved June 1, 2025, from https://chatgpt.com

Choi, M.-H., & Yoon, W. (2025). Predicting Ship Waiting Times Using Machine Learning for Enhanced Port Operations. *IEEE Access*, *13*, 81377–81391. https://doi.org/10.1109/ACCESS.2025.3566429

Elter, M. (2007). *Mammographic Mass* [Dataset]. UCI Machine Learning Repository. https://doi.org/10.24432/C53K6Z

Flow chart of Bayesian optimization. | Download Scientific Diagram. (n.d.). Retrieved June 1, 2025, from https://www.researchgate.net/figure/Flow-chart-of-Bayesian-optimization_fig1_351834872

Gul, E. (2025). Urban flood hazard assessment using FLA-optimized boost algorithms in Ankara, Türkiye. *Applied*

Water Science, *15*(4), 78. https://doi.org/10.1007/s13201-025-02424-2

Introduction to Ant Colony Optimization. (09:31:33+00:00). GeeksforGeeks. https://www.geeksforgeeks.org/introduction-to-ant-colony-optimization/

Jiang, T., Lu, W., Lu, L., Xu, L., Xi, W., Liu, J., & Zhu, Y. (2025). Inlet Passage Hydraulic Performance Optimization of Coastal Drainage Pump System Based on Machine Learning Algorithms. *Journal of Marine Science and Engineering*, *13*(2), 274. https://doi.org/10.3390/jmse13020274

Morales-Hernández, A., Van Nieuwenhuyse, I., & Rojas Gonzalez, S. (2023). A survey on multi-objective hyperparameter optimization algorithms for machine learning. *Artificial Intelligence Review*, *56*(8), 8043–8093. https://doi.org/10.1007/s10462-022-10359-2

Pannakkong, W., Thiwa-Anont, K., Singthong, K., Parthanadee, P., & Buddhakulsomsiri, J. (2022). Hyperparameter Tuning of Machine Learning Algorithms Using Response Surface Methodology: A Case Study of ANN, SVM, and DBN. *Mathematical Problems in Engineering*, *2022*, 1–17. https://doi.org/10.1155/2022/8513719

Particle Swarm Optimization (PSO) – An Overview | GeeksforGeeks. (n.d.). Retrieved June 1, 2025, from https://www.geeksforgeeks.org/particle-swarm-optimization-pso-an-overview/

Pfob, A., Lu, S.-C., & Sidey-Gibbons, C. (2022). Machine learning in medicine: A practical introduction to techniques for data pre-processing, hyperparameter tuning, and model comparison. *BMC Medical Research Methodology*, *22*(1), 282. https://doi.org/10.1186/s12874-022-01758-8

Ravula, V., & Ramaiah, M. (2025). Enhancing phishing detection with dynamic optimization and character-level deep learning in cloud environments. *PeerJ Computer Science*, *11*, e2640. https://doi.org/10.7717/peerj-cs.2640

Vaiyapuri, T., & Sbai, Z. (2025). Bayesian Optimized Boosted Ensemble models for HR Analytics—Adopting Green Human Resource Management Practices. *International Journal of Technology*, *16*(2), 561. https://doi.org/10.14716/ijtech.v16i2.7277

Whale optimizatio algorithm. (2016, October 8). SlideShare. https://www.slideshare.net/slideshow/whale-optimizatio-algorithm/66899593

Whale Optimization Algorithm (WOA). (17:07:46+00:00). GeeksforGeeks. https://www.geeksforgeeks.org/whale-optimization-algorithm-woa/

Breast cancer. (n.d.). Retrieved October 10, 2025, from https://www.who.int/news-room/fact-sheets/detail/breast-cancer