International Journal of Science, Mathematics and Technology Learning
ISSN: 2327-7971 (Print) ISSN: 2327-915X (Online)
Volume 33 No. 1, 2025

CDRPM: Cardiac Disease Risk Prediction Model

Tanvi Rustagi[1], Meenu Vijarania[2]

[1]Research Scholar, Computer Science & Engineering, K.R Mangalam University, Gurugram, Haryana, India

[2]Associate Professor, Computer Science &Engineering,, K.R Mangalam University, Gurugram, Haryana, India

Abstract

The treatment of heart disease, a widespread health problem, requires a rapid and accurate diagnosis. This study improves the diagnosis of heart disease by combining XGBoost with a grey wolf search algorithm (GWSA). Hyperparameters such as regularization, tree depth and learning rate are used by GWSA to optimize the performance of XGBoost classifier. Data pre-processing ensured consistency in the scaling and processed missing data. When combined with XGBoost and GWSA, it improves the accuracy of the cardiac algorithm more than when using traditional parametric tuning methods. Numerous metrics demonstrate the ability of the improved XGBoost model to differentiate between different heart states. The result of the model proposed shows an accuracy of 97.8 percent, which is significantly higher than the traditional ML algorithm. The proposed model has a precision of 97, a recall of 89 and an F1 score of 93. Explanations on the interpretability of the model and the importance of the characteristics for the diagnostic decision are given in the paper. The proposed techniques in clinical practice will improve patient care and health outcomes.

Keywords

Grey Wolf Search Algorithm (GWSA), XGBoost, Hyperparameter Tuning, Cardiac Disorders, Heart Diseases, Diagnosis

1. Introduction

Identification of heart disease is important for several reasons. First, heart problems account for mortality globally [1-3]. Appropriate treatment can slow the course of a disease and

[1]Corresponding Author

Acceptance: 25 March 2025, Publication: 01 April 2025.
[2]Second Author

enhance patient outcomes if detected early. Many cardiac conditions have imprecise or non-existent symptoms, this making testing and diagnosing challenging in the absence of the appropriate resources. Physicians can begin therapy before serious diseases such as myocardial infarction, heart failure, or arrhythmias manifest by identifying patients symptoms early. This preventative method enhances patients quality of life while lowering the expense of treating major heart problems [21]. To identify heart problems, doctors use surgical techniques, high-tech imaging, and minimally invasive screening tests.

Early treatment is made possible by early cardiac diagnosis, which enhances patient outcomes [4]. Physicians can identify a wide range of heart disorders via non-invasive diagnostics, imaging, clinical evaluation, and invasive treatments. To increase the detection of cardiac disease and lessen its worldwide cost, issues with healthcare access, financial constraints, and interpreting abilities need to be addressed [6-8].

Table 1: Review of existing work

Author Name	Technology Used	Observation
Z. Jiang et al (2024)	Facial Expressions and Visual Patterns	AUROCs of 0.72–0.82, Less Accurate
N. Vahab et al (2024)	Convolution Neural Networks (CNN) & Random Forest Classifier(RFC)	AUC 0.94 , Outfitting issues
A. Nyström et al (2024)	Residual neural networks	0.76 AUC, Less Accurate
M.R Kumar et al (2023)	Navie Bayes	90.66% Accuracy, Slower on large datasets
K.M. Mohi Uddin et al (2023)	Gradient boost	92.16% Accuracy, Outfitting issues
Ashir Jawed et al (2022)	SVM,LR,RF,KNN, GA,ANN,NB,DT, DNN	ECG modality-based ML models used signal data for HF and CAD prediction and detection and outperformed other modalities in terms of performance and accuracy.

P. Tripathi et al (2022)	Linear discriminant analysis (LDA) classifier	96.0%Accuracy,Nothypertunned
Rachit Mishra et al(2021)	KNN,DT,LR,NAÏVE BAYES, Random Forest	Random forest has higher accuracy than other ML algorithms
Aadar Pandita et al(2021)	SVM,KNN,LR,RF,NB	KNN and Neural Networks are quite accurate at predicting cardiac problems.
Naresh Kumar et al (2021)	LR (Logistic Regression)	The proposed model can help with the early screening of diseases
I.H Sarkar et al (2021)	K-NN, SVM, Decision Tree, Naïve Bayes, Random Forest	Determines which machine learning model is most appropriate in a given situation.
B.U. Rindhe et al (2021)	ANN, SVM, Random Forest	For the prediction of cardiovascular heart illnesses, the support vector machine has the highest accuracy (84 percent).

2. Motivation

The present diagnostic in cardiology must be understood in order to identify Investigation vulnerabilities in the diagnosis of heart disease [5, 9, 16, and 18]. Research may be lacking in the following areas:

1. Applying AI and machine learning techniques can improve the efficiency and accuracy of CT, MRI, and echocardiogram scans.
2. Researching the potential of telemedicine and remote monitoring technologies to diagnose and treat cardiac problems in underserved or rural locations.
3. Motivate cardiologists, geneticists, and radiologists to work together to improve the diagnosis of complex heart disease by utilising their unique backgrounds and expertise.
4. Research initiatives that put patients' viewpoints and results first to make sure that their values, preferences, and quality of life are considered during the diagnostic process.
5. Researchers and medical professionals may close these Investigation vulnerabilities in cardiac diagnostics to enhance patient care by more correctly and timely diagnosis of cardiac problems.

2.1 Contributions to Research

Research may progress in a number of ways by integrating the different techniques, like gradient boosting algorithm (XGBoost) with the Gray Wolf Search algorithm (GWSA) to precisely diagnose cardiac conditions.

1. The search capability of the GWSA combined with the intricate data pattern capturing of XGBoost, enhance the identification of cardiac problems in this work.
2. XGBoost learns from data to provide precise predictions, while GWSA quickly explores the search space.XGBoost's feature significance has been used by the hybrid approach to choose features efficiently.
3. Proposed model utilized XGBoost interpretability due to feature importance, it may be improved by hybridising it with GWSA. GWSA's optimisation approach highlights important components to aid physicians in understanding the aetiology of heart illnesses.
4. Proposed model performance is analyzed and compared with other boosting algorithms.

GWSA's boost the effectiveness and scalability of the proposed algorithm by population-based search for enormous data sets. The hybrid approach swiftly handles large-scale and high-dimensional datasets because it explores the search space quickly. This makes it perfect for practical real world applications.The following is the general framework of the study.The second section examines and contrasts studies on heart disease prediction. In the third section, materials and techniques are presented. Details on current architecture and procedures, characteristics of a suggested hybrid model, and dataset descriptions are included. The current work on heart disease prediction is concluded in part four, which also outlines its shortcomings and potential future directions.

3. The Proposed Methodology

In the realm of machine learning, where prediction accuracy and efficiency are crucial, the XGBoost algorithm is a strong competitor. XGBoost continuously adds new models to the ensemble in an iterative manner, mitigating a pre-established loss function and attempting to correct mistakes produced by previous models. XGBoost is distinct in that it prioritises model optimisation for both accuracy and computational performance. In order to maintain the ability for prediction, XGBoost employs unique techniques like cache-aware access patterns, parallel and distributed computation, and tree pruning [10,11,13,14]. A loss term quantifies

prediction errors and a regularization term regulates model complexity. XGBoost maximises this unique objective function by controlling the model's complexity and measuring prediction errors using a regularisation term. Because of its versatility, XGBoost may be tailored to meet different learning objectives and evaluation standards. To avoid overfitting, XGBoost uses a dual-pronounce (Lasso) and (Ridge) regularisation. To accurately direct the training process, XGBoost makes use of the gradients and Hessians, which are the beginning and end derivatives of the loss function in relation to the expected scores. With the help of XGBoost's insights on feature relevance, users can choose and understand features more effectively and determine which factors in their datasets have the most influence.

The model can handle massive datasets since XGBoost uses distributed and parallel computation frameworks to speed up the training process on multicore CPUs and distributed clusters [9]. Data scientists and machine learning practitioners use it because of its widespread use in competitions, business, and academia. At the vanguard of the dynamic field of machine learning, XGBoost fosters innovation and offers answers to difficult problems in the actual world [12,22]. XGBoost creates results in an avaricious manner by continually dividing the data based on features that lower the loss function. Choosing the splits is done with the intention of maximising a gain metric that is calculated using the gradients and Hessians.

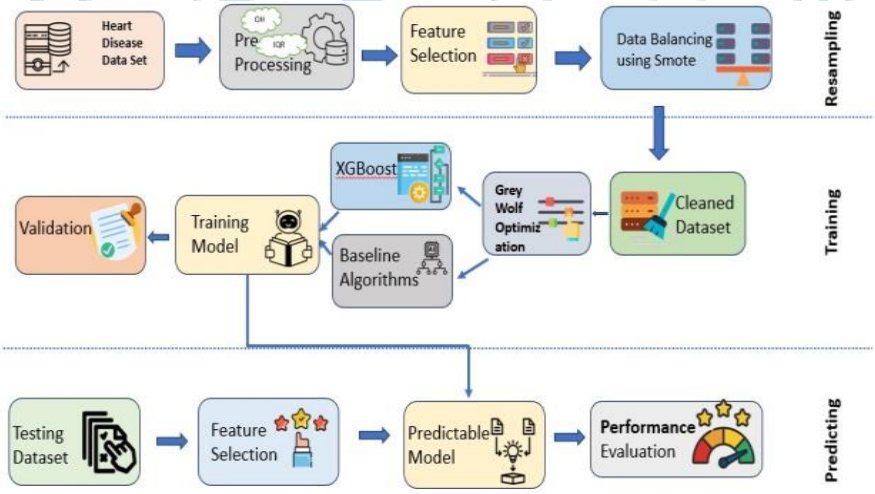

Figure 1: CDRPM: Cardiac Disease Risk Prediction Model

3. Dataset Description

The Data collection, which aims to analyse characteristics associated with heart disease or stroke, includes information about 7190 patients (Table 1).It has 16 columns with information on lifestyle, health, and demographics.

- Classify the subjects based on age and gender.
  • Their educational attainment is reflected in their education.
  • Lifestyle markers include body mass index (BMI), cigarettes per day (cigsPerDay), and smoking status (currentSmoker).
- TheBlood pressure (sysBP, diaBP), cholesterol (totChol), heart rate (heartRate), and glucose levels (glucose) are examples of health measurements.
- Prevalence of stroke (prevalentStroke), hypertension (prevalentHyp), diabetes, and medication usage (BPMeds) are among the medical history information.
  • Heart_stroke target variable, showing whether somebody had a heart attack or a stroke.

3.1 Data preprocessing

Heart disease dataset is split into 20% and 80% test and train sets respectively. We utilise the train dataset to train and optimise the model. Six attributes have missing values 105 values of education, 29 values for cigarettes per day, 53 values for blood pressure medication, 50 values for total cholesterol, 19 values for BMI, 1 value for heart rate and 388 values for glucose. The categorical feature encoding approach (OH encoding), will handle missing values. The target variable "Heart_stroke," has two classes presence and absence of heart disease for the dataset. This becomes a binary classification problem with target variable class labels of 0 and 1, where 1 indicates heart disease and 0 indicates no heart disease.

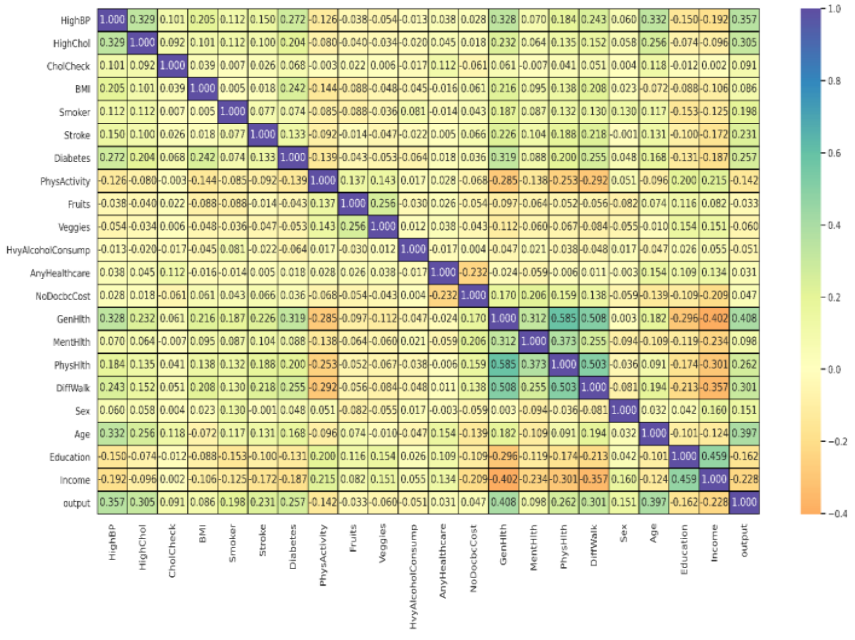

Figure 2: Heat map for Heart disease Dataset

3.2 Algorithm of proposed methodology:

The proposed approach optimizes the number of hyperparameters, such as learning rate, regularization, and subsampling of the XGBoost algorithm, using the GWSA algorithm. The swarm-based optimization algorithm GWSA utilized wolf hunting pattern to optimize the search space. GWSA requires less storage and processing and has a simple structure. Because GWSA requires less storage and processing and has a straightforward structure, Thus, hybridization of XGBoost and GWSA algorithm is employed for enhancing the performance while predicting the Heart disease. Figure 3 shows the process flow proposed methodology

Algorithm 1: XGBoost

Step 1. Construct a Base Model

$$\text{Residual} = \text{Target} - \text{Probability} \quad (1)$$

Step 2.Calculate the Similarity Weight

$$SW = \frac{\sum(\text{Residual})^2}{[\text{Pr}(1-\text{Pr})]+\lambda} \quad (2)$$

Step 3.Calculate Gain

$$SW = \frac{\sum(\text{Residual})^2}{[\text{Pr}(1-\text{Pr})]+\lambda} \quad (3)$$

Step 4.Find Base Model Output

$$\log(\text{odds}) = \log\left(\frac{\text{Pr}}{1-\text{Pr}}\right) \quad (4)$$

Step 5.Find New Probability

$$\text{New Pr} = \sigma(\log(\text{odds}) + \alpha(SWn) \quad (5)$$

Step 6: Repeat Step 2 to 5 till we find Optimal Value

The algorithm optimises a predetermined objective function, usually a mix of regularisation terms and a loss function, throughout the training phase. The hyperparameters of the XGBoost algorithm impact a number of factors, including model complexity, regularisation, learning rate, and tree formation [17,19].

Tree Booster parameters:

The gradient boosting step size each iteration is determined by the learning rate. Slower-learning models are more robust, but they require more iteration [29-31].

1. The ensemble maximum decision tree depth is denoted by Maximum_depth. Greater values might result in over fitting, more complicated models.
2. To control partitioning, minimum_child_weight determines the smallest total of instance weights (hessian) required to build a tree node.
3. Parameter for standardisation Gamma regulates the minimal loss reduction required for a leaf tree node to split.
4. The influence of magnitude on leaf weights is described by the LE2 regularisation term.
5. Alpha promotes feature sparsity by controlling the intensity of LE1 regularisation. Data Sampling Alternatives:
6. At each boosting step, the subsample selects a random fraction of the training data. Lower numbers might help avoid overfitting by introducing unpredictability.
7. The percentage of features (columns) randomly sampled at each level and during tree building is determined by the parameters colsample_bylevel and colsample_bytree.
8. Scale_pos_weight adjusts the positive class weight in binary classification tasks with skewed class distributions to rectify class imbalance.

To identify the ideal configuration, Grid explores and Random Search can explore a grid of possible hyperparameter combinations or randomly sample the hyperparameter space. Bayesian Optimisation use probabilistic models to iteratively search hyperparameter space and identify locations that enhance performance. Cross-validation evaluates the model across a wide range of hyperparameter values, preventing overfitting and enhancing performance estimations [15, 23, 25]. XGBoost hyperparameters have a significant impact on model performance and behaviour. It is crucial to comprehend each hyperparameter and tuning method in order to build reliable and accurate prediction models. As machine learning practitioners work with larger datasets and tasks, they must become proficient in hyperparameter tweaking to fully utilise XGBoost and other complex algorithms. It increases its area and hunts in groups, much like wolves do. Authors define the hyperparameter space of XGBoost. The learning rate, maximum tree depth, number of trees (boosting rounds), column subsampling ratio, subsample ratio, regularisation parameters (gamma and alpha), and subsample ratio are usually contained in this area. Each of the wolves that the authors trained represents a hyperparameter solution for the classification job. The authors use the objective function to compute the validation set performance of an

XGBoost model after it has been trained with the appropriate hyperparameters, resulting in the fitness values of the alpha, beta, and delta wolves. These wolves, in our opinion, are the finest choice for the population. The writers then modify the location of each wolf while remaining inside the hyperparameter range [20, 24, 28]. Set a break condition after that, such as a utmost number of iterations, computing time, or a workable solution. Then, modify the exploration and exploitation parameters of the GWSA to strike a balance between exploring new hyperparameter space and honing in on potential regions. Use the fittest wolf to adjust the hyperparameters if termination is successful. Use the Wolf Search Algorithm to hyperparameter tune XGBoost models and identifies cardiac problems. This tactic enhances the model's ability to detect cardiac irregularities in the actual world [24,27,32].

Algorithm 2: GWSA

Step1: Randomly initialize population of wolves Xi (i=1,2,…,n)

Step2: Initialize value of a=2, A and C

$\vec{A} = 2\vec{a}\vec{r}1-\vec{a}$ and $\vec{C} = 2\vec{r}2$ (1)

Step3: Calculate fitness of each member of population

$$\vec{D} = \left|C \,.\vec{X}\mathrm{p}\,(t) - \vec{X}\,(t)\right| \quad (2)$$

$$\vec{X}(t+1) = \vec{X}_{\mathrm{p}}\,(t) - \vec{A}\,.\,\vec{D}$$

Xα=member with the best fitness value

Xβ=second best member (in terms of fitness value)

Xδ=third best member (in terms of fitness value)

Step4: Variable t = 1 to Max number of iterations:

Update position of all delta wolves by WSA update rules

$$D\alpha = \left|\vec{C}_1.\vec{X}(t) - \vec{X}(t)\right| \quad (3)$$

$$D\beta = \left|\vec{C}_2.\vec{X}(t) - \vec{X}(t)\right|$$

$$D\delta = \left|\,\vec{C}_3.\vec{X}(t) - \vec{X}(t)\right|$$

$$\vec{X}_1(t+1) = \vec{X}(t) - \vec{A}_1.\vec{D}\alpha \quad (4)$$

$$\vec{X}_2(t+1) = \mathrm{X}(t) - \vec{A}_2.\vec{D}\beta$$

$$\vec{X}_3(t+1) = \mathrm{X}(t) - \vec{A}_3.\vec{D}\delta$$

$$\vec{X}(t+1) = \left(\vec{X}_1 + \vec{X}_2 + \vec{X}_3\right) / 3 \quad (5)$$

Revise value of a, A, C

$a = 2(1-t/T)$ (6)

Compute Fitness of all search agents

Revise Xα, Xβ, Xδ.

STOP

Step5: Return Xα

4. Performance Evaluation

The Grey Wolf Search Algorithm (GWSA) for XGBoost's hyperparameter tuning is a new and useful approach to addressing cardiac problems. GWSA optimisation makes advantage of wolf hunting behaviour and use of wolf pack intelligence to efficiently explore the XGBoost hyperparameter space. Analysing complicated, multi-dimensional data from genetic data, patient records, medical imaging, electrocardiograms (ECGs), and medical imaging is frequently necessary for the diagnosis of cardiovascular problems. XGBoost has been known for effectively managing non-linearity, feature interactions, and incomplete data. Heart disease detection may be improved by GWSA optimisation of XGBoost's hyperparameters. Computer science and health informatics collaborate to provide the Wolf Search Algorithm and XGBoost for the early detection of heart ailments. This multidisciplinary collaboration may present ground-breaking concepts that enhance medical diagnosis and care.

4.1 Result and Analysis

The confusion table for a binary classification model is represented in Table 2. It displays the proportion of accurate and inaccurate predictions the model produced. The algorithm correctly detected the positive cases, as evidenced by the 2812 true positives (predicted positive and really positive). The model missed some positive cases, as seen by the 109 false negatives (predicted negative but actually positive). Furthermore, the model misidentified negative cases as positive, resulting in 122 false positives (predicted positive but really negative). Lastly, 2709 real negatives both anticipated and actual—show that negative situations were accurately identified.

Table 2: Confusion Table of Proposed Model (Training)

	Predicted Positive	Predicted Negative
Actual Positive	2812	109
Actual Negative	122	2709

The confusion table for a binary classification model during testing (Table 3) represented. By contrasting expected and actual findings, a confusion matrix aids in evaluating model's performance. We can assess primary performance indicators including accu, prec, recall, and F1 score using confusion matrix. The fraction of accurate predictions (true positives and true negatives) to total number of occurrences, for example, indicates the model's accuracy. With very few incorrect classifications, this matrix shows that the model is operating effectively.

Table 3: Confusion Table of Proposed Model (Testing)

	Predicted Positive	Predicted Negative
Actual Positive	714	20
Actual Negative	18	686

Performance metrics for the suggested model, assessed on both training and testing datasets, are shown in Table 4. Accu, prec, recall, and F1 score are primary measures. Suggested model's comprehensive assessment metrics on the training and testing datasets are shown in the table. These metrics are essential for evaluating model's performance in terms of fitting to training data and generalising to unknown data (testing).

Table 4: Proposed Model Performance (Testing & Training Set)

Data	Accuracy	Precision	Recall	F1 Score
Train	96.0	95.83	86.43	90.09
Test	97.8	97.0	89.0	93.0

The model's capability to generalise is demonstrated by the similar performance metrics between the testing and training sets. The test set's slight improvements in accuracy, precision, recall, and F1 score imply that the proposed model hasn't overfitted to the training set. Most metrics on the test set show a little improvement, especially accuracy and recall, suggesting that the model is not just accurate but also effective at recognising positives without producing a large number of erroneous predictions. The strong F1 and accuracy

scores in both sets indicate that the model is resilient, minimising mistakes while reliably and successfully identifying the positive class.

4.2 Comparison with the Recent work

The proposed model's performance along with various other machine learning methods, including AdaBoost, Gradient Boost, XGBoost, CatBoost, and XGBoost with Bayesian optimization is shown in Table 5 and Figure 4 shows the graphical representation of result comparison.

Table 5: Result Comparison of Proposed Model

Algorithms	Accuracy	Precision	Recall	F1-Score
Adaboost	0.770	0.759	0.787	0.779
Gradient Boost	0.773	0.749	0.816	0.781
Xgboost	0.769	0.748	0.807	0.776
Catboost	0.773	0.750	0.814	0.781
Xgboost With Bayesian Optimization	0.968	0.969	0.857	0.905
Proposed Model	0.973	0.970	0.890	0.930

The moderate dataset and the expanded dataset are utilized to validate the suggested model. Table 6 and Figure 5 shows that extended dataset shows better performance across all metrics, indicating that large dataset improves the model's accuracy, precision, recall, and F1-score.

Table 6: Proposed Model Result Validation on Different Datasets

Algorithms	Accuracy	Precision	Recall	F1-Score
Proposed Model With Moderate Dataset	0.973	0.960	0.885	0.921
Proposed Model With Extended Dataset	0.978	0.970	0.890	0.930

5 Discussion and Conclusion

The XGBoost with Bayesian optimization performs admirably with an accuracy of 96.8%; however, it falls short of the suggested model in terms of accu, prec, recall, and F1-Score.

This demonstrates that even if Bayesian optimization enhances XGBoost's performance, the proposed model still surpasses it in terms of total classification performance.In summary, the suggested model outperforms all other evaluated algorithms in terms of predictive capacity, completing the classification work with higher accu, prec, recall, and F1 score. This demonstrates its resilience and appropriateness for use in Heart disease prediction where the accuracy and optimal performance is essential.

The suggested model performs well on both the moderate and expanded datasets, showing notable gains in F1-Score, recall, precision, and prediction accuracy when trained on the expanded dataset. The suggested model is a dependable option for real-world deployment, especially in tasks where high accuracy and a balance between precision and recall are crucial. Personalised medicine initiatives can customise therapies to each patient's risk factors by accurately identifying cardiac disease. Enhancing XGBoost with GWSA may lead to models that more accurately detect cardiac problems and identify risk factors and potential treatments. Enhancing the diagnosis of heart disease can have positive therapeutic effects on prognosis, early identification, and treatment planning. Researchers and physicians may use GWSA and XGBoost to build more accurate prediction models that will enhance patient outcomes, save healthcare expenditures, and ease system strain.

6 Future Scope

GWSA an metaheuristic optimisation techniques is inspired by wolf hunting. In classification tasks, XGBoost hyperparameter tweaking may reliably identify heart abnormalities. GWSA can quickly determine the best cardiac disease classification settings by exploring the hyperparameter space of XGBoost. To maximise XGBoost performance, WSA may change the tree count, regularisation parameters, learning rate, and tree depth. The importance of characteristics in predicting the target variable is scored by XGBoost. When diagnosing cardiac sickness, feature evaluation might yield the most pertinent clinical traits or biomarkers. Model predictions may be analysed and feature selection aided by this study. It is necessary to assess the accuracy of the XGBoost classifier following hyperparameter adjustment and model training. By removing false positives and negatives, these tests evaluate how well the model classifies cardiac conditions. It is necessary for the trained model to withstand cross-validation. On unidentified data, these elements are able to predict model generalisation. This is essential to prevent overfitting and evaluate how well the model performs on various data subsets. Plots of feature importance and decision trees aid in the

interpretation of XGBoost models. Researchers and physicians working in cardiology might gain insight by knowing how the model makes decisions. Interpretability is particularly crucial for healthcare applications where domain specialists are required to comprehend the decision-making processes of machine learning models. Cardiac diseases may be accurately recognised by using XGBoost for classification and GWSA for hyperparameter tweaking. The improved XGBoost model does a good job of using clinical data and biomarkers to differentiate heart diseases. This approach may aid medical professionals in early detection, risk assessment, and medical planning of heart illness. A strong foundation for developing interpretable heart disease detection models is offered by GWSA and XGBoost, which can enhance clinical decision-making in cardiology and patient outcomes.

References:

1. Chicco, D., & Jurman, G. (2020). Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone. BMC medical informatics and decision making, 20, 1-16.
2. Alsaeedi, A. H., Al-Shammary, D., Hadi, S. M., Ahmed, K., Ibaida, A., & AlKhazraji, N. (2024). A proactive grey wolf optimization for improving bioinformatic systems with high dimensional data. International Journal of Information Technology, 16(8), 4797-4814.
3. Xu, Y., Han, D., Huang, T., Zhang, X., Lu, H., Shen, S., & Wang, H. (2022). Predicting ICU mortality in rheumatic heart disease: comparison of XGBoost and logistic regression. Frontiers in Cardiovascular Medicine, 9, 847206.
4. Ahmed, S., Biswas, A., & Khairuzzaman, A. K. M. (2024). An experimentation of objective functions used for multilevel thresholding based image segmentation using particle swarm optimization. International Journal of Information Technology, 16(3), 1717-1732.
5. Dubey, A. K., Sinhal, A. K., & Sharma, R. (2024). Heart disease classification through crow intelligence optimization-based deep learning approach. International Journal of Information Technology, 16(3), 1815-1830.
6. Guo, C., Zhang, J., Liu, Y., Xie, Y., Han, Z., & Yu, J. (2020). Recursion Enhanced Random Forest with an Improved Linear Model (RERF-ILM) for Heart Disease Detection on the Internet of Medical Things Platform. IEEE Access, 8, 59247–59256. https://doi.org/10.1109/ACCESS.2020.2981159

7. Deviaene, M., Borzee, P., Van Gilst, M., Van Dijk, J., Overeem, S., Buyse, B., Testelmans, D., Van Huffel, S., & Varon, C. (2020). Multilevel Interval Coded Scoring to Assess the Cardiovascular Status of Sleep Apnea Patients Using Oxygen Saturation Markers. IEEE Transactions on Biomedical Engineering, 67(10), 2839–2848. https://doi.org/10.1109/TBME.2020.2972126
8. Kalita, K., Ganesh, N., Jayalakshmi, S., Chohan, J. S., Mallik, S., & Qin, H. (2023). Multi-Objective artificial bee colony optimized hybrid deep belief network and XGBoost algorithm for heart disease prediction. Frontiers in Digital Health, 5, 1279644.
9. Behera, M. P., Sarangi, A., Mishra, D., & Sarangi, S. K. (2022). A Hybrid Machine Learning algorithm for Heart and Liver Disease Prediction Using Modified Particle Swarm Optimization with Support Vector Machine. Procedia Computer Science, 218(2022), 818–827. https://doi.org/10.1016/j.procs.2023.01.062
10. Tripathi, P., Ansari, M. A., Gandhi, T. K., Mehrotra, R., Heyat, M. B. Bin, Akhtar, F., Ukwuoma, C. C., Muaad, A. Y. M., Kadah, Y. M., Al-Antari, M. A., & Li, J. P. (2022). Ensemble Computational Intelligent for Insomnia Sleep Stage Detection via the Sleep ECG Signal. IEEE Access, 10(August), 108710–108721. https://doi.org/10.1109/ACCESS.2022.3212120
11. Blanchard, M., Feuilloy, M., Sabil, A., Gerves-Pinquie, C., Gagnadoux, F., & Girault, J. M. (2022). A Deep Survival Learning Approach for Cardiovascular Risk Estimation in Patients With Sleep Apnea. IEEE Access, 10(December), 133468–133478. https://doi.org/10.1109/ACCESS.2022.3231743
12. Zhang, X., Lu, D., Hu, J., Banaei, A., & Abedi-Firouzjah, R. (2022). The role of ultrasound and MRI in diagnosing of obstetrics cardiac disorders: A systematic review. Journal of Radiation Research and Applied Sciences, 15(3), 261–269. https://doi.org/10.1016/j.jrras.2022.06.007
13. Dai, Q., Sherif, A. A., Jin, C., Chen, Y., Cai, P., & Li, P. (2022). Machine learning predicting mortality in sarcoidosis patients admitted for acute heart failure. Cardiovascular Digital Health Journal, 3(6), 297–304. https://doi.org/10.1016/j.cvdhj.2022.08.001
14. Pashikanti, R. S., Patil, C. Y., & Shinde, A. A. (2022). Cardiac Arrhythmia Classification using Deep Convolutional Neural Network and Fuzzy Inference

System. Proceedings - 2022 International Conference on Artificial Intelligence of Things, ICAIoT 2022, 1–6. https://doi.org/10.1109/ICAIoT57170.2022.10121894

15. Bhan, A., Mangipudi, P., & Goyal, A. (2023). An assessment of machine learning algorithms in diagnosing cardiovascular disease from right ventricle segmentation of cardiac magnetic resonance images. Healthcare Analytics, 3(December 2022), 100162. https://doi.org/10.1016/j.health.2023.100162
16. Rapp, E. S., Pawar, S. R., & Longoria, R. G. (2022). Hybrid Mock Circulatory Loop Simulation of Extreme Cardiac Events. IEEE Transactions on Biomedical Engineering, 69(9), 2883–2892. https://doi.org/10.1109/TBME.2022.3156963
17. Khan, A. H., Hussain, M., & Malik, M. K. (2021). Cardiac disorder classification by electrocardiogram sensing using deep neural network. Complexity, 2021(1), 5512243.
18. Mahmud, T., Barua, A., Islam, D., Hossain, M. S., Chakma, R., Barua, K., Monju, M., & Andersson, K. (2023). Ensemble Deep Learning Approach for ECG-Based Cardiac Disease Detection: Signal and Image Analysis. 2023 International Conference on Information and Communication Technology for Sustainable Development, ICICT4SD 2023 - Proceedings, 70–74. https://doi.org/10.1109/ICICT4SD59951.2023.10303625
19. Weiss, A. J., Yadaw, A. S., Meretzky, D. L., Levin, M. A., Adams, D. H., McCardle, K., Pandey, G., & Iyengar, R. (2023). Machine learning using institution-specific multi-modal electronic health records improves mortality risk prediction for cardiac surgery patients. JTCVS Open, 14(June), 214–251. https://doi.org/10.1016/j.xjon.2023.03.010
20. Deshwal, D., Sangwan, P., Dahiya, N., Lilhore, U. K., Dalal, S., & Simaiya, S. (2023). COVID-19 Detection using Hybrid CNN-RNN Architecture with Transfer Learning from X-Rays. *Current Medical Imaging Reviews*, *20*, 1–17. https://doi.org/10.2174/1573405620666230817092337
21. Joy, S. I., Kumar, K. S., Palanivelan, M., & Lakshmi, D. (2023). Review on Advent of Artificial Intelligence in Electrocardiogram for the Detection of Extra-Cardiac and Cardiovascular Disease. IEEE Canadian Journal of Electrical and Computer Engineering, 46(2), 99–106. https://doi.org/10.1109/ICJECE.2022.3228588
22. Venkatesan, C., Thamaraimanalan, T., Ramkumar, M., Sivaramakrishnan, A., & Marimuthu, M. (2023). ECG Signal Feature Extraction and SVM Classifier Based

Cardiac Arrhythmia Detection. 2023 2nd International Conference on Electrical, Electronics, Information and Communication Technologies, ICEEICT 2023, 1–4. https://doi.org/10.1109/ICEEICT56924.2023.10157789

23. Budholiya, K., Shrivastava, S. K., & Sharma, V. (2022). An optimized XGBoost based diagnostic system for effective prediction of heart disease. Journal of King Saud University-Computer and Information Sciences, 34(7), 4514-4523.
24. Mohi Uddin, K. M., Ripa, R., Yeasmin, N., Biswas, N., & Dey, S. K. (2023). Machine learning-based approach to the diagnosis of cardiovascular vascular disease using a combined dataset. Intelligence-Based Medicine, 7(April), 100100. https://doi.org/10.1016/j.ibmed.2023.100100
25. Hannan, A., Cheema, S. M., & Pires, I. M. (2024). Machine learning-based smart wearable system for cardiac arrest monitoring using hybrid computing. Biomedical Signal Processing and Control, 87(PB), 105519. https://doi.org/10.1016/j.bspc.2023.105519
26. Sarvani, K. N. S., Nalluri, P., Srujana, A. B., & Ramtej, K. S. (2024). Prediction of Sudden Cardiac Arrest using Machine Learning. 2023 2nd International Conference on Automation, Computing and Renewable Systems (ICACRS), 965–969. https://doi.org/10.1109/icacrs58579.2023.10404350
27. Vahab, N., Bonu, T., Kuhlmann, L., Ramialison, M., & Tyagi, S. (2024). Uncovering co-regulatory modules and gene regulatory networks in the heart through machine learning-based analysis of large-scale epigenomic data. Computers in Biology and Medicine, 171(February), 108068. https://doi.org/10.1016/j.compbiomed.2024.108068
28. Wang, X., Rao, J., Zhang, L., Liu, X., & Zhang, Y. (2024). Identification of circadian rhythm-related gene classification patterns and immune infiltration analysis in heart failure based on machine learning. Heliyon, 10(6), e27049. https://doi.org/10.1016/j.heliyon.2024.e27049
29. Hudson, J. I., Hudson, Y., Kanyama, G., Schnabel, J., Javaras, K. N., Kaufman, M. J., & Pope, H. G. (2024). Causal factors in childhood and adolescence leading to anabolic-androgenic steroid use: A machine learning approach. Drug and Alcohol Dependence Reports, 10(August 2023), 100215. https://doi.org/10.1016/j.dadr.2023.100215