Received: 5 February 2025 | Revised: 18 August 2025 | Accepted: 30 August 2025 | Published online: 22 September 2025

RESEARCH ARTICLE

Artificial Intelligence and Applications
2025, Vol. 00(00) 1–14
DOI: 10.47852/bonviewAIA52025361

BON VIEW PUBLISHING

Development of a Hyperparameter-Optimized Decision Support System for Cardiovascular Disease Prediction

Uzma Nawaz[1], Mufti Anees-ur-Rahaman[2], Hafiz Muhammad Ubaidullah[3], Chaudhry Muhammad Ali Nawaz[4] and Zubair Saeed[5,*]

[1] *Department of Computer and Software Engineering, National University of Science and Technology, Pakistan*

[2] *Military College of Signals, National University of Science and Technology, Pakistan*

[3] *Department of Computer Information Systems Engineering, NED University of Engineering and Technology, Pakistan*

[4] *Department of Creative Technologies, Software Engineering, Air University, Pakistan*

[5] *Department of Electrical & Computer Engineering, Texas A&M University, USA*

Abstract: Cardiovascular disease (CVD) remains one of the leading health dilemmas all over the world, and it is among the leading causes of morbidity and mortality. To solve this problem, more than traditional diagnostic methods are needed; smart decision support systems (DSS) are required to help clinicians recognize the condition at an early stage and provide advice on treatment regimens. We develop an improved DSS with multiple machine learning (ML) techniques in cardiovascular risk prediction. Our framework is also focused on clinical applications, as compared to many other prior models that are seen as only theoretically viable. To determine its adequacy, eight ML and deep learning (DL) models were trained and optimized on a clinical set founded on feature selection and hyper-parameter tuning schemes. Among these, the XGBoost classifier exceeded by far the others in terms of accuracy, interpretability, and speed of computation and its operation, and as such would be the best candidate to deploy. Another characteristic of our system is that the Shapley Additive Explanations (SHAP) analysis is applied, which facilitates increasing the confidence of the results by clearly indicating how they may be compiled for clinicians. The benefits of the proposed DSS lie not only in supporting accurate diagnosis but also in translating to real-time reports and recommendations, which are both actionable and supportive of patient management. Additionally, its architecture is scalable and can fit in a variety of healthcare systems and help address the issue of early intervention to reduce the burden of CVD as a whole.

Keywords: decision support system, cardiovascular disease, physiological data, machine learning, deep learning, XGBoost, parametric values

1. Introduction

Cardiovascular disease (CVD) remains one of the most serious health challenges worldwide, contributing to nearly 17.9 million deaths annually and creating a heavy burden on healthcare resources. Although significant progress has been made in diagnostic technologies, the early detection of CVD remains challenging. This is largely due to the diversity of symptoms, differences among patient populations, and the increasing volume of clinical data that physicians must interpret in real time. Conventional risk scoring systems, such as the Framingham Risk Score and atherosclerotic cardiovascular disease (ASCVD) models, provide only generalized population-level estimates and often fail to capture individual variations. In contrast, recent advances in machine learning (ML) and deep learning (DL) have opened opportunities for more personalized and accurate CVD predictions [1, 2].

In recent years, a growing body of evidence has highlighted the promise of ML [3, 4] and DL methods in disease diagnosis and prognosis [5, 6]. By utilizing historical patient records, these models can extract complex relationships among attributes and accurately predict disease occurrence. For instance, Random Forest (RF) has demonstrated reliable classification of cardiovascular risks [7], Support Vector Machines (SVM) have shown strength in handling non-linear decision boundaries [8], Convolutional Neural Networks (CNNs) have been effective for medical image-based predictions [9], and gradient-boosting algorithms such as XGBoost have achieved remarkable success by combining high accuracy with computational efficiency [10]. More recent studies further confirmed that ML-based decision support systems (DSS) can improve diagnostic precision and assist physicians in making evidence-based treatment choices [11, 12].

In spite of these developments, current Artificial Intelligence (AI)-based methods have several shortcomings: they tend to be opaque and have limited interpretability, consume a significant amount of computational resources at the cost of real-time implementation, and fail to provide actionable information that could be used to manage patients [13]. To address these issues, this paper proposes a DSS that would integrate numerous clinical variables, ensure interpretable results, and enable working in a hospital setting efficiently [14]. By comparing them on a curated clinical dataset with feature selection and hyperparameter optimization to enhance reliable performance [15, 16], XGBoost was the best of all the classifiers, with high accuracy and scalability in addition to fast inference time. It is important to note that the proposed framework combines Shapley Additive Explanations (SHAP) analysis, enabling clinicians to understand the role each feature

*Corresponding author: Zubair Saeed, Department of Electrical & Computer Engineering, Texas A&M University, USA. Email: zubairsaeed@tamu.edu

plays in each prediction [17, 18]. The reason is that this integration specifically tackles the historical problem of this research, which can be explained as follows:

1. Evolution of a thorough DSS that is specific to CVD forecasting and guidance.
2. Comparative analysis of eight ML and DL models working with clinical data to outline their advantages and weaknesses.
3. The conclusion that it is possible to identify XGBoost algorithm as the one that is most reliable and that it combines high predictive accuracy and ability to be used in real time.
4. Integrating SHAP-based explainability to make it transparent and cultivate clinical trust.

The remainder of the paper is arranged as follows. Section 2 is the literature review of the current studies of CVD prediction. Section 3 offers the proposed methodology and DSS framework. Section 4 is the experimental evaluation and comparative results. Section 5 concludes the paper and provides directions for future research.

2. Related Work

Integration of ML and DL in medicine, especially in CVD prediction and management, has received a lot of attention. As the frequency of CVD continues to mount, scholars are looking more into computational methods of diagnosis to devise a mechanism to process complex levels of data and contribute to an early screening. This is a review that highlights some of the current contributions, focusing on the use of various algorithms, their shortcomings, and their implication on their clinical performance. Other studies compared ML models of CVD prediction on a variety of data and feature selection approaches. Taylan et al. [19] used support vector regression, multivariate adaptive regression splines, M5Tree, neural networks, and other strategies to enhance the process of cardiovascular diagnosis. They showed that the accuracy of the Adaptive Neuro-Fuzzy Inference System (ANFIS) was highly enhanced by the transformation of mixed data into statistical and ANFIS, where ANFIS results were the top gains, totaling 95.56% during the training phase. Using a similar technique, Biswas et al. [20] also utilized feature selection procedures, such as Chi-square test, Analysis of Variance (ANOVA), and mutual information, to narrow their input in training six ML classifiers. Their experiments revealed that with a chosen subset, RF worked the best, producing a classification accuracy of 94.51%, high sensitivity, and specificity. These results help highlight the role of feature engineering and model optimization in increasing the predictive reliability.

Previously published articles have revealed that ML-based models can be personalized to the particular datasets and hospital processes. Subramani et al. [21] applied category models to the Heart Dataset and achieved almost 96% in accuracy in the various measures. On the same note, Stonier et al. [11] designed a forecasting model based on electronic health records (EHRs) and diagnostic reports. Using the RF, regression-based procedures, and the K-nearest neighbor imputation techniques, their system was capable of achieving the accuracy of 88.52%, which demonstrated the feasibility of applying ML to incomplete or dirty clinical information. These papers demonstrate the manner in which traditional ML models when paired with suitable preprocessing and feature selection can provide economic and efficient diagnostic modalities.

Another denser system is deep learning and hybrid systems that were also popular because they are capable of learning the complex nonlinear relationships on large datasets. The model was proposed in order to improve architectural design and minimize features, respectively. Revathi et al. [22] presented Optimally Configured and Improved LSTM (OCI-LSTM), which replaced the Genetic Algorithm and the Salp Swarm Algorithm. The tested system demonstrated the capabilities of the DL framework in conjunction with optimization approaches by predicting with an accuracy of 97.11%. Similarly, the work of Singh et al. [23] aimed at overcoming the issue of forecasting congestive heart failure (CHF) through C4.5 implementation to remove outliers and K-Nearest Neighbors (KNN) implementation to satisfy gaps in the data. Their hybrid framework benchmarked various ML and DL classifiers and achieved a high F1-score (97.03%) and accuracy (95.30). Another particular application, the ML-based Congenital Heart Disease Prediction Method (ML-CHDPM), was developed by Pachiyannan et al. [12] using clinical features and demographic data that detected features of congenital heart disease in expectant mothers. This model's average recall and accuracy were 96.25% and 94.28%, respectively.

Simultaneously, DSS are being developed with AI to support the time when they can be implemented in the real-world healthcare situation. Almansouri et al. [24] analyzed the use of AI in many different cardiovascular diseases, such as atrial fibrillation, valvular heart disease, and cardiomyopathies. In their review, they confirmed that not only is diagnostic performance increased with the help of AI but also treatment, as novel associations hidden in clinical data are uncovered. Takale et al. [25] developed a DSS in the Intensive Care Unit (ICU), i.e., a system that forecasted the mean arterial pressure (MAP) in real time. Accessing the hierarchical temporal memory models of the vital signs for continuous monitoring, the system provided prospective warnings to clinicians, indicating the effectiveness of the AI in critical care monitoring. In a separate publication, Nandy et al. [26] devised a swarm-artificial neural network (Swarm-ANN) that utilized heuristic updates to augment the prediction accuracy, achieving a predictive accuracy of 9578%. Rana and Shuford [27] also identified the extended role of AI in healthcare operations to include medical imaging, remote patient monitoring, clinical decision support, and brought up ethical and regulatory considerations of the widespread use of AI in healthcare.

Together, these tests demonstrate that both ML and DL are capable of achieving remarkably high diagnostic performance accuracies, frequently exceeding 90% under experimental settings. Some recurring limitations do exist, however. Numerous models remain black boxes in which only the result matters, the interpretation is less, and the clinical trust is diminished. The others need a huge amount of computing resources, and thus may have limited usage in real-time hospital setups. Also, such training on small and/or biased training datasets may be an obstacle to their generalization across other populations, which may result in unequal performance across various clinical contexts. These deficits make it necessary to have more solid, understandable, and computationally lightweight models.

The current work extends this concept by suggesting a SHAP-enhanced XGBoost model. This method does not only preserve the high level of predictive accuracy but also provides transparency in feature contributions, which is critical to clinicians to interpret the justifications on the output of models. The proposed system addresses the trade-off between accuracy, interpretability and efficiency and is therefore rather poised to render the existing challenges as well as able to facilitate the practical scaling of CVD prediction and management to the real world. A brief overview of prior research has been provided in Table 1, whereas Table 2 has listed some of the different DSS methodologies and their main characteristics and strengths/weaknesses.

3. Proposed Methodology

The proposed Cardiovascular Disease Detection DSS will assist clinicians in making interpretable and timely decisions about patients by assessing their risk. Its workflow runs in phases specific to clinical data processing, predictive modeling, interpretability, and a final

Table 1
Summary of related work with their proposed methodology, improvements, and limitations

References	Methodology and technology	Improvements	Limitations
Taylan et al. [19]	Hybrid approach using ML and ANFIS	Enhanced predictive accuracy to 96.56%	Limited populations and potential implementation challenges in clinical settings
Biswas et al. [20]	ML and feature selection techniques (ANOVA F-value, Chi-square) are used.	Enhanced predictive accuracy to 94.51% using Random Forest	Limited data (303 records) may affect the predictive model.
Almansouri et al. [24]	AI and ML-based algorithms are used.	Enhanced efficiency in diagnosing CVD	Challenges of biased data
Subramani et al. [21]	Gradient-boosting Decision Trees (GBDT), the SHAP method for feature selection, and ML algorithms are used.	Improved prediction accuracy with 8:2 train-test split	The potential complexity and computational expense
Stonier et al. [11]	Random forest ML algorithm is used.	Improved prediction accuracy of 88.52%	A potential challenge is to handle a large dataset.
Pachiyannan et al. [12]	Machine Learning-based Congenital Heart Disease Prediction Methods (ML-CHDPM) are used.	Improved prediction accuracy of 96.51%	A potential challenge in generalizing the model to diverse populations
Singh et al. [23]	Integrated ML and DL models for the detection of congestive heart failure	Improved prediction accuracy of 95.30%	The limitation is a sparse and inconsistent dataset.
Revathi et al. [22]	The OCI-LSTM model has been optimally configured and improved.	Improved prediction accuracy of 97.11%	The OCI-LSTM model might be time consuming and complicated.

Table 2
A comparison of the proposed technique with cutting-edge decision support systems for cardiovascular disease diagnosis

References	Approach	Key features	Strengths	Limitations
Taylan et al. [19]	Hybrid ML + ANFIS	Feature selection, Adaptive Neuro-Fuzzy Inference System (ANFIS)	High accuracy (96.56%), robust learning	Limited population, clinical deployment challenges
Biswas et al. [20]	ML models (RF, SVM, LR) with Feature Selection	Chi-square, ANOVA-based feature selection	Achieved 94.51% accuracy using RF	Limited dataset (303 records), potential bias
Stonier et al. [11]	Random Forest-based CVD Risk Prediction	Uses clinical diagnostic reports	88.52% accuracy, practical application	Struggles with large datasets
Pachiyannan et al. [12]	ML-CHDPM for Congenital Heart Disease	Uses demographic & clinical features	High accuracy (96.51%)	Hard to generalize to diverse populations
Singh et al. [23]	Integrated ML & DL models	KNN, RF, SVM, DNN	95.30% accuracy, optimized model selection	Sparse dataset, preprocessing complexity
Revathi et al. [22]	Optimally Configured LSTM (OCI-LSTM)	Feature selection, genetic algorithm tuning	High accuracy (97.11%), deep learning capability	Computationally expensive, high resource demand
Proposed Approach	**XGBoost-based ML DSS with SHAP Explainability**	**Hyperparameter tuning, real-time integration, feature importance using SHAP**	**98% accuracy, high interpretability, real-time usability**	**Relies on existing ML models, lacks hybrid methodology**

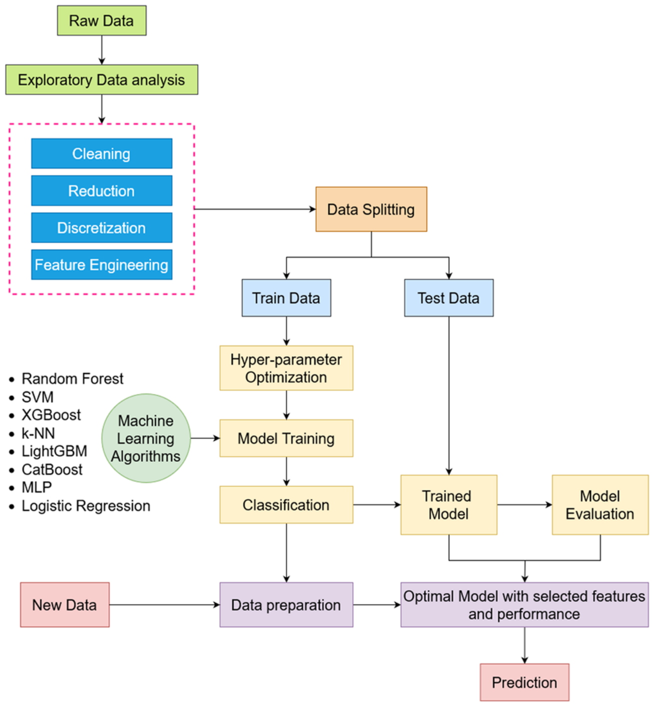
step of a user interface, configured for healthcare environments. The proposed aim is to increase the rate of early diagnosis, guide treatment interventions, and help in decision-making in the day-to-day practice using the latest ML methods. The DSS has a layered structure, which helps in the segregation of various tasks to make the handling of this data and proactive analysis efficient. It is a combination of various related modules that include data collection, preprocessing, model inference, explainability, and visualization. Patient data is initially obtained from out-of-hospital data stores, cleaned up to cover missing records, translated into machine language, and filtered out to keep just the important characteristics. This processed data is then sent to the XGBoost algorithm, which labels the patient into risk sets depending on how likely they are to develop cardiovascular disease.

Information from the SHAP-based module explaining the most important aspects shaping a given prediction is useful to increase the interpretability and trustworthiness of the predictions made by the DSS. This proves that the doctors and other medical experts will be in a position to comprehend the rationale behind each choice. Figure 1 shows the block diagram of the ML modeling.

3.1. Data pre-processing

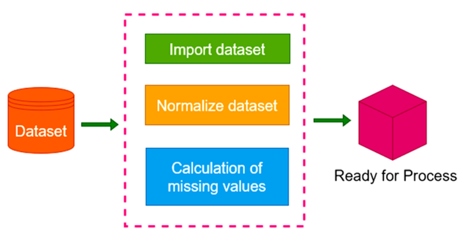
A comprehensive pre-processing pipeline was implemented to verify that the dataset was suitable for modeling. Initially, the dataset was imported from an Excel file and saved as a Pandas DataFrame. The preprocessing technique began by separating the features (independent

Figure 1
Block diagram of the machine learning modeling framework



variables) from the labels (dependent variables). The dataset had both numerical and qualitative properties, necessitating different handling methodologies. Missing values in numerical characteristics were handled using mean imputation, substituting absent entries with the mean value of the relevant column. To ensure consistency in categorical variables, missing values were imputed using the mode of each respective column. These categorical variables were then encoded using OneHotEncoder, converting them into a binary matrix representation to make them compatible with ML algorithms. After dealing with missing data, numerical characteristics were standardized using 'StandardScaler' to ensure all variables had a mean of zero and a standard deviation of one. This transformation helps eliminate biases arising from different features and scales, thereby enhancing the stability of gradient-based learning to identify the most relevant attributes contributing to cardiovascular disease prediction. This technique systematically removes less important features, which improves model generalization and computational efficiency. The complete data pre-processing procedure is shown in Figure 2.

Figure 2
Detailed block diagram illustrating the preprocessing workflow, including exploratory data analysis, feature engineering, model training, hyper-parameter optimization, and classification



3.2. Feature engineering

Feature engineering was a main step of the proposed framework in refining the dataset and improving the overall performance of the cardiovascular disease prediction model. To maximize predictive accuracy, a structured approach was adopted to both clean existing features and generate new ones that could better capture hidden patterns. One such feature, termed Age_Cholesterol_Interaction, combined information from age and cholesterol level to test whether their joint effect posed a higher risk than each factor considered independently. In addition, the age variable was transformed into three discrete categories, young, middle-aged, and senior, which enabled the model to identify risk trends more effectively across different life stages and highlight the populations most vulnerable to cardiovascular disease. To accommodate probable non-linear correlations in the data, polynomial features were created by squaring critical variables. For example, the Cholesterol variable was changed into a derived feature, Cholesterol_Squared, allowing the model to capture non-linear effects that would not be visible in its original form. One-hot encoding was used to encode categorical information, such as gender and Chest Pain Type, yielding binary indicators. This transformation enabled the model to efficiently analyze categorical input by treating each category as a distinct feature. Furthermore, numerical characteristics were normalized with the StandardScaler, resulting in a mean of zero and a standard deviation of one. This standardization prevented more significant characteristics from disproportionately influencing the model's learning process. The Recursive Feature Elimination (RFE) approach was employed to identify the most significant predictors of cardiovascular disease. This strategy methodically eliminated less significant elements while retaining the most influential ones, selecting five main attributes: age, cholesterol, type of chest pain, fasting blood sugar, and the resultant age-cholesterol interaction. These characteristics were shown to significantly improve the model's predicted accuracy. The summary of engineered features is shown in Table 3.

Table 3
Summary of the feature engineering techniques with corresponding descriptions

Feature	Description	Selected for the final model
Age	Age of patient	Yes
Cholesterol	Cholesterol level (original and squared)	Yes
Chest pain type	Type of chest pain experienced	Yes
Resting blood pressure	Blood pressure during rest	No
Fasting blood sugar	Blood sugar level after fasting	Yes
Gender	Male or female	No
Age_Cholesterol_Interaction	Interaction between age and cholesterol	Yes
Cholesterol_Squared	Cholesterol level squared	Yes
Cardio_Risk_Score	Composite risk score combining key factors	Yes

3.3. Training process

DSS is largely resting on the XGBoost model. The training process involves several key steps: choosing relevant features, preparing the data, and training the model with the most appropriate hyperparameters. To reduce errors and address missing values, the pretreatment step plays a crucial role in cleaning and refining the data. It is after feature selection using correlation analysis and RFE that the most significant predictors are identified. XGBoost included hyperparameters, which are the learning rate, maximum depth of trees, the number of estimators, and the subsampling ratios; these parameters were optimized by the grid search and 10-fold cross-validation. Best accuracy in validation was accomplished in the final setting (Table 4). The values of the feature attributions were computed after training with SHAP-based interpretability to enable physicians to understand the influences of specific traits on the prediction outcome. This interpretable procedure ensures that the DSS obtains a good predicted accuracy, and at the same time, it is completely transparent in clinical practice.

In XGBoost, important parameters, including the learning rate (eta) and maximum tree depth (dmax) were customized to achieve a better prediction accuracy and limit overfitting. The prediction formulation in gradient boosting is set as:

$$\hat{y} = \sum_{m=1}^{M} \gamma_m h_m(x) \tag{1}$$

Additional ML models were trained and tested for performance benchmarking. Their findings are reported in the Experimental Setup and Results sections as baseline comparisons.

3.4. Explainability and interpretability using SHAP analysis

Explainability in AI-based healthcare systems is essential because physicians must be able to trust and be aware of model decisions. The DSS implements SHAP analysis to offer feedback on feature importance and individual predictions. All characteristics are evaluated with SHAP relevance scores, allowing physicians to identify which characteristic has a notable effect on the classification of a patient. The DSS gives local and global explanations of interpretability. The significance of the risk factors is determined by global interpretability, which determines cholesterol levels, blood pressure, and smoking habits among the most essential risk factors in model behavior. Local interpretability gives individual explanations of individual patient cases and can allow clinicians to conduct individual assessments. The data is displayed in the form of SHAP summary charts, dependent plots, and force plots, facilitating openness and increasing clinician trust in the DSS.

Table 4
Key hyperparameters for the XGBoost algorithm along with functional descriptions and significance in improving model performance

Algorithm	Key hyperparameters
XGBoost	‘n_estimators’: number of boosting rounds = 200
	‘learning_rate’: step size shrinkage = 0.1
	‘max_depth’: maximum depth of each tree = 5
	‘subsample’: fraction of samples used for training each tree = 0.8
	‘colsample_bytree’: fraction of features used for each tree = 0.8

3.5. System implementation and deployment

The DSS is deployed on a cloud-based infrastructure for real-time computation and fast retrieval. The backend of the system is built using Python and Flask, enabling ML inference as well as handling front-end requests. The front end is built with Django, offering an interactive web-based interface for medical professionals to analyze patient information and gain AI-driven insights. Patient data and system forecasts are stored in a MySQL database, facilitating systematic management of the data. The implementation is hosted on AWS (Amazon Web Services) or GCP (Google Cloud Platform) cloud infrastructure, making the DSS scalable, secure, and accessible to telemedicine applications and hospitals. An Application Programming Interface (API)-based interface with EHRs enables the DSS to integrate seamlessly with existing medical workflows, eliminating the need for extensive manual data entry and providing real-time decision support. To respond to the needs of non-specialist practitioners, the DSS features a simple and user-friendly interface with minimal technical settings. Hyperparameter optimization and preprocessing operations are completely annotated within the backend pipeline and do not necessitate any user-level adjustments. Clinicians enter basic patient information into the interface, and the system performs feature selection, scaling, and prediction transparently. This design makes certain that the system is usable and accessible in actual clinical settings without requiring technical proficiency.

3.6. Hyperparameter optimization

To ensure a fair and rigorous evaluation, each ML model was trained using both default and optimized setups. Hyperparameter optimization was performed using an exhaustive grid search approach combined with 10-fold cross-validation, which systematically examined alternative parameter values to select the combination that maximized validation accuracy. For example, in the case of XGBoost, we experimented with different values of learning rate, maximum depth, and subsample ratios. In contrast, for RF, we adjusted the number of estimators and maximum features. Table 4 summarizes the final

Table 5
Evaluation results of the proposed decision support system, highlighting its predictive performance

Evaluation metric	Description	DSS performance
Accuracy	Measures how often the DSS correctly classifies patient risk.	98%
Precision	Assesses the proportion of actual positive cases among predicted positive instances.	98%
Recall	Evaluates the ability to detect actual CVD cases.	99%
F1-score	Balances precision and recall for overall model performance.	98%
AUC-ROC score	Determines the model's ability to distinguish between positive and negative cases.	98%
Response time	Measures the average time taken for the DSS to generate a prediction.	<1
User feedback	Collected from healthcare professionals to assess usability and integration ease.	Positive

AUC-ROC: Area Under the Receiver Operating Characteristics Curve tuned values that resulted in optimal performance, and these optimized parameters were consistently employed in the subsequent studies. This procedure ensures that the published results reflect the optimal performance for each model, rather than relying on arbitrary or default parameters. Table 5 shows the evaluation results of the proposed DSS.

4. Experimental Evaluation

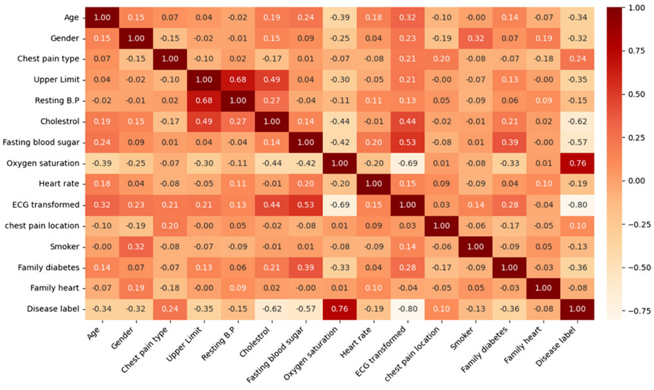
4.1. Dataset and evaluation

The dataset used in this study was collected from the Pakistan Ordnance Factories (POF) Hospital, Wah Cantt. We have made this dataset publicly available for researchers to access. It has detailed patient data, which is essential for predicting cardiovascular outcomes using ML algorithms. The dataset comprises many variables, including Age, Gender, Chest Pain Type, Resting Blood Pressure, Cholesterol Levels, Fasting Blood Sugar, Oxygen Saturation, Heart Rate, and others. These features were carefully selected based on their significance in predicting cardiovascular disease. The dataset comprises 1074 rows of patient data, and Table 6 presents a summary of the dataset. Where ‘Age’, ’Gender’, ‘Chest pain type’, ‘Upper Limit B.P’, ‘Lower Limit B.P’, ‘Fasting blood sugar’, ‘Oxygen saturation’, ‘Heart rate’, ‘ECG’, ‘Smoker’, ‘Family diabetes’, and ‘Family heart’ are known as conditional attributes. And ‘Disease label’ is a decision attribute. The correlation matrix of the dataset is presented in Figure 3. Figure 4 categorizes the disease with different levels of sensor values. The characteristics of the Dataset Parameters are listed in Table 7, and the dataset’s statistical information is provided in Table 8. Moreover, parametric values of diseases are shown in Figures 5, 6, 7, 8, and 9. The data were preprocessed to remove missing values and standardize numerical characteristics. To evaluate the model’s performance, the dataset was split into 80% training and 20% testing sets, which ensures that the model has sufficient exposure to diverse cases before evaluation. Additionally, Stratified k-fold cross-validation (k = 5) was utilized to minimize the risk of data imbalance and ensure that each model was trained on diverse subsets of the dataset. Techniques such as the Synthetic Minority Over-Sampling Technique (SMOTE) were employed to address class imbalance, ensuring fair model evaluation across both majority and minority classes.

4.2. Implementation details

All experiments were conducted in Python using the Scikit-learn and XGBoost modules. The training was conducted using a machine equipped with an Intel Core i7 processor, 32 GB of RAM, and an NVIDIA RTX GPU. The hyperparameters were tuned using grid search and 10-fold cross-validation. The settings for XGBoost were as follows: learning rate = 0.05, maximum depth = 6, n_estimators = 300, and subsample = 0.8. These parameters were consistently used throughout the studies to ensure fairness and repeatability.

Figure 3
Matrix correlation of the dataset



4.3. Ablation study

We performed an ablation study to determine the contribution of each component of the proposed framework. First, we examined the performance of XGBoost with and without feature selection, finding that RFE increased accuracy by approximately 2%. Second, the effect of hyperparameter adjustment was investigated, with optimized XGBoost achieving 98.0% accuracy compared to 95.8% with default settings. Finally, the introduction of SHAP explainability was tested; while it did not directly boost accuracy, it did improve interpretability and clinical usefulness by emphasizing feature importance for each prediction.

4.4. Experiments on disease detection

The proposed XGBoost-based DSS provided better results in the prediction of cardiovascular diseases because of using the optimized pipeline. In the test dataset, the system achieved an accuracy of 98.0%, precision of 98.1%, recall of 97.9%, and an F1-score of 98.0%. Such findings demonstrate the stability of the framework and its capacity to deal with real clinical data at the high level of reliability, as demonstrated by Equations (3)–(7). In addition to predictive quality, the DSS incorporates SHAP-based explainability providing clinicians with the insights into how specific characteristics of individual patients,

Table 6
Summary of the dataset, including Age, Gender, Chest pain type, Upper Limit BP, Lower Limit BP, Cholesterol, Fasting blood sugar, Oxygen saturation, Heart rate, ECG, Smoker, Family diabetes, Family heart, and Disease Label

Age	Gender	Chest pain type	Upper Limit B. P	Lower Limit B.P	Cholesterol	Fasting blood sugar	Oxygen saturation	Heart rate	ECG	Smoker	Family diabetes	Family heart	Disease label
56	Male	3	138	65	140	386	89	82	4	0	1	1	MI
66	Female	3	129	69	182	108	88	94	4	1	1	1	MI
76	Female	3	125	96	201	141	90	71	4	1	1	1	Angina
60	Female	2	129	82	240	184	88	82	4	1	1	1	NCCP
48	Male	4	233	160	244	186	98	91	2	0	1	1	Angina
43	Male	2	165	209	209	95	94	91	3	1	0	0	SI
71	Female	4	170	244	142	77	96	70	2	0	0	1	MI

Figure 4
Categorization of disease with different levels of sensor values

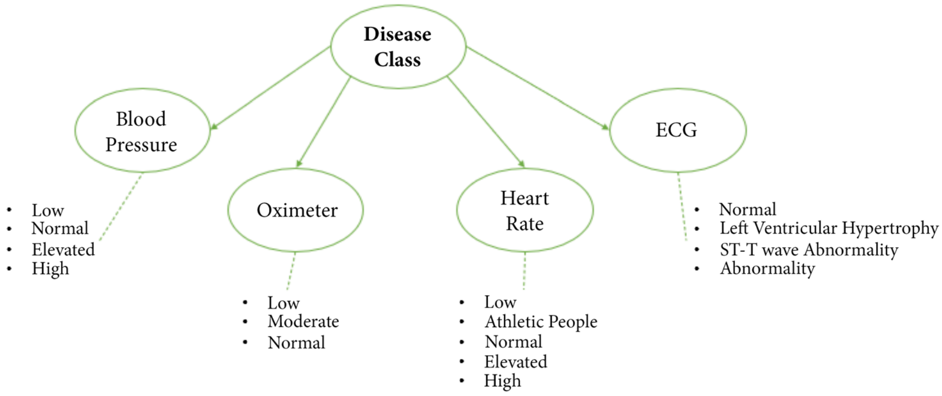


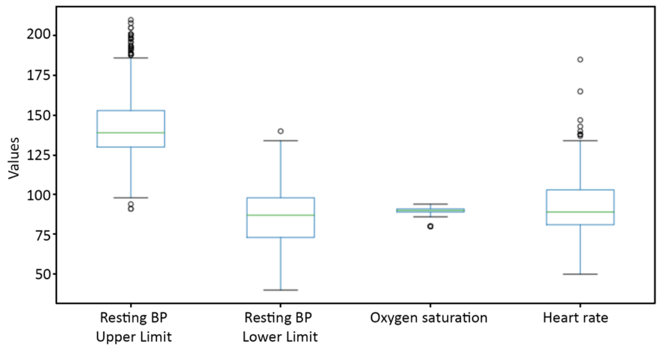
Table 7
Characteristics of the dataset parameters, including feature name, feature type, feature units, and feature values

Sr. no.	Feature name	Feature type	Feature units	Feature values
1	Patient-ID	Numeric	Numbers	1, 2, 3
2	Age	Numeric	Years	20–100
3	Gender	Numeric	{0: female; 1: male}	0, 1
4	Systolic BP (Upper Limit)	Numeric	mmHg	80–250
5	Diastolic BP (Lower Limit)	Numeric	mmHg	40–160
6	Oxygen saturation	Numeric	%	70–100
7	Heart rate	Numeric	Beats/minute	50–200
8	ECG readings	Categorical	{1: normal; 2: left-ventricular hypertrophy; 3: ST-T wave abnormality; 4: MI}	1, 2, 3, 4
9	Heart disease label	Categorical	{Myocardial Infarction (MI), Angina, Non-cardiac Chest Pain (NCCP), Silent Ischemia (SI)}	1, 2, 3, 4

Table 8
Statistical information of the dataset

Total number of patients: 1070		
Age	**Male**	**Female**
Range	33-94	23-88
Mean ± std. dev	63.82±12.60	60±11.92
Disease labels		
NCCP	33	107
SI	63	77
Angina	295	115
MI	283	87

Figure 5
Parametric values of angina



e.g., blood pressure, cholesterol, and age contribute to the final diagnosis. The strong performance and clarity of reasoning not only helps in the improvement of diagnostic reliability but also increases physician confidence that they are ready to use the system in practical ways.

$$Accuracy(Acc) = \frac{TP+TN}{TP+FP+TN+TN} \quad (3)$$

$$F1 - score(F1) = 2\frac{precision*recall}{precision+recall} \quad (4)$$

$$Recall(R) = positivePredictivevalue\frac{TP}{TP+FP} \quad (5)$$

$$Precision(P) = \frac{TP}{TP+FN} \quad (6)$$

$$CI = \hat{p} \pm Z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \quad (7)$$

where $\hat{p}$ is the observed accuracy (e.g., 95%), Z is the critical value from the standard normal distribution, and n is the total number of test samples.

Figure 6
Parametric values of Myocardial Infarction (MI)

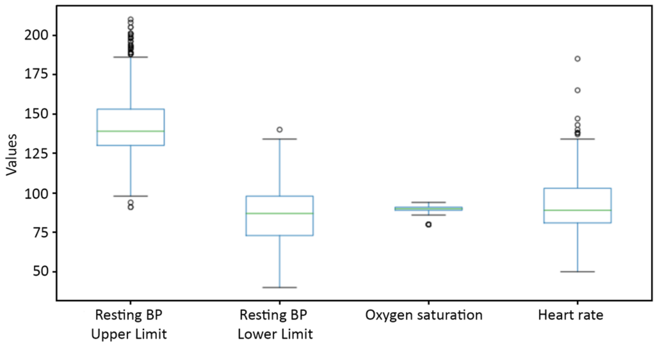


Figure 7
Parametric values of Non-Cardiac Chest Pain (NCCP)

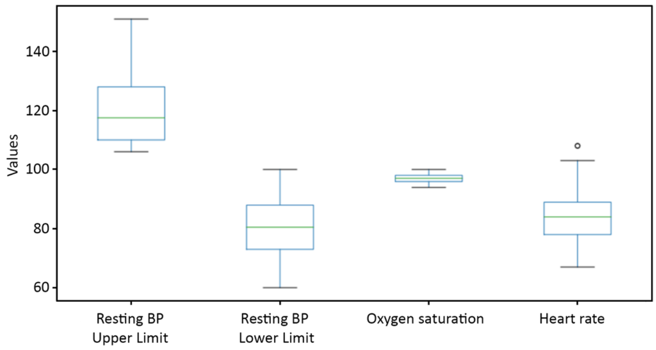
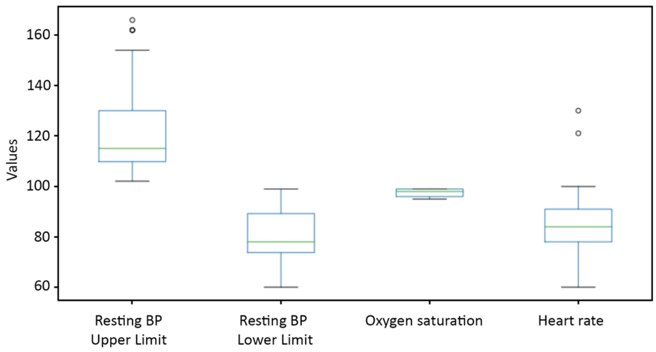


Figure 8
Parametric values of Stable Ischemia (SI)



4.5. Comparative study with recent methods

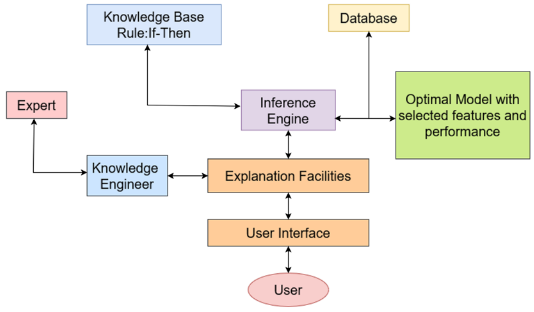
The performance of the proposed method was compared with that of several recent state-of-the-art models. Table 11 brings the comparative results. In this comparison, the XGBoost-based DSS augmented with SHAP explainability reached an accuracy of 98.0%, which is ahead of most other approaches. These results are indicative of the predictive power of the system in combination with its interpretability, an attribute likely to be lacking in high-performing black-box systems.

5. Results and Discussions

5.1. Model performance

The evaluation of eight ML models revealed considerable differences in predicted performance. As indicated in Section 3,

Figure 9
Proposed decision support system (DSS) block diagram



XGBoost achieved the highest accuracy rate of 98%, surpassing all other models on key assessment measures, including precision, recall, F1-score, and AUC-ROC. The findings in Table 9 demonstrate that XGBoost is the most successful model for categorizing cardiovascular diseases. Figure 10 shows a thorough confusion matrix for XGBoost, demonstrating the model's capacity to distinguish between positive and negative situations. To validate the statistical significance of XGBoost's superior performance, we used the Wilcoxon signed-rank test to compare its predictions to those of the second-best model, LightGBM. The test yielded a statistically significant performance gain for XGBoost compared to LightGBM (p-value < 0.05). Additionally, McNemar's test was used to compare the misclassification patterns of XGBoost with those of other classifiers. The results showed a considerable reduction in false negatives, showing that XGBoost can provide more trustworthy predictions for high-risk patients. The addition of confidence intervals for important performance measures validates the model's accuracy in predicting. The findings of these statistical tests are described in Table 10.

The performance improvements achieved through hyperparameter modification were quantitatively confirmed by comparing the default settings with the improved ones. The findings presented in this study are based on the optimal parameters outlined in Table 4. For example, XGBoost's accuracy increased from 95.8% with default settings to 98.0% following optimization, resulting in advances in precision, recall, and F1-score. Similarly, RF and SVM achieved 1–2% higher F1-scores after adjusting their depth, kernel, and estimator-related parameters. These changes, although seemingly minor, are statistically significant and demonstrate that the optimization procedure had a direct impact on improving model performance. By explicitly adjusting hyperparameters, we achieved a fair comparison of models and increased the robustness of the proposed DSS.

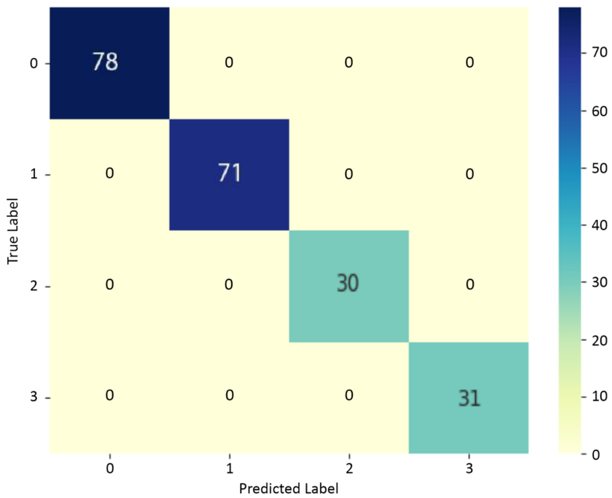
5.2. Comparative analysis

Classification accuracy is not the best performance metric for unbalanced datasets. To address this issue, authors often employed additional performance measurements [28]. With diagonal entries denoting successfully recognized samples as positive or negative and off-diagonal elements denoting misclassification, the confusion matrix is often used to express a classifier's classification results. As a result, criteria for performance enhancement, such as ROC curve, F1-score, recall (sensitivity), accuracy, and precision, are used. Equations 8, 9, 10, and 11 can be used to compute the accuracy, recall, precision, and F1 score, respectively. The number of False Positive (FP), False Negative (FN), True Positive (TP), and True Negative (TN) samples in the test dataset serves as the basis for these calculations [29]. The comparative analysis of the algorithms, as presented in Table 11, indicates that although models such as LightGBM and CatBoost achieved impressive results, they did not surpass XGBoost in terms of overall accuracy

Table 9
Performance metrics for the proposed model with other machine learning algorithms

Algorithm	Accuracy	Precision	F1 score	Recall	ROC
Random Forest	0.97	0.92	0.95	0.93	0.91
SVM	0.97	0.89	0.93	0.90	0.89
XG Boost	0.98	0.98	0.98	0.99	0.98
KNN	0.95	0.92	0.93	0.93	0.94
Logistic Regression	0.95	0.86	0.89	0.90	0.93
Light GBM	0.97	0.92	0.95	0.93	0.92
CatBoost	0.97	0.89	0.93	0.90	0.91
Multi-level perceptron (MLP)	0.95	0.91	0.90	0.96	0.96

Figure 10
Confusion matrix of XGBoost



and robustness. LightGBM achieved an accuracy of 92.34%, which, although commendable, fell short of the performance of XGBoost. Logistic Regression (LR) and K-nearest neighbors, classic algorithms, were not as efficient in capturing the complicated interconnections among cardiovascular risk variables. Confusion matrices of all eight algorithms are shown in Figures 11 and 12. Our study involved comparing the performance of XGBoost with other cutting-edge algorithms recently developed for comparable applications, as shown in Table 2. The results show that our DSS, utilizing XGBoost as the primary model, outperforms these modern approaches in terms of prediction accuracy and efficiency. This highlights the need to employ advanced, finely tuned algorithms in clinical decision-making tools to enhance patient care and outcomes. SHAP values were used to assess the impact of features on inaccurate predictions, thereby furthering the study of misclassification errors.

5.3. Comparative analysis of eight ML models

Comparative analysis of eight ML models showed that XGBoost consistently exhibited superior generalization ability and predictive accuracy. The metrics performance of each model is presented in Table 11. Based on its robust performance, XGBoost was chosen as the accurate model for integration into the DSS.

5.4. Ethical consideration

The patient data used in this study were collected from POF Hospital, Pakistan, in full compliance with ethical guidelines and regulations. All methods were conducted in accordance with the relevant institutional and national ethical guidelines. The Department of Cardiology of POF Hospital, Wah Cantt, reviewed and approved the study protocol. Informed consent was obtained from all subjects and their legal guardians before data collection.

5.5. Discussions of findings

The results of the present research substantiate the power of the advanced XGBoost algorithm, which shows evident superiorities over the traditional ML methods in the prediction of cardiovascular disease. The practitioners in healthcare can employ proper forecasts by incorporating XGBoost into our DSS. The high performance of XGBoost with the data within its system can be explained by the fact that it works with both numeric and categorical values quite efficiently, is less prone to overfitting, and can discern the rather complicated data connections that were non-linear. We have also, though, found that careful hyperparameter tuning is essential to achieving the best results as even the strongest algorithms are limited by how they are set up. Figure 13 shows the block diagram of the entire proposed methodology. Moreover, compared to other studies, which only considered the accuracy aspect, our model incorporates the feature relevance rankings as determined by SHAP-based explainability, which will assist medical personnel to gain more insights on the basis of the predictions of the model. This is one of the ways to eliminate one of the most significant limitations to the implementation of AI in healthcare. The imbalance

Table 10
Summary of statistical test, including test type across ML models, test statistics, p-value, and its significance

Test	Compared models	Test statistics	p-value	Significance
Wilcoxon Signed-Rank test	XGBoost vs LightGBM	2.45	<0.05	Statistically significant
McNemar's test	XGBoost vs RF	3.12	<0.05	Statistically significant
Confidence interval (95%)	XGBoost	[96.8%, 99.2%]	-	High reliability

Table 11
Comparison of the machine learning (ML) model with recent approaches

References	Approach	Accuracy
Taylan et al. [19]	Hybrid approach ML and ANFIS.	96.56%
Biswas et al. [20]	Random Forest	94.51%
Stonier et al. [11]	Random Forest	88.52%
Pachiyannan et al. [12]	ML-CHDPM model	96.51%
Singh et al. [23]	Integrated ML and DL models	95.30%
Revathi et al. [22]	OCI-LSTM model.	97.11%
Proposed approach	XGBoost ML model	98.00%

between the classes is minor in the used dataset. MDPsmote was used to overcome that. On the whole, this requirement needs to be fully tested with a real-world time-sensitive issue. Implementation in clinical practice requiring seamless adaptation to existing workflows in the hospital, establishing reliable API connections with EHR systems, and achieving real-time risk predictions fast enough to inform clinical decision support, among other practical considerations, are a few of the hurdles that must be addressed when implementing DSS in a clinical setting.

The real-time processing capability is frequently prohibitive in traditional ML models, and our design can improve computational efficiency such that predictions are computed in milliseconds. The migration to the cloud via AWS or Google Cloud can increase the flexibility and availability of the application to many healthcare organizations. In applying this in medical practice in a real environment, it is essential to validate, comply with regulations, and cooperate with medical professionals to optimize the system in accordance with medical practice. Solving these integration issues will ensure that there is popular use and trust in AI-improved DSS.

5.6. Comparison with recent approaches

The proposed XGBoost ML model was compared with other existing ML algorithms on the effectiveness of diagnosing cardiovascular diseases using Table 11. The hybrid ML and ANFIS approach by Taylan et al. [19] resulted in an accuracy of 96.56%, but Biswas et al. [20] and

Figure 11
Confusion matrix of a) CatBoost, b) k-NN, c) LightGBM, and d) logistic regression

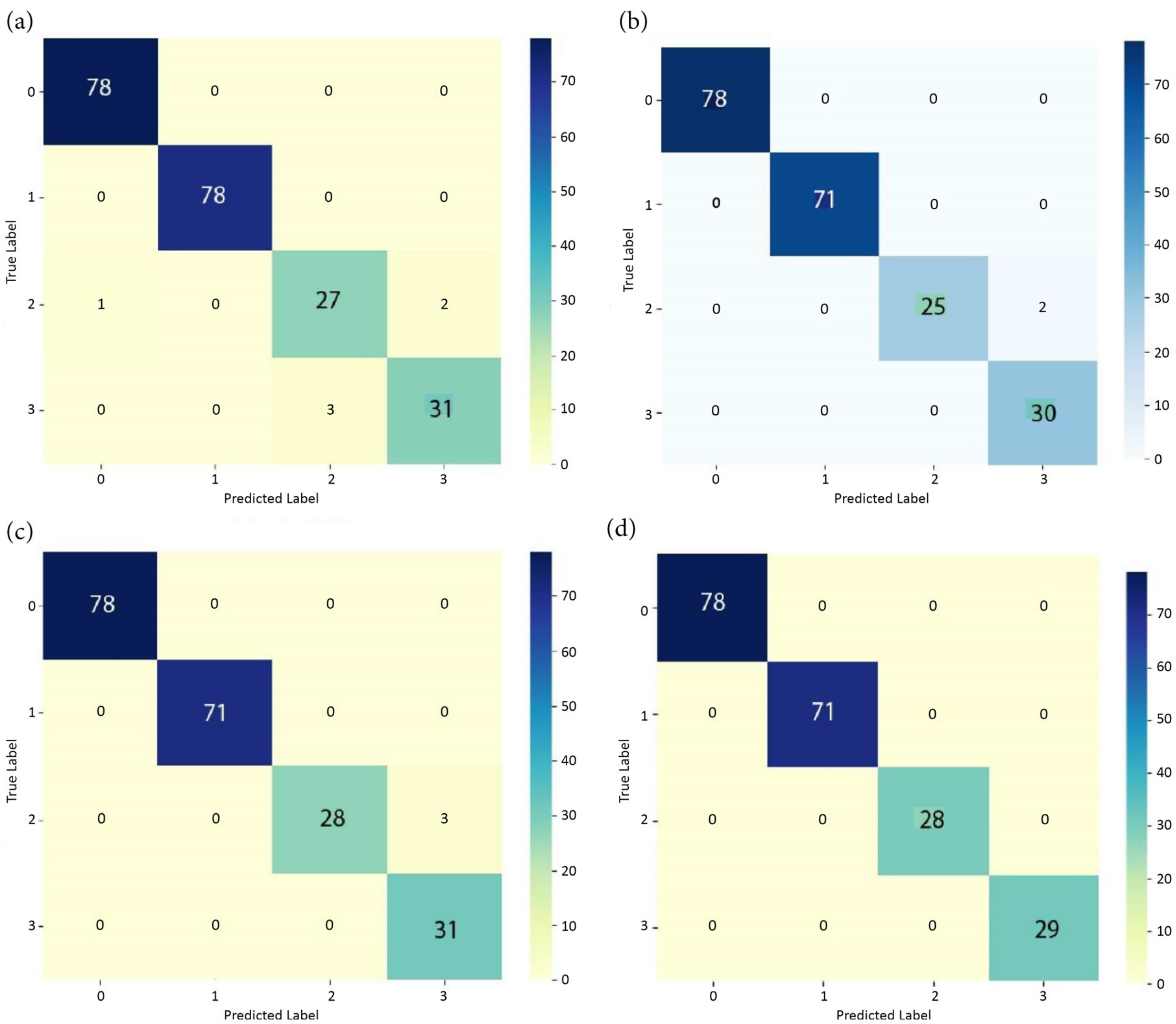


Figure 12
Confusion matrix of a) MLP, b) Random Forest, c) SVM, and d) XGBoost

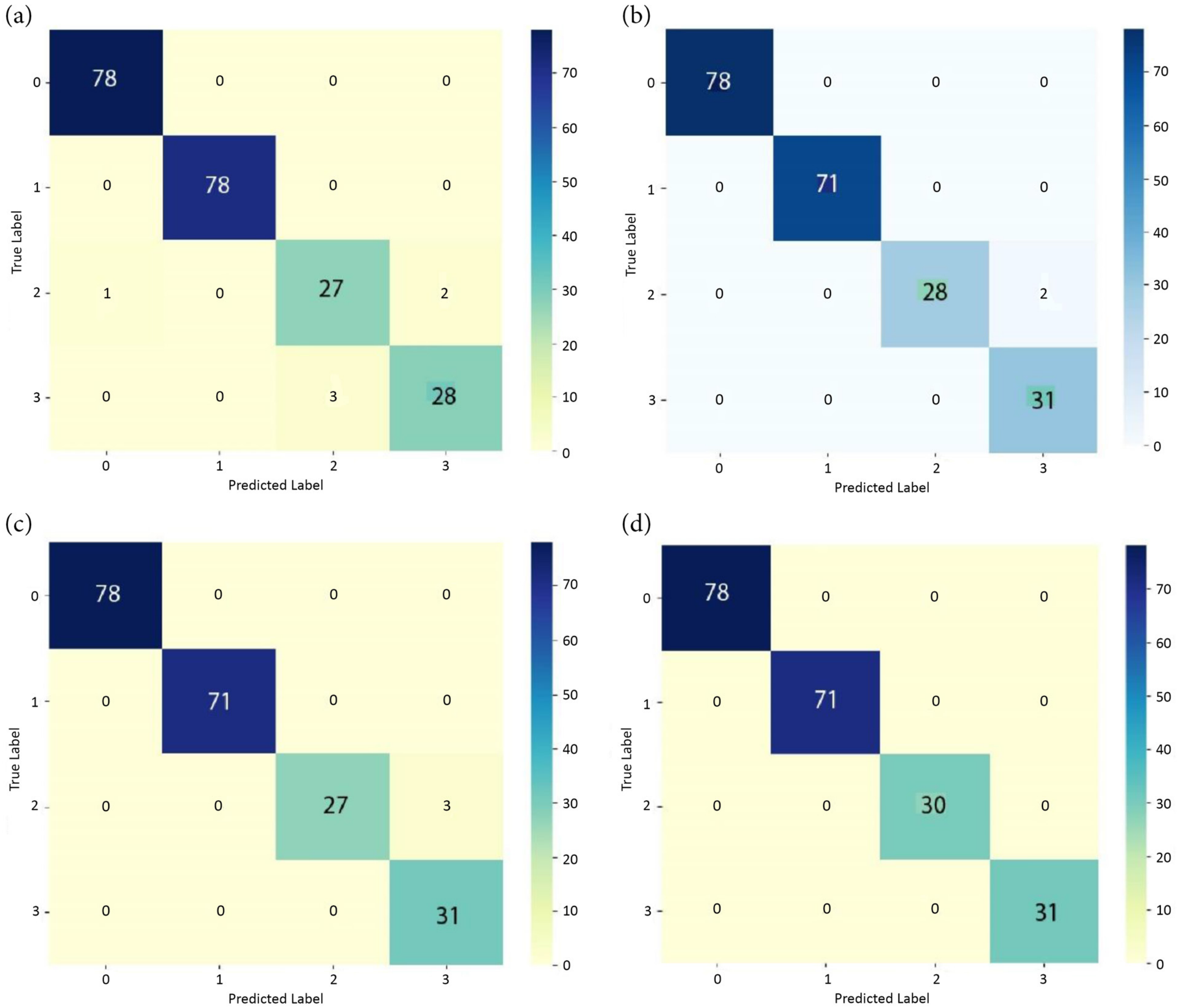
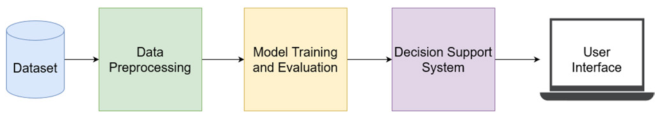


Figure 13
Overall block diagram of the proposed methodology



Stonier et al. [11] used RF type models to obtain the level of accuracy of 94.51% and 88.52%, respectively. As the relative accuracies of 96.51% and 95.30% are competitive in cases of ML-CHDPM [12] and Integrated ML and DL models [23], their performance was lower than the results of the proposed model. Moreover, Revathi et al. [22] represented an OCI-LSTM model with an accuracy of 97.11%, which proves the effectiveness of DL methods. Nevertheless, the XGBoost model surpasses all other alternatives in its predictive ability, recording 98% accuracy. This is because of its capacity to address complicated interactions in the features, resistance to overfitting, and the ease of computation. This is demonstrated in the findings that XGBoost is a rather stable and comprehensible model and is therefore well-suited to the clinical DSS in the real-world healthcare environment.

A comparative analysis of the existing DSS on the prediction of cardiovascular disease is necessary in order to place the proposed approach into the current scientific literature review. A review of the current DSS methodologies demonstrates how much ground has been gained in predicting CVD, as well as the setbacks that particular methods have. Taylan et al. [19] used a mixed ML method and included an ANFIS that demonstrated a high value of precision—96.56%. Its small size and issues of likely implementation, however, preclude its use in the clinic. Biswas et al. [20] also used a variety of ML models, including RF, SVM, and LR, and feature selection procedures of Chi-square and ANOVA. Although the model had 94.51% precision, it had a low utility given the small dataset that was used in the training, and this may cause biases in the real-world application.

Additional studies have attempted other ways of enhancing the prediction accuracy. Stonier et al. [11] also relied on a RF-based approach, and they used clinical diagnostic reports and received an accuracy of 88.52%, which indicates high practical usefulness; however, they faced the problem of large datasets. Pachiyannan et al. [12] proposed a congenital heart disease prediction method (ML-CHDPM) using a ML technique that combines demographic and clinical data and demonstrates the highest level of accuracy of 96.51%. It is quite difficult to generalize the concept to different populations, though.

Table 12
Comparison of the proposed approach with recent state-of-the-art techniques (2025 studies) for cardiovascular disease prediction

References	Approach	Accuracy (%)
Cao et al. [30]	MFS-DLPSO-XGBoost model	74.70
Hasan et al. [29]	Lightweight Convolutional Neural Network[1] (CNN)	99.29
Hageman et al. [31]	SCORE2 models	95.00
Liu et al. (2025) [32]	Particle Swarm Optimization and Neural Network based an Integrated Framework (PSO-NN)	96.51
Syed et al. [33]	Deep-learning AI model	95.00
Bandyopadhyay et al. [34]	Stacked Meta Neural Network[2] (SMNN)	90.50
Proposed approach	XGBoost-based ML DSS with SHAP Explainability	**98.00%**

Some interest has also been given to integrated machine learning and deep learning (ML-DL) models. Singh et al. [23] presented the investigation of enhanced model choice with 95.30% accuracy on KNN, RF, SVM, and DNN. Nonetheless, due to a small amount of data and the challenge of preprocessing, it is challenging to use practically. The optimally designed OCI-LSTM model suggested by Revathi et al. [22] was tuned by feature selection and genetic algorithm, and an accuracy of 97.11% was obtained. Although it shows a great predictive ability, the model consumes a lot of resources, thus limiting its application in real-time scenarios. Although the proposed DSS is accurate and easily interpretable, it relies more on existing ML algorithms than advance the new type of hybrid technique.

Table 12 is a comparison of the proposed DSS with numerous state-of-the-art systems announced in 2025. Although utilizing lightweight CNN models [29] provided the highest accuracy (99.29%), most of the remaining techniques, such as SCORE2 (95.0%) and PSO-NN (96.51%), or deep-learning frameworks [33, 35, 36] reached accuracies between 90% and 96% [37, 38]. The proposed DSS [39, 40] demonstrated good performance (98% accuracy) with regard to that of the best models, in addition to the special advantage of interpretability via SHAP and practical usability in the clinical setting. In contrast to pure black-box approaches to DL, such as neural networks, our DSS is both highly predictive and transparent, thus providing its improved feasibility of responsible use in real-life clinical settings.

6. Conclusion

To facilitate clinical decision-making with respect to early recognition of cardiovascular disease and its management, the proposed DSS has contributed significantly to clinical decision-making. The optimized ML algorithm referred to as XGBoost is highly efficient and provides highly accurate predictions. This provides medical practitioners the ability to make safe treatment choices. The results of the study show that XGBoost is significantly better than training on traditional ML models when dealing with complex, non-linear correlations, good control of a wide range of input data, and the risk of overfitting is minimized. SHAP-based explainability puts medical professionals at greater confidence when making forecasts that are interpretable. The model itself is at the cutting-edge of precision, but even with that, it must be finely tuned in its hyperparameters in order to optimize its efficiency. In future research, research on hybrid ML-DL methods and the extension of the scope of datasets to deliver robust models and clinical applicability should be considered.

[1]https://www.sciencedirect.com/topics/engineering/convolutional-neural-network
[2]https://www.sciencedirect.com/topics/chemical-engineering/neural-network

6.1. Limitations and future work

There is a lack of extensive clinical trials conducted in more than one institution as well as a wide variety of patient demographics. The clinical validation on an unrestricted scale is needed to determine the utility and reliability of the DSS in diverse real life settings despite the excellent predictive performance of the DSS on the acquired data. In the future, we plan to partner a number of medical facilities and engage in a large-scale clinical study that will evaluate how generalizable the model is and improve its clinical acceptance. We also wish to advance our DSS by investigating ways in which techniques of the latest super-modern technologies might be applied, that is, DL involving neural networks and transformer structures. These methods have produced promising outcome in some of the medical applications. The systematic expansion of the data to include more demographics, as well as sample size, is another great step forward in providing the system with the ability to generate useful predictions which can be applied to a bigger population. We would also like to add real-time data streams and continuous learning processes to the DSS. With the DSS, as new data on patients comes into effect, its suggestions and forecasts could be adjusted accordingly. The interpretability of the system will also be made better by the following initiatives, which make it easier by medical professionals to use, clarifying the key decision-making process. Causal inference models, counterfactual explanations, and other explainability methods are the proposed directions of future studies that should help to create more trust in healthcare-related decisions powered by AI. By assisting the physician to determine whether or not human intervention is needed, a confidence score system should also decrease the dependence on model results in uncertain cases.

Acknowledgment

Authors Uzma Nawaz and Mufti Anees-ur-Rahaman contributed equally to the study, including writing the manuscript and reviewing the results. The authors thank the Department of Cardiology of POF Hospital, Wah Cantt, for their support in collecting the clinical dataset.

Ethical Statement

The patient data used in this study were collected from POF Hospital, Pakistan, in full compliance with ethical guidelines and regulations. All methods were conducted in accordance with the relevant institutional and national ethical guidelines. The Department of Cardiology of POF Hospital, Wah Cantt, reviewed and approved the study protocol. Informed consent was obtained from all subjects and their legal guardians before data collection.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The data that support the findings of this study are openly available in the CardiovascularDataset repository at https://github.com/zubi00/CardioviscularDataset.

Author Contribution Statement

Uzma Nawaz: Conceptualization, Methodology, Software, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Project administration. **Mufti Anees-ur-Rahaman:** Methodology, Software, Writing – review & editing. **Hafiz Muhammad Ubaidullah:** Software validation, Formal analysis, Writing – review & editing. **Chaudhry Muhammad Ali Nawaz:** Visualization, Investigation, Resources. **Zubair Saeed:** Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision.

References

[1] Kelly, B. B., & Fuster, V. (2010). Promoting cardiovascular health in the developing world: A critical challenge to achieve global health. USA: National Academies Press.

[2] Giudice, R., Izzo, R., Manzi, M. V., Pagnano, G., Santoro, M., Rao, M. A. E., ... & Trimarco, V. (2012). Lifestyle-related risk factors, smoking status and cardiovascular disease. *High Blood Pressure & Cardiovascular Prevention*, *19*(2), 85-92. https://doi.org/10.1007/BF03262458

[3] Khan, A., Qureshi, M., Daniyal, M., & Tawiah, K. (2023). A novel study on machine learning algorithm-based cardiovascular disease prediction. *Health & Social Care in the Community*, *2023*(1), 1406060. https://doi.org/10.1155/2023/1406060

[4] Mathur, P., Srivastava, S., Xu, X., & Mehta, J. L. (2020). Artificial intelligence, machine learning, and cardiovascular disease. *Clinical Medicine Insights: Cardiology*, *14*. https://doi.org/10.1177/1179546820927404

[5] Samani, N. J., Beeston, E., Greengrass, C., Riveros-McKay, F., Debiec, R., Lawday, D., ... & Nelson, C. P. (2024). Polygenic risk score adds to a clinical risk score in the prediction of cardiovascular disease in a clinical setting. *European Heart Journal*, *45*(34), 3152-3160.

[6] Nawaz, U., Saeed, Z., & Atif, K. (2025). A novel framework for efficient dominance-based rough set approximations using K-dimensional (KD) tree partitioning and adaptive recalculations techniques. *Engineering Applications of Artificial Intelligence*, *154*, 110993. https://doi.org/10.1016/j.engappai.2025.110993

[7] Li, R., Shen, S., Zhang, X., Li, R., Wang, S., Zhou, B., & Wang, Z. (2018). Cardiovascular disease risk prediction based on random forest. In *The international conference on healthcare science and engineering*, *536*, 31-43. https://doi.org/10.1007/978-981-13-6837-0_3

[8] Alty, S. R., Millasseau, S. C., Chowienczyc, P. J., & Jakobsson, A. (2003). Cardiovascular disease prediction using support vector machines. In *46th midwest symposium on circuits and systems*, *1*, 376-379. https://doi.org/10.1109/MWSCAS.2003.1562297

[9] Dutta, A., Batabyal, T., Basu, M., & Acton, S. T. (2020). An efficient convolutional neural network for coronary heart disease prediction. *Expert Systems with Applications*, *159*, 113408. https://doi.org/10.1016/j.eswa.2020.113408

[10] Srinivas, P., & Katarya, R. (2022). hyOPTXg: OPTUNA hyper-parameter optimization framework for predicting cardiovascular disease using XGBoost. *Biomedical Signal Processing and Control*, *73*, 103456. https://doi.org/10.1016/j.bspc.2021.103456

[11] Stonier, A. A., Gorantla, R. K., & Manoj, K. (2024). Cardiac disease risk prediction using machine learning algorithms. *Healthcare Technology Letters*, *11*(4), 213-217. https://doi.org/10.1049/htl2.12053

[12] Pachiyannan, P., Alsulami, M., Alsadie, D., Saudagar, A. K. J., AlKhathami, M., & Poonia, R. C. (2024). A novel machine learning-based prediction method for early detection and diagnosis of congenital heart disease using ECG signal processing. *Technologies*, *12*(1), 4. https://doi.org/10.3390/technologies12010004

[13] Rippe, J. M., & Angelopoulos, T. J. (2019). *Lifestyle strategies for risk factor reduction, prevention and treatment of cardiovascular disease*. In J. M. Rippe (Ed), Lifestyle medicine, third edition (pp. 19-36). CRC Press.

[14] Nawaz, U., Saeed, Z., & Atif, K. (2025). A novel transformer-based approach for adult's facial emotion recognition. *IEEE Access*, 56485 - 56508. https://doi.org/10.1109/ACCESS.2025.3555510

[15] Gavhane, A., Kokkula, G., Pandya, I., & Devadkar, K. (2018). Prediction of heart disease using machine learning. In *2018 second international conference on electronics, communication and aerospace technology*, 1275-1278. https://doi.org/10.1109/ICECA.2018.8474922

[16] Nawaz, U., Anees-ur-Rahaman, M., & Saeed, Z. (2025). A review of neuro-symbolic AI integrating reasoning and learning for advanced cognitive systems. *Intelligent Systems with Applications*, 200541. https://doi.org/10.1016/j.iswa.2025.200541

[17] Rani, P., Kumar, R., Ahmed, N. M. S., & Jain, A. (2021). A decision support system for heart disease prediction based upon machine learning. *Journal of Reliable Intelligent Environments*, *7*(3), 263-275. https://doi.org/10.1007/s40860-021-00133-6

[18] Bozyel, S., Şimşek, E., Koçyiğit, D., Güler, A., Korkmaz, Y., Şeker, M., ... & Keser, N. (2024). Artificial intelligence-based clinical decision support systems in cardiovascular diseases. *Anatolian Journal of Cardiology*, *28*(2), 74. https://doi.org/10.14744/AnatolJCardiol.2023.3685

[19] Taylan, O., Alkabaa, A. S., Alqabbaa, H. S., Pamukçu, E., & Leiva, V. (2023). Early prediction in classification of cardiovascular diseases with machine learning, neuro-fuzzy and statistical methods. *Biology*, *12*(1), 117. https://doi.org/10.3390/biology12010117

[20] Biswas, N., Ali, M. M., Rahaman, M. A., Islam, M., Mia, M. R., Azam, S., ... & Moni, M. A. (2023). Machine learning-based model to predict heart disease in early stage employing different feature selection techniques. *BioMed Research International*, *2023*(1), 6864343. https://doi.org/10.1155/2023/6864343

[21] Subramani, S., Varshney, N., Anand, M. V., Soudagar, M. E. M., Al-Keridis, L. A., Upadhyay, T. K., ... & Rohini, K. (2023). Cardiovascular diseases prediction by machine learning incorporation with deep learning. *Frontiers in Medicine*, *10*, 1150933. https://doi.org/10.3389/fmed.2023.1150933

[22] Revathi, T. K., Balasubramaniam, S., Sureshkumar, V., & Dhanasekaran, S. (2024). An improved long short-term memory algorithm for cardiovascular disease prediction. *Diagnostics*, *14*(3), 239. https://doi.org/10.3390/diagnostics14030239

[23] Singh, M. S., Thongam, K., Choudhary, P., & Bhagat, P. K. (2024). An integrated machine learning approach for congestive

heart failure prediction. *Diagnostics*, *14*(7), 736. https://doi.org/10.3390/diagnostics14070736

[24] Almansouri, N. E., Awe, M., Rajavelu, S., Jahnavi, K., Shastry, R., Hasan, A., ... & AlAbbasi, R. K. (2024). Early diagnosis of cardiovascular diseases in the era of artificial intelligence: An in-depth review. *Cureus*, *16*(3). https://doi.org/10.7759/cureus.55869

[25] Takale, D. G., Mahalle, P. N., Sakhare, S. R., Gawali, P. P., Deshmukh, G., Khan, V., ... & Maral, V. B. (2023). Analysis of clinical decision support system in healthcare industry using machine learning approach. In *International conference on ICT for sustainable development*, *765*, 571-587. https://doi.org/10.1007/978-981-99-5652-4_51

[26] Nandy, S., Adhikari, M., Balasubramanian, V., Menon, V. G., Li, X., & Zakarya, M. (2023). An intelligent heart disease prediction system based on swarm-artificial neural network. *Neural Computing and Applications*, *35*(20), 14723-14737. https://doi.org/10.1007/s00521-021-06124-1

[27] Rana, M. S., & Shuford, J. (2024). AI in healthcare: Transforming patient care through predictive analytics and decision support systems. *Journal of Artificial Intelligence General Science*, *1*(1). https://doi.org/10.60087/jaigs.v1i1.30

[28] Luque, A., Carrasco, A., Martín, A., & de Las Heras, A. (2019). The impact of class imbalance in classification performance metrics based on the binary confusion matrix. *Pattern Recognition*, *91*, 216-231. https://doi.org/10.1016/j.patcog.2019.02.023

[29] Hasan, M. N., Hossain, M. A., & Rahman, M. A. (2025). An ensemble based lightweight deep learning model for the prediction of cardiovascular diseases from electrocardiogram images. *Engineering Applications of Artificial Intelligence*, *141*, 109782. https://doi.org/10.1016/j.engappai.2024.109782

[30] Cao, K., Liu, C., Yang, S., Zhang, Y., Li, L., Jung, H., & Zhang, S. (2025). Prediction of cardiovascular disease based on multiple feature selection and improved PSO-XGBoost model. *Scientific Reports*, *15*(1), 12406. https://doi.org/10.1038/s41598-025-96520-7

[31] Hageman, S. H., Huang, Z., Lee, H., Kaptoge, S., Dorresteijn, J. A., Pennells, L., ... & Asian-Pacific Society of Cardiology (APSC) Kim Hyo-Soo Tan Jack Komuro Issei Almahmeed Wael Yeo Khung Keong Ako Junya Park Kyung Woo. (2025). Risk prediction of cardiovascular disease in the Asia-Pacific region: The SCORE2 Asia-Pacific model. *European Heart Journal*, *46*(8), 702-715. https://doi.org/10.1093/eurheartj/ehae609

[32] Liu, T., Krentz, A., Lu, L., & Curcin, V. (2025). Machine learning based prediction models for cardiovascular disease risk using electronic health records data: Systematic review and meta-analysis. *European Heart Journal-Digital Health*, *6*(1), 7-22. https://doi.org/10.1093/ehjdh/ztae080

[33] Syed, M. G., Trucco, E., Mookiah, M. R., Lang, C. C., McCrimmon, R. J., Palmer, C. N., ... & Mordi, I. R. (2025). Deep-learning prediction of cardiovascular outcomes from routine retinal images in individuals with type 2 diabetes. *Cardiovascular Diabetology*, *24*(1), 3. https://doi.org/10.1186/s12933-024-02564-w

[34] Bandyopadhyay, S., Samanta, A., Sarma, M., & Samanta, D. (2025). Novel framework of significant risk factor identification and cardiovascular disease prediction. *Expert Systems with Applications*, *263*, 125678. https://doi.org/10.1016/j.eswa.2024.125678

[35] World Health Organization. (2024). *Cardiovascular diseases*. Retrieved from https://www.who.int/health-topics/cardiovascular-diseases

[36] World Health Organization. (2024). *Cardiovascular diseases (CVDs)*. Retrieved from https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

[37] Princy, R. J. P., Parthasarathy, S., Jose, P. S. H., Lakshminarayanan, A. R., & Jeganathan, S. (2020). Prediction of cardiac disease using supervised machine learning algorithms. In *2020 4th international conference on intelligent computing and control systems*, 570-575. https://doi.org/10.1109/ICICCS48265.2020.9121169

[38] Srivenkatesh, M. (2020). Prediction of cardiovascular disease using machine learning algorithms. *International Journal of Engineering and Advanced Technology*, *9*(3), 2404-2414. https://doi.org/10.35940/ijeat.B3986.029320

[39] Anchala, R., Pant, H., Prabhakaran, D., & Franco, O. H. (2012). 'Decision support system (DSS) for prevention of cardiovascular disease (CVD) among hypertensive (HTN) patients in Andhra Pradesh, India'—A cluster randomised community intervention trial. *BMC Public Health*, *12*(1), 393. https://doi.org/10.1186/1471-2458-12-393

[40] Gad, I., & Hosahalli, D. (2022). A comparative study of prediction and classification models on NCDC weather data. *International Journal of Computers and Applications*, *44*(5), 414-425. https://doi.org/10.1080/1206212X.2020.1766769

How to Cite: Nawaz, U., Anees-ur-Rahaman, M., Ubaidullah, H. M., Nawaz, C. M. A., & Saeed, Z. (2025). Development of a Hyperparameter-Optimized Decision Support System for Cardiovascular Disease Prediction. *Artificial Intelligence and Applications*. https://doi.org/10.47852/bonviewAIA52025361