RESEARCH ARTICLE

INTERNATIONAL RESEARCH JOURNAL OF
MULTIDISCIPLINARY TECHNOVATION

Asian Research Association

DOI: 10.54392/irjmt2647

A Modified Genetic Algorithm–Based Feature Optimization Framework for Cardiovascular Disease Risk Prediction

Shamal Salunkhe [a], Smita Bharne [b,*], Poonam Patil [c], Pragati Akre [b], Kairavi Patra [a], Anmol Thakur [a], Anya Thakur [a]

[a] Ramrao Adik Institute of Technology, D.Y. Patil Deemed to be University, Navi Mumbai, India.
[b] Department of Engineering and Technology, Bharati Vidyapeeth Deemed to be University, Navi Mumbai, India.
[c] Department of Instrumentation Engineering, Bharati Vidyapeeth College of Engineering, Navi Mumbai.
* Corresponding Author Email: smita146@gmail.com
DOI: https://doi.org/10.54392/irjmt2647
Received: 19-01-2026; Revised: 18-06-2026; Accepted: 23-06-2026; Published: 07-07-2026

Abstract: The proper diagnosis of cardiovascular disease (CVD) risk rises as a significant issue in clinical practice because the risk is directly associated with patient survival. Thus, correct and timely assessment of the risk is crucial. This work suggests a CVD probability prediction model, which combines ensemble learning and Deep Learning (DL) algorithms with a feature choice approach supported by a Modified Genetic Algorithm (MGA). A balanced clinical dataset consisting of 1,025 patient records and 14 medically relevant attributes was used which was retrieved via a publicly available repository with benchmarks dataset. The comparison of a broad range of predictive models, including classical Machine Learning (ML) algorithms like Logistic Regression, Support Vector Machines (SVM), K-Nearest Neighbors, Naïve Bayes, Decision Trees, Random Forests, XGBoost, and LightGBM, as well as DL models, including Convolutional Neural Networks (CNN), Long Short-Term Memory networks (LSTM), and Multi-layer Perceptrons (MLP), was made. The MGA based feature collection is used to select the most informative clinical attributes, decrease feature redundancy, and increase discriminative ability. The outcomes of experimental findings prove that the models with the assistance of MGA are most effective as they are more predictive and stable across various classifiers. The results demonstrate the imperative need of intelligent feature selection to enhance model generalization and predictability, which reinforces predictive modelling framework with feature optimization against cardiovascular risk prediction. The paper presents the framework that mediates between optimization methods and predictive modeling to provide valuable information on the next generation of data-driven cardiovascular diagnostics.

Keywords: Feature Selection, Cardiovascular Disease Prediction, Deep Learning, Machine Learning, Modified Genetic Algorithm.

1. Introduction

Cardiovascular disease is a cardinal public health trouble in the world and therefore it is a significant cause of expiry and morbidity in the world. Early diagnosis and proper diagnosis are necessary. Conventional diagnostic systems tend to be inefficient and have a low predictive accuracy due to redundant or inappropriate characteristics that are present in clinical data [1]. Premature and precise forecasting of heart disease not only offers preventive opportunities but also maximizes the use of healthcare supplies, therefore, minimizing the load on healthcare institutions [2]. Recent developments in ML and DL have been indicating possibility of enhancing predictive ability of heart disease diagnosis models. These algorithms apply advanced algorithms to identify patterns and associations among large dimensions of data and are more accurate than conventional methods [3, 4]. However, clinical data is of large dimensionality, which makes it difficult to apply effective feature selection methods to sort out noise and inoperable data. The features that optimize model performance are those that are selected to determine features that are most discriminatory. Chi-square, Multiple Information (MI), and Analysis of Variance (ANOVA) are methods applied to improve the accuracy of the ML models in predicting heart diseases [5-7]. Moreover, evolutionary algorithms such as Particle Swarm, Genetic Algorithms (GA) and Hybrid algorithms have been applied to improve process of feature selection and have directed to the creation of more perfect predictive models [8, 9]. The existing cardiovascular disease prediction models have high-dimensional data redundancy, low generalizability, and poor interpretability. Current feature selection algorithms do not consider the inter-feature dependency and do not dynamically adapt to the varying characteristics of data [10]. In addition, the conventional techniques like grid search and random search of hyperparameter optimization, are not efficient in computations.

The main contributions of this study are:

- This work proposes MGA based feature selection framework to advance the predictive accuracy of CVD risk assessment using a range of ML and DL models.
- MGA includes mean-based crossover, adaptive mutations is proposed to make advances with feature selection effectiveness.
- The results display that optimized feature selection enhances both prediction performance and clinical interpretability, supporting the development of reliable and scalable decision support systems for initial cardiovascular disease diagnosis.

2. Literature Survey

Precise and early estimate of CVD is an urgent priority in medical data science. With cardiovascular disease continuing to be a top cause of worldwide mortality, scientists have increasingly utilized Artificial Intelligence (AI) and ML to get strong prediction that can enable early diagnosis and treatment. Much of the recent literature investigates feature selection, model tuning, and classification methods to enhance predictive power and transferability.

The study [7] illustrates the implementation of ten various ML algorithms, namely base classifiers and ensemble, to compare CVD prediction on public and private datasets. This research emphasizes necessity of preprocessing steps like data cleaning, normalization, and balancing by employing the Synthetic Minority Oversampling Technique (SMOTE). In order to select the most impactful predictors, three feature selection strategies were employed: chi-square, analysis of variance (ANOVA), and mutual information (MI). Each of them was used to create three various subsets (SF-1, SF-2, SF-3) using which classifiers were trained. Among them, XG Boost with SF-2 feature subset trained with SMOTE was the best with the highest accuracy (97.57), F1-score (92.68), and AUC (98). SHAP values also addressed explainability as clinicians could know the impact of parameters on model predictions.

Study in [8] reports an alternative point of view with greater interest in the use of GA as a metaheuristic method to optimize the hyper parameters of ML classifier, namely SVM. The paper identifies the weaknesses of the popular optimization algorithms like grid search and random search and the hypothesis that the algorithm is either too exhaustive or too random to work in large search spaces. GA outperform this through an increase in the number of solutions that are an outcome of selection, crossover, and mutation functions that ultimately produces an optimal set of parameters. The dataset used in the work consists of 1000 records comprising of 14 attributes and it uses GAs to optimize important SVM parameters such as cost (C) and gamma. The resultant GA-optimized SVM classifier reached a 97 percent accuracy rate, which showed that evolutionary methods can be used to very large extent in improving the performance of the classifier in CVD diagnosis.

The article [11] highlights the significance of early detection of CVD to enable clinicians to diagnose and make decisions on time. The authors obtained good prediction accuracy with 1D-Convolutional Neural Networks (1D-CNNs), with 99.95% accuracy when using a train-test split, 99.12% accuracy with 5-Fold cross-validation and 98.53% accuracy with 10-Fold cross-validation. This performance is superior to the models available with the same data. Deep learning solutions have a potential of decreasing dependency on expensive clinical tests, hence saving on the expenses of patients and the health care provider.

The hybrid explainable with machine learning model (HXAI-ML) and is suggested to be used to detect CVD at initial stage and with interpretability, and it is important to note that it tackles the main issues, such as, the lack of interpretability that reduces the generalizability of the machine learning model to clinical practice [9]. HXAI-ML model has a potential of practical application into clinical decision support systems for early with average accuracy of 96%.

In [12], feature selection that integrates PSO and ML classifiers is effective in enhancing the accuracy of predicting CVD. The best performing combination is SVM with PSO, which makes it suitable in decision-making and early intervention in the clinical setting. SVM was best suited to nonlinear and high dimensional data but, retrospective data use can impair generalization. When applied to CVD prediction, the choice of the relevant features and the choice of a suitable machine learning ML or DL model is an important concern because it can greatly determine the quality and the effectiveness of the diagnostic systems. The data set of the heart diseases with various patient features is used. It is a starting point to create such predictive models. Nevertheless, not every feature is useful in predicting model performance; irrelevant or redundant features may also decrease predictive and model interpretability. In order to address this issue a MGA for feature selection is presented.

3. Proposed Methodology

MGA is a technique that utilizes evolutionary principles to select a best feature subset, thereby improving the working of the predictive model. Figure 1 indicates the architecture of CVD prediction. This diagram illustrates methodology to developing a CVD prediction model with ML and DL techniques, optimized through GA-based feature selection. Figure 2 shows the system flow for CVD prediction.

3.1 Data Collection

The method begins with acquiring a structured dataset containing patient histories with attributes relevant to cardiovascular health. The Heart Disease Dataset is used from Kaggle. This dataset contained 1025 records and 14 features, with the target being the presence or absence of heart disease as shown in Table 1. The features included categorical and continuous variables: age, sex, type of chest pain, blood pressure, level of cholesterol, which are of crucial importance for predicting heart disease. This dataset has broad usage in the medical field and ML research, hence making it the best fit for improvement of CVD prediction models [12, 13].

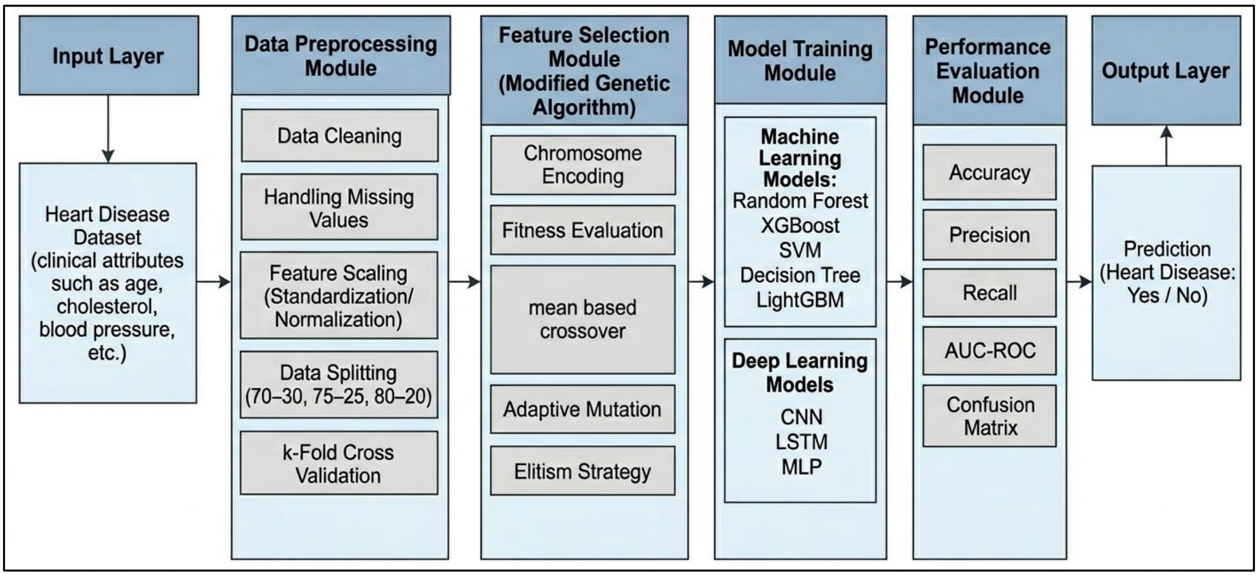


Figure 1. Architecture of CVD prediction.

Table 1. Dataset attributes

S.No	Dataset Column Name	Description	Value Type
1	Age	Age of the individual in years	Numerical
2	Sex	Gender (0 = Female, 1 = Male)	Categorical
3	Cp	Chest pain type (0–3: Typical, Atypical, Non-anginal, Asymptomatic)	Categorical
4	Trestbps	Resting blood pressure (in mm Hg)	Numerical
5	Chol	Serum cholesterol in mg/dl	Numerical
6	Fbs	Fasting blood sugar > 120 mg/dl (1 = True, 0 = False)	Binary
7	Restecg	Resting ECG results (0 = Normal, 1 = ST-T abnormality, 2 = LV hypertrophy)	Categorical
8	Thalach	Maximum heart rate achieved during the test	Numerical
9	Exang	Exercise-induced angina (1 = Yes, 0 = No)	Binary
10	Oldpeak	ST depression induced by exercise relative to rest	Numerical
11	Slope	Slope of the ST segment (0 = Upsloping, 1 = Flat, 2 = Down sloping)	Categorical
12	Ca	Number of major vessels (0–3) coloured by fluoroscopy	Numerical
13	Thal	Thalassemia (0 = Normal, 1 = Fixed defect, 2 = Reversible defect)	Categorical
14	Target	Presence of heart disease (1 = Disease, 0 = No disease)	Binary

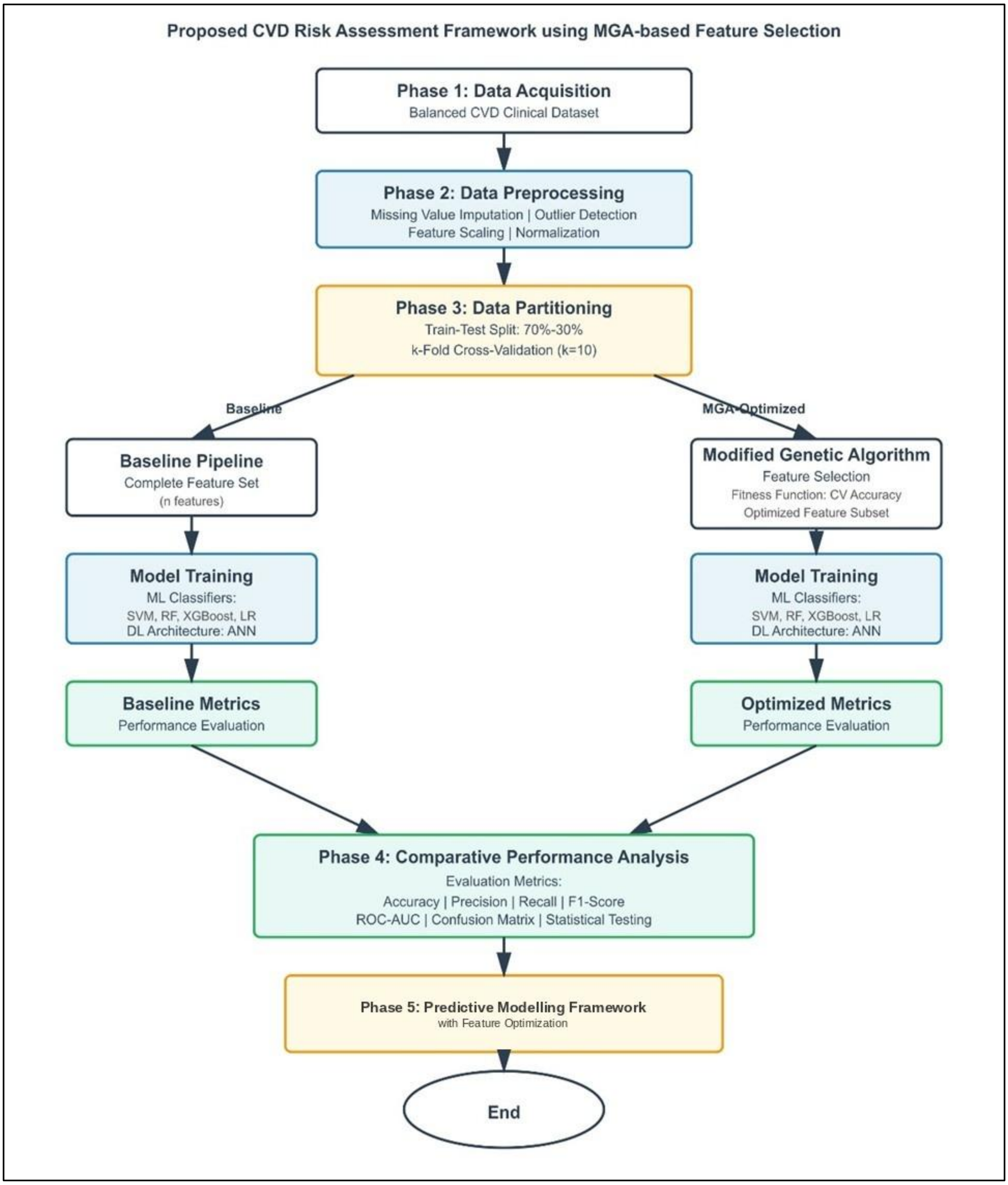


Figure 2. System flow for CVD prediction.

3.2 Data Preprocessing

Data preprocessing was carried out to prepare the dataset for uses by cleaning, transformation, and formatting. Data cleaning handles vanished values, removes duplicates, and corrects inconsistencies. It sought to make sure that the information employed on the modelling was consistent, relevant, and prepared for analysis. This was done in various operations like data cleaning. Balancing datasets, feature scaling and feature scaling [14]. The first stage in preprocessing was data cleaning. In this step, we concentrated on working with missing or inconsistent data. Data was analyzed in a great deal and it was identified to be devoid of any missing values thereby ensuring that the trained models are correct [15]. In the dataset, the target attribute (heart disease present or absent) consisted of 526 samples of class 1 (disease present) and 499 samples of class 0 (disease absent), indicating that the data was balanced. This was significant since unbalanced datasets have the potential to result in skewed predictions, where the model tends to favor the majority class [16]. Figure 3 showing, the percentage distribution of the target classes of dataset.

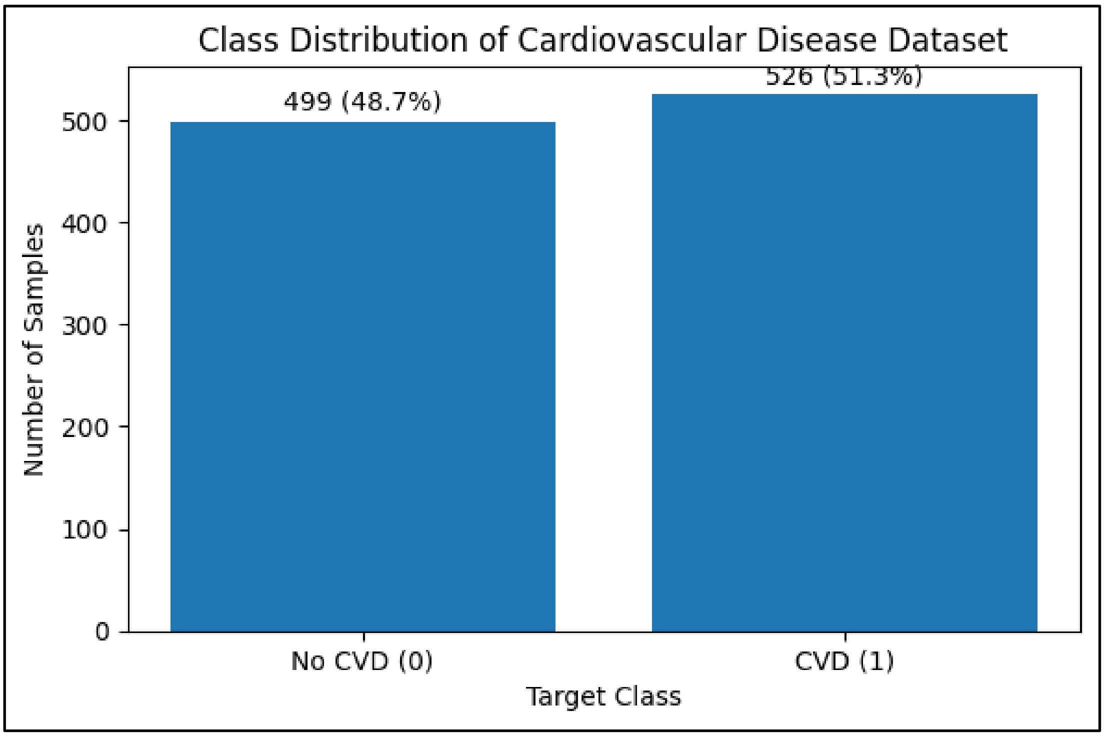


Figure 3. Cardiovascular disease class distribution.

Next, feature scaling is performed, which Normalization and standardization are two popular methods of feature scaling. Normalization scales data into an assortment between 0 and 1, according to the formula shown in Equation (1).

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

Where, X is the original value, X_{min} is the minutest value, and X_{max} is the extreme value in the feature. It is helpful when models are based on distances, like in K-Nearest Neighbors (KNN) [13]. Conversely, standardization rescales the data to have a mean of 0 and a standard deviation of 1, which suits models where data is assumed to follow a Gaussian distribution, such as LS and SVM. The formula is given in Equation (2).

$$X_{standard} = \frac{X - \mu}{\sigma} \quad (2)$$

Where, X is the original value, μ is the mean, and σ is the standard deviation of the feature. Then, the dataset is split to training and test sets based on several strategies. First, the dataset was split using three different strategies: 70% training and 30% testing, 75% training and 25% testing, and 80% training and 20% testing. This allowed to test how well the models performed with different amounts of training data. In addition to these divisions, k-fold cross-validation used with k=5, where data were split into five equal-sized sections (or folds), and the model trained on four sections and tested on the other section [17]. We did this five times, with each segment being the test set once, to give a less variable measure of model performance.

3.3 Feature Selection using MGA

The choice of features is a significant process in ML because it is responsible for choosing the most significant variables from a database and hence enhancing the accuracy of the model without increasing its complexity. MGA to perform feature selection in this research to enhance the performance of CVD with various ML and DL models. Feature selection is a significant process in ML as it helps to choose the best features from a data set by removing dimensionality and enhancing the performance of a model. Evolutionary methods like GA are frequently utilized for feature selection since they are stable in navigating large and complex search spaces [18, 19]. The MGA adapted accomplishes this by maximizing feature selection which most leads to accurate predictions. Detailed phases involved in MGA are discussed below.

Step 1: Initialization

The initial population $P(0)$ with N chromosomes is indicated in Equation (3),

$$P = \{C_1, C_2, \dots, C_n\} \quad (3)$$

Where each chromosome is presented in next Equation (4),

$$C_i \in \{0,1\}^d, \quad C_i = [c_{i1}, c_{i2}, \dots, c_{id}] \quad (4)$$

Each chromosome C_i is a binary vector of length d (number of features) is given in Equation (5).

$$C_i \in \{0,1\}^d \quad (5)$$

Let $G_{\max}$ be the maximum number of generations, p_m the mutation probability, and ε the convergence threshold.

The fitness function $f(C_i)$ evaluates the classification accuracy (from an ML model) on features selected by chromosome C_i and shown in Equation (6).

$$f(C_i) = \text{Accuracy}_{ML}\big(\text{SelectedFeatures}(C_i)\big) \quad (6)$$

The selection probability $P(C_i)$ of each chromosome based on its fitness. It is shown in Equation (7).

$$P(C_i) = \frac{f(C_i)}{\sum_{j=1}^{N} f(C_j)} \quad (7)$$

Step 2: Fitness Evaluation

For each chromosome C_i, define the selected feature set as shown in Equation (8).

$$S_i = \{j \mid c_{ij} = 1\} \quad (8)$$

Then Compute the fitness using stratified k-fold cross-validation with Equation (9)

$$F(C_i) = \frac{1}{k}\sum_{l=1}^{k} A_i^{(l)} \quad (9)$$

Step 3: Evolutionary Loop (for $g = 1$ to $G_{\max}$)

a. Selection

Select the top m chromosomes with highest fitness as given in Equation (10)

$$S = \{P_1, P_2, \dots, P_m\}, \quad F(P_1) \geq F(P_2) \geq \cdots \geq F(P_m) \quad (10)$$

b. Mean-Based Crossover

- Compute offspring as given in Equation (11),

$$O_j = \frac{1}{m}\sum_{i=1}^{m} p_{ij}, \quad \text{for } j = 1,2,\dots,d \quad (11)$$

- Binarize each gene as shown in Equation (12)

$$o_j = \begin{cases} 1, & \text{if } O_j \geq 0.5 \\ 0, & \text{otherwise} \end{cases} \quad (12)$$

c. Mutation

Apply mutation to each gene o_j as shown in Equation (13),

$$o'_j = \begin{cases} 1 - o_j, & \text{with probability } p_m \\ o_j, & \text{otherwise} \end{cases} \quad (13)$$

Resulting in mutated offspring is shown in Equation (14),

$$O' = [o'_1, o'_2, \dots, o'_d] \quad (14)$$

d. Fitness Evaluation of Offspring is indicated in Equation (15)

$$F(O') = \frac{1}{k}\sum_{l=1}^{k} A_{O'}^{(l)} \quad (15)$$

e. Replacement

The worst chromosome is detected as per Equation (16)

$$C_w = \underset{C_i}{\operatorname{argmin}} F(C_i) \quad (16)$$

If,

$$F(O') > F(C_w) \Rightarrow P := P \setminus \{C_w\} \cup \{O'\} \quad (17)$$

f. Convergence Check

Compute convergence as shown in Equation (18),

$$F_{\min} = \min_{C_i} F(C_i), \quad F_{\max} = \max_{C_i} F(C_i) \quad (18)$$

$$\text{If, } F_{\max} - F_{\min} < \varepsilon \Rightarrow \text{Stop} \quad (19)$$

Step 4: Final Output

Select the best chromosome,

$$C^* = \underset{C_i}{\operatorname{argmax}} F(C_i) \quad (20)$$

The chromosome C^* represents the optimal subset of features. Preserving the best-performing chromosome $\hat{C}$ across generations is called Elitism, shown in Equation (21).

$$\hat{C} = \underset{C_i}{\operatorname{argmax}} f(C_i) \quad (21)$$

Ensure $\hat{C}$ is included in the next generation.

Convergence Criteria

Algorithm terminate as per Equation (22)

$$\left|\max f\left(C_i^{(g+1)}\right) - \max f\left(C_i^{(g)}\right)\right| < \varepsilon \quad \text{or} \quad g = G_{\max} \quad (22)$$

The MGA process, from population initialization and fitness evaluation to crossover, mutation, elitism, and convergence for feature selection. We start by initializing the feature subset as a chromosome. Every chromosome is a binary value where each entry leads to a feature of the dataset. The entry value of 1 signifies that the respective feature belongs to the subset, whereas the value 0 signifies exclusion of the feature. At start, the randomly generated population contains several chromosomes, each one corresponding to a potential feature subset.

The core of the MGA is the fitness function, which estimates the quality of each feature subset by training a ML model and its performance. We employed a Stratified K-Fold Cross-Validation (CV) method to guess the performance of every chromosome. The data is split into 5 folds such that each fold has an equal distribution of classes. The training set is trained on and validated on the model for each fold. The fitness value of a chromosome is taken as the average accuracy over all the folds in the cross-validation. The aim of the algorithm is to maximize this fitness value, i.e., to identify the feature subset that provides the maximum classification accuracy. Mathematically, the fitness function is expressed in Equation (23).

$$fitness = \frac{1}{k}\sum_{i=1}^{k} accuracy_i \quad (23)$$

Where k is the number of folds, and $accuracy_i$ is the accuracy obtained from i^{th} fold. After the fitness of all chromosomes is computed, the second step is to produce offspring by crossover. In the proposed MGA, a simple approach is used where parents are chosen on the basis of their fitness. The first 4 chromosomes are selected as parents. These parents contribute to the creation of new chromosomes (offspring) that may inherit their features. Instead of using traditional two-parent crossover, a mean-based crossover is used, where the offspring is generated by averaging the feature selection patterns of the parents. The offspring is calculated as the mean of the selected features from the parent chromosomes as shown in Equation (11). This approach guarantees that the offspring will inherit traits from both parents with the ability to generate a diverse but good-performing feature subset. Following the creation of the offspring, we add mutation to keep genetic diversity and avoid the algorithm getting stuck in local optima. During the mutation step, every gene (feature) in the offspring has a possibility to flip its value with a probability that is set by the mutation rate. This step makes sure that the search process keeps exploring new sets of features. For maintaining the optimal, an elitism is employed, such that the best individuals from the existing generation are carried to the next generation without any changes. This way, the optimal of feature sets do not get eliminated in the process of evolution. In this model, when generating offspring, offspring's fitness compares with the least fit one in the present population. If the new offspring is better, it replaces the least fit one. This guarantees that the population develops so that it never loses or deviates from, but rather keeps or enhances, the best so far solution found. The algorithm stops when the fitness score difference between the best and worst chromosomes is less than a specified threshold, which means that convergence has been reached. This criterion guarantees that the search for an optimal solution has arrived at a stage where additional iterations are not likely to produce much better results.

3.3.1 MGA Configuration with Parameters

The proposed MGA is implemented after careful selection of parameters to balance the efficiency and convergence. The size of the population set size is 50 with max 100 generations for search depth. The selection of parents was performed by tournament selection with the size of 3 for high fitness of individual. The crossover probability of 0.8 with mean-based method is used for feature subset combination. After that the mutation is applied with mutation rate of 0.01 and 0.05 to maintain the diversity. To maintain the best performing solutions, an elitism is used by preserving the top 5% feature values across the generations. Convergence was determined based on the fitness threshold of 0.001 over 10 continuous generations. The selection of final feature subset ensures the stability of selected feature subset along with frequently occurring features.

3.3.2 Deep Learning Model Configuration

To ensure a fair and reproducible comparison, the architectures and training configurations of the DL models, CNN, LSTM, and MLP were explicitly defined and consistently applied across all experiments.

Input Representation: Since the dataset is tabular with 14 features, the input was pre-processed and normalized before being fed into deep learning models. For CNN, the feature vector was reshaped into a 1D format (1 × 14) to enable convolution operations. For LSTM, the same feature vector was treated as a sequence of length 14 with one feature per timestep. The CNN model consists of two 1D convolutional layers followed by a fully connected network. The first convolutional layer uses 32 filters with a kernel size of 2 and ReLU activation with a learning rate of 0.001, followed by a max-pooling layer. The second convolutional layer uses 64 filters with ReLU activation, followed by another pooling layer. The extracted feature maps are flattened and passed to a dense layer of 64 neurons with ReLU activation. A dropout layer with a rate of 0.3 is applied to reduce overfitting. The final output layer uses a sigmoid activation function for binary classification. The LSTM model is designed to capture sequential dependencies in the input representation. LSTM model consist of 50 hidden units used followed by a dropout layer (rate = 0.2) to prevent overfitting. The output is then passed to a dense layer with 32 neurons (ReLU activation), followed by a sigmoid output layer. The MLP model consists of two fully connected hidden layers with 64 and 32 neurons respectively, both using ReLU activation. A dropout layer (rate = 0.2) is applied between layers to improve generalization. The final layer uses a sigmoid activation function. All models were trained using the Adam optimizer with a learning rate of 0.001 and binary cross-entropy as the loss function. Early stopping was employed with a patience of 5 epochs based on validation loss to prevent overfitting. Model performance was evaluated using the same train-test splits and cross-validation strategies as the machine learning models to ensure consistency.

3.4 Comparison with the Standard GA

The proposed MGA typically distinguishes from the standard GA in several ways. In the standard GA, crossover is performed with only two parents chromosomes with the uniform or single crossover unlike the proposed MGA uses the mean-based crossover where the offsprings are produced by combining the feature selection pattern across the multiple parents. This will make the more knowledgeable search space and decreasing the probability of the receiving suboptimal features.

In addition to this the mutation process is modified to maintain the multiplicity to mitigate the early convergence which is usually occur in the standard GA. The combination of the best execution feature subset across the generations with fitness variance convergence criteria permits the early stopping. The enhanced efficiency of the algorithm to recognize the distinguish the feature subset while maintaining the computational efficiency makes algorithm robust within various classifiers.

3.5 Validation Protocol and Leakage Prevention

For unbiased evaluation, feature selection was firmly performed within each taring fold with cross validation. The dataset is split into the training and validation subset with MGA was applied on train data to identify the optimal feature subset. This feature subset is used to train the predictive model and evaluated with consistent validation fold. This fold wise selection techniques prevent the information leakage and ensure that no data from test and validation set effects the feature selection process. The same strategy is also applied through all evaluation including k-fold cross validation. By dividing the feature selection within train data, the performance metric ensures the model generalization capability.

3.6 Model Training

In the initial step of this approach is comparing, a number of popular ML and DL models, including Random Forest, XGBoost, Logistic Regression, SVM, K-Nearest Neighbors (KNN), Naïve Bayes, Decision Tree, LightGBM, CNN, LSTM, and MLP [20-23]. These models were validated under typical train-test protocols, employing different split ratios (70–30, 75–25, and 80–20) as well as k-fold cross-validation to evaluate their generalization and robustness without feature selection. Subsequently, the MGA was used to conduct feature selection for all of the models. This algorithm iterated over a population of potential feature subsets, evaluated their performance against predictive precision, and selected top-performing sets of features upon which to train models.

The final aim of this process was to find out how the feature selection process could enhance the performance of both the ML and DL models by extracting the most useful features and eliminating superfluous complexity. In this approach, we aimed at demonstrating how the Modified GA when used in the context of feature selection can enhance predictive performance of existing ML and DL models to heart disease diagnosis to generate more accurate, yet easier to interpret and more efficient, models. According to the chosen model, the system is predictive of a patient being at risk of CVD (YES) and not being at risk (NO) [24, 25].

4. Results and Discussion

This section provides experimental comparisons of ML and DL classifiers on a heart disease dataset with and without feature selection by the proposed MGA. The MGA was used for feature selection to find the most meaningful features in order to achieve better model accuracy, lower computational complexity, and higher efficiency [26, 27]. Classifier performance was evaluated on various train-test splits (70–30, 75–25, 80–20) and 5-fold cross-validation fold cross-validation and mostly on accuracy values. The findings prove that the MGA is significantly select useful features, greatly improving tree-based models like Random Forest, LightGBM, and Decision Tree accuracy as well as efficiency, whereas DL models had mixed results towards feature reduction. All the results are based on leakage free evaluation for which the selection of feature is performed within each training fold. This makes the performance metrics are not biased with any unseen data and precisely reflecting the penalization.

4.1 Experimental Setup Details

To ensure the fair comparison across the ML models all the experiments are conducted using a constant evaluation framework using the Scikit-learn. The hyperparameters are selected based on experiential tuning. Random Forest was configured with 100 estimators with a maximum depth of 10. XGBoost used with a learning rate of 0.1 and with 100 boosting rounds. With a regularization strength of 1.0 SVM is implemented with RBF kernel and KNN is used with k-5 value. All the DL models were implemented with TensorFlow framework. They are trained with a 50 epoch, batch size of 32 based on validation loss. A fixed random seed (42) is used across all experiments. The experiments are executed on Intel i7 processor with 32 GB Ram, GPU acceleration wherever needed.

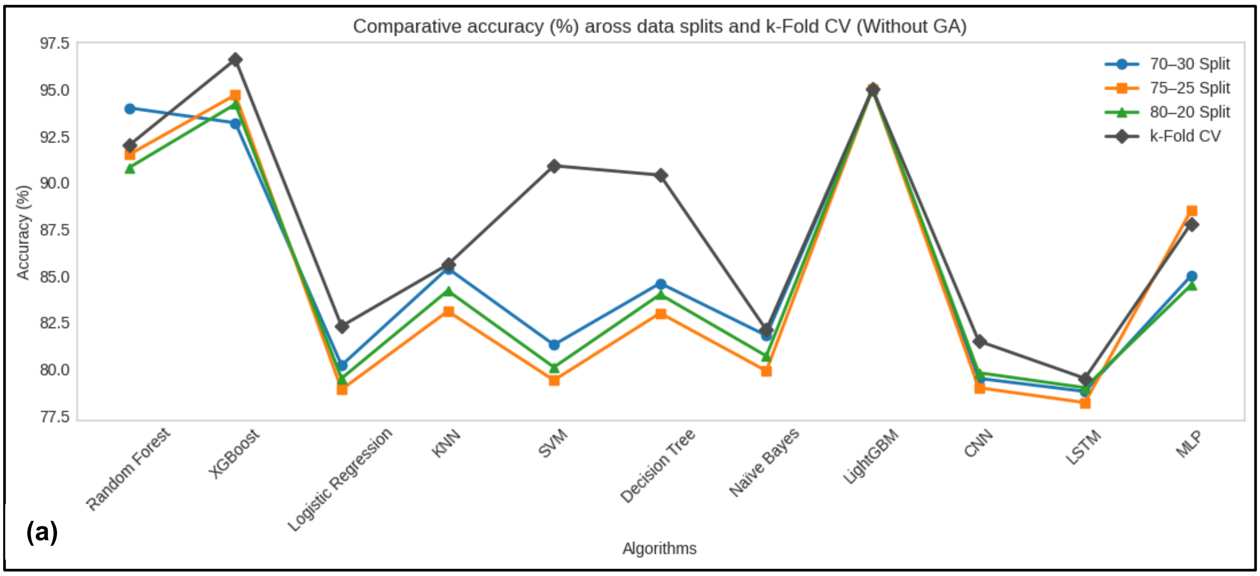
4.2 Classification Accuracy

The baseline model classification accuracy without the use of MGA for feature selection is measured based on various train–test splits: 70–30, 75–25, 80–20, and 5-fold cross-validation, as shown in Figure 4 (a).

Among the models under evaluation, the best performing classifiers were Random Forest, LightGBM, and XGBoost that produced accuracies of 94.00%, 95.00%, and 93.20%, respectively, for the 70–30 split. These classifiers performed well on other splits also. Traditional classifiers like Logistic Regression, Naïve Bayes, and SVM performed moderately with accuracy values varying from 80% to 90%, while deep learning classifiers like CNN, LSTM, and MLP are also performing well. The classifiers' predictive accuracy was enhanced after being trained on the Modified GA selected respect to the train–test split ratios of 70–30, 75–25, 80–20 and 5-fold cross-validation.

As shown in Figure 4 (b) Decision Tree, Random Forest, and LightGBM were the top-performing classifiers with high classification accuracy, thus verifying the effectiveness of the MGA in improving model performance through efficient feature selection. This finding explains the capability of the MGA to discard useless or redundant features so that classifiers could concentrate on those most informative features. However, an accuracy drop was noticed for deep models such as CNN and LSTM. This is due to the fact that the GA-based feature removal can remove the features that, although not necessary for traditional models, have significant spatial or sequential patterns which are needed for deep models [25]. Therefore, such a decrease in feature richness constrained the capacity of CNNs and LSTMs to learn richer internal representations and thus their predictive performance decreased. These results demonstrate that although MGA considerably outperforms traditional classifiers, DL models would need either a richer feature set or other choice techniques that preserve richer data relationships.

The performance of the baseline algorithms on different data splits and K-Fold Cross Validation is shown in Figure 4 (a) and (b). The classification performance of the evaluated ML and DL models was analysed before and after applying the MGA for feature selection across multiple data splits (70–30, 75–25, 80–20) and 5-fold cross-validation. Before feature optimization, ensemble and tree-based models such as Random Forest, LightGBM, and XGBoost consistently achieved higher accuracy compared to other classifiers. For instance, in the 70–30 split, these models reached accuracy levels above 93%, indicating their strong capability in handling structured clinical data.



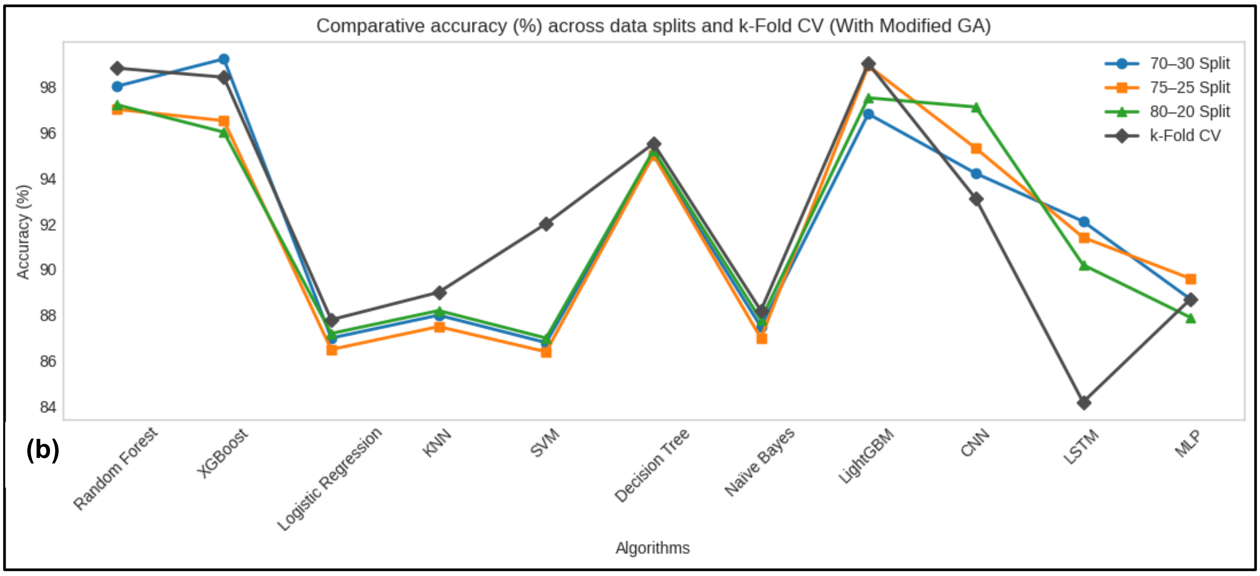


Figure 4. (a) Comparative performance of algorithms without GA, **(b)** Comparative performance of algorithms with modified GA.

Traditional models such as Logistic Regression, SVM, and Naïve Bayes demonstrated moderate performance, while DL models achieved competitive but slightly variable results depending on the data split.

After applying MGA-based feature selection, a clear improvement in classification accuracy was observed for most ML models, particularly tree-based and ensemble approaches. Decision Tree, Random Forest, and LightGBM showed significant gains, achieving near-perfect or highly consistent accuracy across different validation strategies. This improvement can be attributed to the ability of MGA to eliminate redundant and less informative features, allowing these models to focus on the most discriminative attributes and reduce overfitting. In contrast, deep learning models exhibited a decline in performance after feature selection. Models such as CNN and LSTM, which rely on learning complex feature interactions and hierarchical representations, were adversely affected by the reduced feature space. The removal of features, although beneficial for traditional models, likely limited the ability of these models to capture underlying patterns in the data. As a result, their classification accuracy decreased across most evaluation settings. Overall, the results indicate that feature optimization using MGA is highly effective for traditional and ensemble learning models operating on tabular clinical datasets, but may not be suitable for DL architectures that depend on richer input representations. This highlights the importance of aligning feature selection strategies with the characteristics of the learning model.

4.3 Performance Comparison of Models

Figure 4 (a) and (b) displays the comparison of performance of different ML and DL models with and without feature selection using GA. The models employed are Random Forest, XGBoost, Logistic Regression, KNN, SVM, Decision Tree, Naïve Bayes, LightGBM, CNN, LSTM and MLP. Each was evaluated on four data split methods: 70–30, 75–25, 80–20, and k-fold cross-validation. With respect to accuracy, tree-based Random Forest, Decision Tree, and LightGBM scored perfect or close-to-perfect accuracy when paired with MGA. XGBoost generally performed well with or without MGA. Logistic Regression, KNN, and SVM improved when paired with MGA, while deep-learning models CNN and LSTM dropped, suggesting incompatibility between features selected with MGA and their inner workings. Precision and Recall values indicated the same pattern, and Decision Tree and LightGBM reported high recall when used with MGA as shown in Table 2 and Table 3 respectively.

Table 2. Precision values for ML and DL models with and without MGA feature selection on different data splits

Algorithm	Condition	70-30	75-25	80-20	k-fold
Random Forest	without MGA	91.77	89.31	88.28	92.20
	with MGA	98.23	97.69	97.69	98.90
XGBoost	without MGA	95.23	94.69	94.69	94.90
	with MGA	98.07	98.20	98.13	98.61
Logistic Regression	without MGA	76.31	73.79	75.63	82.26
	with MGA	83.13	81.42	82.05	83.54
KNN	without MGA	84.31	80.91	80.00	85.61
	with MGA	88.51	87.30	87.61	89.98
SVM	without MGA	76.47	74.82	75.60	90.91
	with MGA	80.00	79.60	77.95	87.55
Decision Tree	without MGA	77.29	75.31	78.40	90.10
	with MGA	98.70	97.20	96.13	95.61
Naïve Bayes	without MGA	76.43	75.00	75.40	82.05
	with MGA	84.04	82.73	80.34	83.37
LightGBM	without MGA	94.66	93.66	91.58	94.00
	with MGA	98.98	97.90	96.90	98.90
CNN	without MGA	98.48	95.20	98.01	93.07
	with MGA	78.04	76.59	75.21	79.87
LSTM	without MGA	91.83	92.80	90.38	84.16
	with MGA	75.56	78.46	73.60	75.75
MLP	without MGA	87.89	84.90	81.20	83.80
	with MGA	89.00	88.21	86.98	88.73

Table 3. Recall performances of different classifiers on different splits of data, comparing with and without MGA

Algorithm	Condition	70-30	75-25	80-20	k-fold
Random Forest	without MGA	97.31	93.60	95.14	95.42
	with MGA	98.10	98.43	97.56	97.90
XGBoost	without MGA	95.97	97.60	97.08	99.42
	with MGA	96.83	98.48	99.00	98.47
Logistic Regression	without MGA	87.00	85.60	87.37	89.44
	with MGA	87.34	86.36	91.42	87.25
KNN	without MGA	86.57	84.80	89.32	85.47
	with MGA	82.91	83.34	87.61	82.31
SVM	without MGA	87.24	85.60	90.29	93.15
	with MGA	91.13	91.66	94.28	92.01
Decision Tree	without MGA	95.97	95.20	95.14	95.95
	with MGA	98.97	97.20	98.14	99.42
Naïve Bayes	without MGA	89.26	86.40	89.32	86.28
	with MGA	86.70	87.12	89.52	88.21
LightGBM	without MGA	95.30	97.60	95.14	95.40
	with MGA	97.97	96.20	97.14	98.42
CNN	without MGA	87.24	95.20	96.11	92.19
	with MGA	81.01	81.81	83.80	87.06
LSTM	without MGA	90.60	92.80	91.26	80.48
	with MGA	86.07	77.27	87.61	86.88
MLP	without MGA	97.08	97.60	97.08	97.72
	with MGA	87.34	90.15	92.38	91.25

Table 4. AUC-ROC values of MGA-based feature selection models and non-MGA-based feature selection models under various evaluation strategies

Algorithm	Condition	70-30	75-25	80-20	k-fold
Random Forest	without MGA	89.55	87.78	87.68	89.90
	with MGA	98.08	96.60	96.08	95.72
XGBoost	without MGA	89.55	87.78	87.68	99.79
	with MGA	99.61	99.95	100	99.96
Logistic Regression	without MGA	90.00	87.78	87.68	91.48
	with MGA	91.03	90.95	91.52	90.69
KNN	without MGA	88.18	86.81	87.07	95.69
	with MGA	95.26	95.19	96.45	96.43
SVM	without MGA	90.00	87.78	87.68	97.44
	with MGA	90.05	89.80	91.02	94.74
Decision Tree	without MGA	89.52	87.78	87.68	99.47
	with MGA	97.08	97.60	96.80	95.15
Naïve Bayes	without MGA	89.55	87.78	87.68	90.28
	with MGA	90.67	90.74	91.37	91.39
LightGBM	without MGA	89.55	87.78	87.68	88.90
	with MGA	97.08	96.10	96.53	95.72
CNN	without MGA	97.32	98.10	98.83	98.52
	with MGA	89.29	89.36	88.46	90.44
LSTM	without MGA	96.95	98.52	97.89	90.46
	with MGA	84.13	84.44	83.53	83.77
MLP	without MGA	99.25	99.38	99.25	99.48
	with MGA	87.33	85.47	81.69	92.13

Conventional models once more indicated improvements, whereas deep models indicated other or poorer performance with MGA, which implies that MGA can eliminate input features essential for sequential or spatial pattern learning.

AUC-ROC values comparison is indicted in Table 4, MGA significantly improved scores for the majority of conventional and ensemble models, particularly Decision Tree, Random Forest, XGBoost, and LightGBM. KNN and Naïve Bayes significantly improved as well. CNN and LSTM, however, saw the drop in AUC-ROC when combined with MGA, again indicating that MGA would not work well with deep models based on hierarchical feature extraction. Generally, MGA -based feature selection was good for most shallow models and ensemble models in the sense that it improved precision, recall, as well as AUC-ROC. Its effect on deep learning models was however not so encouraging, though, to suggest other specific feature selection algorithms for these models.

A key observation from the experimental results is the contrasting impact of MGA-based feature selection on different model families. Tree-based and ensemble methods such as Random Forest, Decision Tree, and LightGBM show significant improvement after feature optimization, whereas DL models such as CNN and LSTM exhibit noticeable performance degradation. This behaviour can be explained by the nature of tabular clinical data and the learning mechanisms of these models.

Tree-based models inherently perform implicit feature selection and are robust to irrelevant or redundant attributes. The MGA further enhances this capability by eliminating non-informative features, thereby reducing noise and improving generalization performance. As a result, these models benefit from a more compact and discriminative feature space. In contrast, DL models rely on high-dimensional feature representations to learn complex, non-linear relationships. Models such as CNN and LSTM are designed to capture spatial or sequential dependencies and often require richer input features to build meaningful internal representations. MGA removes redundant features thus improves generalization but damages latent patterns. As a result, aggressive feature reduction through MGA negatively impacts their ability to learn complex patterns, leading to the observed performance degradation (e.g., CNN precision drop from 98.48 to 78.04 Table 2 and LSTM AUC drop from 96.95 to 84.13 Table 4). Feature compression by MGA leads to reduced representation learning capacity in DL models. To provide a more comprehensive evaluation suitable for medical diagnosis, additional performance metrics were considered, including F1-score, and Matthews Correlation Coefficient (MCC). Sensitivity is particularly important in cardiovascular risk prediction as it reflects the model's ability to correctly identify patients with the disease, while specificity indicates correct identification of healthy individuals. MCC provides a balanced measure even in the presence of class imbalance.

F1 Score and MCC values of MGA-based feature selection models and non-MGA-based feature selection models under various evaluation strategies are shown in Table 5

To evaluate whether the performance improvements achieved by the proposed Modified Genetic Algorithm (MGA) are statistically significant, a paired t-test was conducted comparing F1-scores of baseline models (without MGA) and MGA-enhanced models across multiple validation strategies (70–30, 75–25, 80–20 splits, and k-fold cross-validation).

The results indicate that for ensemble and tree-based models such as Random Forest, Decision Tree, and LightGBM, the improvements in F1-score after applying MGA are statistically significant ($p < 0.05$). These models consistently show large positive differences across all validation splits, confirming that MGA contributes meaningfully to performance enhancement.

XGBoost also demonstrates statistically significant improvement, although the magnitude of gain is comparatively smaller. Traditional models such as Logistic Regression, KNN, and Naïve Bayes exhibit moderate improvements, with statistical significance observed in most cases.

In contrast, DL models, CNN, LSTM, and MLP show a consistent decline in performance after MGA-based feature selection. The paired t-test results for these models indicate that the changes are either statistically insignificant or significantly negative ($p \geq 0.05$), suggesting that feature reduction adversely affects deep learning architectures.

These findings highlight that while feature selection is beneficial for traditional machine learning models on structured data, it must be applied cautiously for deep learning architectures, where preserving feature richness is critical for optimal performance.

4.4 Confusion Matrix Analysis

To further assess the classification accuracy of the top models following feature selection by the Modified Genetic Algorithm, confusion matrices were contrasted. Figure 5 illustrates that both the Decision Tree and LightGBM classifiers obtained perfect test set classification, with all the true positive and true negative instances correctly classified. The Random Forest model also did the same, having misclassified three positive instances and registering a test accuracy of 99% in 70-30 split and an AUC-ROC value of 1.00 in 75-25, 80-20 and k-fold cv splits.

Table 5. F1 Score and MCC values of MGA-based feature selection models and non-MGA-based feature selection models under various evaluation strategies

Algorithm	F1 70-30 (W/O)	F1 70-30 (MGA)	MCC 70-30 (W/O)	MCC 70-30 (MGA)	F1 75-25 (W/O)	F1 75-25 (MGA)	MCC 75-25 (W/O)	MCC 75-25 (MGA)	F1 80-20 (W/O)	F1 80-20 (MGA)	MCC 80-20 (W/O)	MCC 80-20 (MGA)	F1 k-fold (W/O)	F1 k-fold (MGA)	MCC k-fold (W/O)	MCC k-fold (MGA)
Random Forest	94.45	98.16	0.89	0.96	91.40	98.05	0.83	0.97	91.60	97.62	0.83	0.95	93.78	98.39	0.87	0.97
XGBoost	95.60	97.44	0.91	0.95	96.12	98.34	0.93	0.97	95.87	98.56	0.92	0.97	97.11	98.54	0.94	0.97
Logistic Reg.	81.31	85.16	0.62	0.68	79.20	83.75	0.57	0.65	80.90	86.45	0.60	0.70	85.69	85.38	0.69	0.70
KNN	85.42	85.65	0.71	0.72	82.83	85.29	0.66	0.70	84.58	87.61	0.67	0.74	85.54	85.87	0.70	0.72
SVM	81.50	85.20	0.62	0.68	79.60	85.28	0.58	0.64	82.40	85.67	0.62	0.68	91.99	89.73	0.83	0.79
Decision Tree	85.60	98.83	0.71	0.98	84.34	97.20	0.67	0.96	86.70	97.13	0.70	0.96	92.96	97.49	0.86	0.97
Naïve Bayes	82.34	85.33	0.64	0.69	80.17	84.89	0.59	0.68	81.90	84.79	0.61	0.67	84.07	85.73	0.64	0.66
LightGBM	94.98	98.47	0.90	0.98	95.61	97.04	0.91	0.96	93.30	97.02	0.86	0.96	94.70	98.66	0.89	0.98
CNN	92.50	79.48	0.86	0.57	95.20	79.10	0.90	0.56	96.99	79.21	0.94	0.63	92.62	83.36	0.85	0.62
LSTM	91.21	80.29	0.83	0.57	92.80	77.86	0.86	0.55	90.81	80.06	0.81	0.56	82.21	81.19	0.64	0.53
MLP	92.26	88.14	0.87	0.75	91.05	89.17	0.86	0.71	88.90	89.61	0.82	0.65	90.52	89.96	0.84	0.65

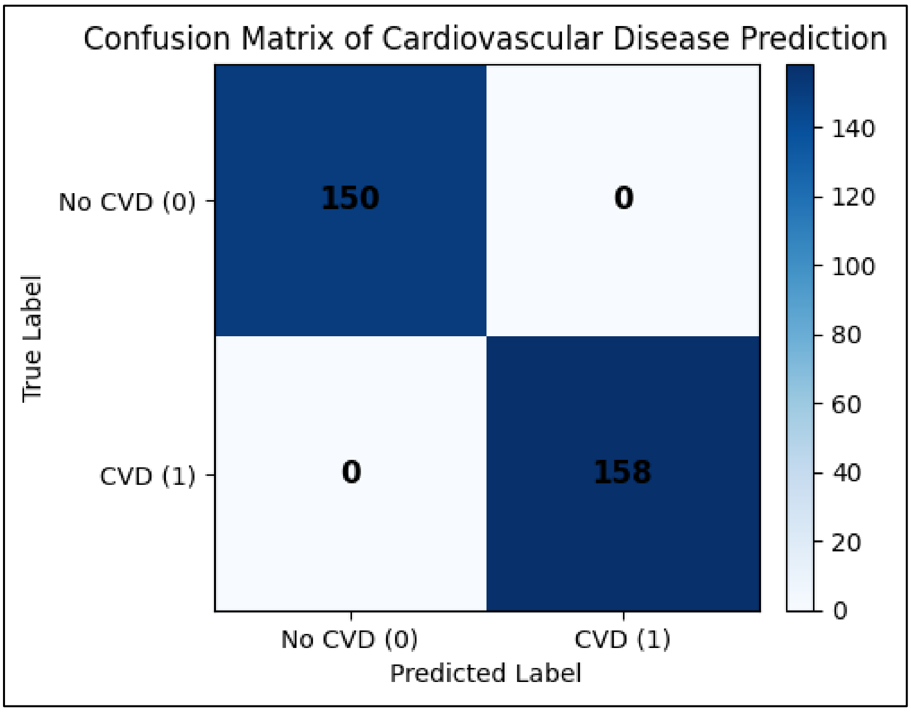


Figure 5. Confusion matrix of top performing classifiers.

These results guarantee the stability of our extracted features and attest to the efficacy of the MGA in optimizing the interpretability of the model alongside performance is indicated in Figure 5. This research developed MGA feature selection in an attempt to increase CVD prediction accuracy and efficiency using different machine learning and deep learning algorithms. Unlike the traditional statistical feature selection techniques, with the use of a MGA, sets of features could be optimized dynamically, and this brought about big improvements in class performance for the majority of the algorithms used. One significant finding of this research is the superior classification performance achieved using LightGBM and Decision Tree classifiers, both achieving perfect 100% accuracy in a number of splits of the data and using less than 5-fold cross-validation. These were accompanied by Random Forest, which also achieved near-perfect performance, being found to have the best predictive capability when trained from the attributes selected by the MGA. This observation points to the capacity of the algorithm to draw out very informative features with strong correlation to tree-based model robustness and structure.

The same could not be said of the DL models which did not enjoy the same improvements. On the contrary, it was even worse than before feature selection when their performance was determined. The reason is that the deep learning models are typically more reliant on richer and more complex features spaces, in particular, deep learning models specialized to the relationship of space or sequence [19]. The effective sparse feature space though applicable in traditional models is probably one of the factors limiting the learning ability of such neural networks. This therefore poses a very critical point of observation. The other important contribution of the work is that extensive testing was done on different data splits (70 30, 75 25, 80 20) and 5-fold cross-validation. This rigorous validation was done to make sure that the performance gains could not only be restricted to a specific split but the instability was at least partially transferable under different training conditions. Relative to the past literature that concentrates on ensemble learning where the statistical features are predetermined, the present literature demonstrates how evolutionary computation can be used to achieve adaptive and autonomous feature optimization. This is because the MGA is able to choose discriminative features in a manner independent of the pre-statistical assumptions hence more flexible and better generalizability to other datasets and classification problems. In general, the results have proven the effectiveness and robustness of the proposed MGA-based feature selection framework for cardiovascular disease risk prediction, particularly in enhancing the performance and generalizability of traditional machine learning classifiers.

4.5 Validation of the Proposed MGA

To validate the efficiency of the proposed MGA, comparison is done with standard GA, particle swarm optimization and filter-based methods Chi-square.

Table 6. Performance comparison of MGA, standard GA, PSO and Filter bases methods

Method	Accuracy (Avg)
Filter	~88%
PSO	~92%
GA	~94%
MGA	98%

All the methods are assessed with the identical data split to ensure the fairness in results. The results shows that proposed MGA perform best amongst all methods in term of accuracy. MGA shoes the improved prediction performance in comparison with the classical ML models. Filter based methods are efficient but observed lack of adaptability leading to degrade the performance. Table 6 shows the performance comparisons of MGA with different methods.

4.6 Stability and Variance Analysis

For robustness and reliability of the results all the experiments were repeated five times with random runs and data split. Table 7 shows results are computed with mean results across multiple runs the along with computed standard deviation. The obtained results shows that the performance is consistent even with low variance representing the proposed MGA based feature selection approach has stable outcomes. Particularly, for certain model like the XGBoost in specific split obtained perfect AUC values confirms that such performance is not due to the random chance but shows the steady predictive competence. The values described in the table 5 represent the mean values of the results obtained over the five runs with random initialization. The standard deviations for all the model's evolution are low typically ranges from the value of ±0.2 to ±1.5 based on the complexity of the model. Traditional ML model shows the low variance while DL models shows high variance due to training variations. The observations shows that the performance is steady and are not inclined towards any favourable data split.

Table 7. Variance Analysis

Model	Accuracy Std Dev	AUC Std Dev
RF	±0.45	±0.30
XGB	±0.35	±0.10
CNN	±1.20	±0.85
LSTM	±1.40	±1.10

4.7 Ablation Study

To examine the role of MGA feature selection, an ablation study is done with selective features from the feature subset. We examine the full feature set and progressively reduce the feature during MGA process. The results shown Table 8, that the model is improved initially with top 12 and top 10 features. Beyond this point further reduction leads to decrease in the accuracy and AUC. This confirms that the selected features signify the best balance between the information and model generality. This shows the efficiency of the MGA within compressed feature space.

4.8 Limitations and Future Work

Despite the fact that the proposed MGA based feature selection method was found to have a good performance in the prediction of CVD, some limitations remain that should be met [20]. Although the proposed framework demonstrates strong predictive performance, it is important to note that the results are derived from a single dataset. In this research single Kaggle dataset with 1025 samples is used therefore, cannot be generalized without further validation. External validation on independent cohorts, along with calibration assessment and testing in real-world clinical settings, is necessary to establish the robustness, reliability, and broader applicability of the proposed model. The adaptive nature of the MGA model adds computational overhead that might limit its use in computational benchmark study.

This method proved to be more effective using traditional ML models as compared to DL models, which would indicate necessary improvements. It negatively impacts DL models due to loss of feature richness, indicating the need for alternative feature selection strategies for such architectures. It has also taken into account only structured clinical data, which is why in the future, other studies should be conducted to increase the flexibility of the model and the ability to incorporate data [21]. These limitations are important research directions in future research work.

Future work will focus on multi-dataset validation, inclusion of heterogeneous and real-world clinical data, and development of hybrid feature selection for deep learning models. These steps are necessary before considering integration into clinical decision-making environments.

Table 8. Ablation Study

No. of Features	Selected Features (Example)	Accuracy (%)	AUC (%)
14 (All)	All features	92.50	94.00
12	Top 12 features (MGA ranking)	95.20	96.80
10	Top 10 features	96.90	98.10
8	Top 8 features	98.90	99.60
6	Top 6 features	97.80	98.50
5	Top 5 features	96.40	97.20

5. Conclusion

In this study, a strong prediction model of heart diseases was designed on the basis of ML algorithms and an adapted genetic algorithm to find the best features to predict CVD. The model addressed the high-dimensional medical data issue through reducing the computational complexity and improving prediction accuracy. LightGBM, as well as Decision Tree gave the highest score in terms of precision, recall, F1-score, and AUC-ROC. Random Forest fared almost equally as well, and it had few or no misclassifications. The MGA enhanced the model performance by determining a best but sparse and highly informative set of features making it more interpretable and resistant to overfitting. The stability and validity of the proposed approach in real-life healthcare practice is evidenced by the high consistency of the outputs in different types of validation approaches. Future studies might include considering hybrid approaches to feature selection, increasing the size of the sample, and considering multi-modal data. The MGA-enhanced models displayed better performance compared to traditional models with all features, which validates the importance of the custom feature selection to the predictive modeling framework. This method will also help to come up with effective and data-driven solutions that will assist clinicians to detect cardiovascular diseases early and accurately. Future efforts will be on incorporating the deep learning methods, conducting real-time verification in personality care environment, and the diversity of the data to enhance the reliability of the models in various groups of people.

References

[1] M.U. Rehman, S. Naseem, A.U.R. Butt, T. Mahmood, A.R. Khan, I. Khan, J. Khan, Y. Jung, Predicting Coronary Heart Disease with Advanced Machine Learning Classifiers for Improved Cardiovascular Risk Assessment. Scientific Reports, 15(1), (2025) 13361. https://doi.org/10.1038/s41598-025-96437-1

[2] H.F. El-Sofany, Predicting Heart Diseases Using Machine Learning and Different Data Classification Techniques. IEEE Access, IEEE, 12, (2024) 106146–106155. https://doi.org/10.1109/ACCESS.2024.3437181

[3] A.K. Mandal, S. Dehuri, P.K.D. Sarma, Analysis of Machine Learning Approaches for Predictive Modeling in Heart Disease Detection Systems. Biomedical Signal Processing and Control, 106, (2025) 107723. https://doi.org/10.1016/j.bspc.2025.107723

[4] R. Kumar, S. Garg, R. Kaur, M.G.M. Johar, S. Singh, S.V. Menon, P. Kumar, A.M. Hadi, S.A. Hasson, J. Lozanović, A Comprehensive Review of Machine Learning for Heart Disease Prediction: Challenges, Trends, Ethical Considerations, and Future Directions. Frontiers in Artificial Intelligence, 8, (2025) 1583459. https://doi.org/10.3389/frai.2025.1583459

[5] N.A. Hussain, A.A. Mohammed, Early Heart Attack Detection using Hybrid Deep Learning Techniques. Information, 16(5), (2025) 334. https://doi.org/10.3390/info16050334

[6] K. Venkatesh N. Mohanprabu, B. Sakthi Karthi Durai, N. Nithiyanandam, Early Prediction and Diagnosis of Heart Attack using Machine Learning Techniques. In AIP Conference Proceedings, AIP Publishing LLC, 3345(1). https://doi.org/10.1063/5.0300215

[7] D.G. Honi, L. Szathmary, A One-Dimensional Convolutional neural network-based Deep Learning Approach for Predicting Cardiovascular Diseases. Informatics in Medicine Unlocked, 49, (2024) 101535. https://doi.org/10.1016/j.imu.2024.101535

[8] S. Paliwal, S. Praveen, M.A. Alam, J. Ahmed, Optimizing Heart Disease Prediction Models Using Genetic Algorithms: A Metaheuristic Approach. Journal of Theoretical and Applied Information Technology, 102(9), (2024) 3868–3876. https://jatit.org/volumes/Vol102No9/14Vol102No9.pdf

[9] M.A. Talukder, A.S. Talaat, M. Kazi, HXAI-ML: A Hybrid Explainable Artificial Intelligence based Machine Learning Model for Cardiovascular Heart Disease Detection. Results in Engineering, 25, (2025) 104370. https://doi.org/10.1016/j.rineng.2025.104370

[10] B.A.M. Metwally, N.E. Mekky, I.M. Elhenawy, Heart Disease Prediction using Genetic Algorithm with Machine Learning Classifiers. International Journal of Advanced Science, Engineering and Information Technology, 12(5), (2022) 1887. https://doi.org/10.18517/ijaseit.12.5.16015

[11] E.Y. Hidayat, Y.P. Astuti, I.N. Dewi, A. Salam, M.A. Soeleman, Z.A. Hasibuan, A.S. Yousif, Genetic Algorithm-based Convolutional Neural Network Feature Engineering for Optimizing Coronary Heart Disease Prediction Performance. Healthcare Informatics Research, 30(3), (2024) 234–243. https://doi.org/10.4258/hir.2024.30.3.234

[12] M. Alshraideh, N. Alshraideh, A. Alshraideh, Y. Alkayed, Y. Al Trabsheh, B. Alshraideh, Enhancing Heart Attack Prediction with Machine learning: A study at Jordan University Hospital. Applied Computational Intelligence and Soft

Computing, 2024(1), (2024) 5080332. https://doi.org/10.1155/2024/5080332

[13] C. Bou Rjeily, G. Badr, A. Hajjar El Hassani, E. Andres, Medical Data Mining for Heart Diseases and the Future of Sequential Mining in Medical Field. In Machine Learning Paradigms, Intelligent Systems Reference Library, Springer, Cham, 149, (2018) 71–99. https://doi.org/10.1007/978-3-319-94030-4_4

[14] P.K. Shrivastava, M. Sharma, A. Kumar, HCBiLSTM: A Hybrid Model for Predicting Heart Disease using CNN and BiLSTM Algorithms. Measurement: Sensors, 25, (2023) 100657. https://doi.org/10.1016/j.measen.2022.100657

[15] K. Maharana, S. Mondal, B. Nemade, A Review: Data Pre-Processing and Data Augmentation Techniques. Global Transitions Proceedings, 3(1), (2022) 91–99. https://doi.org/10.1016/j.gltp.2022.04.020

[16] M.T. Mustapha, O.O. Olugbara, T. Marwala, Class Imbalance and its Impact on Predictive Models for Binary Classification of Disease: A Comparative Analysis. Artificial Intelligence and Image Processing in Medical Imaging, (2024) 389–408. https://doi.org/10.1016/B978-0-323-95462-4.00014-5

[17] D. Normawati, D.P. Ismi, K-fold Cross Validation for Selection of Cardiovascular Disease Diagnosis Features by Applying Rule-Based Datamining. Signal Processing Letters, 1(2), (2019) 62–72. https://doi.org/10.31763/simple.v1i2.3

[18] M.A. Bouqentar, O. Terrada, S. Hamida, S. Saleh, D. Lamrani, B. Cherradi, A. Raihani, Early Heart Disease Prediction using Feature Engineering and Machine Learning Algorithms. Heliyon, 10(19), (2024) e38731. https://doi.org/10.1016/j.heliyon.2024.e38731

[19] I.S. Al-Mahdi, S.M. Darwish, M.M. Madbouly, Heart Disease Prediction Model Using Feature Selection and Ensemble Deep Learning with Optimized Weight. Computer Modeling in Engineering & Sciences, 143(1), (2025) 875–909. https://doi.org/10.32604/cmes.2025.061623

[20] T. Banerjee, İ. Paçal, A Systematic Review of Machine Learning in Heart Disease Prediction. Turkish Journal of Biology, 49(5), (2025) 600–634. https://doi.org/10.55730/1300-0152.2766

[21] M. Hamid, F. Hajjej, A.S. Alluhaidan, N.W. bin Mannie, Fine-tuned CatBoost Machine Learning Approach for Early Detection of Cardiovascular Disease through Predictive Modeling. Scientific Reports, 15(1), (2025) 31199. https://doi.org/10.1038/s41598-025-13790-x

[22] A. Rancea, I. Anghel, T. Cioara, Edge Computing in Healthcare: Innovations, Opportunities, and Challenges. Future Internet, 16(9), (2024) 329. https://doi.org/10.3390/fi16090329

[23] T. Datsi, B.A. Benali, M. Srifi, M. Hachimi, K. El Kharrachi, A Short Survey on Multimodal Deep Learning Models in Healthcare. In Advances in Intelligent Systems and Digital Applications, Lecture Notes in Networks and Systems, Springer, Cham, 1485, (2025) 199–211. https://doi.org/10.1007/978-3-031-95326-2_20

[24] M.T. Mustapha, D.U. Ozsahin, Class Imbalance and its Impact on Predictive Models for Binary Classification of Disease: A Comparative Analysis. Artificial Intelligence and Image Processing in Medical Imaging, (2024) 389–408. https://doi.org/10.1016/B978-0-323-95462-4.00014-5

[25] P.J.M. Ali, R.H. Faraj, E. Koya, P.J.M. Ali, R.H. Faraj, Data Normalization and Standardization: A Technical Report. Machine Learning Technical Reports, 1(1), (2014) 1–6. https://docs.google.com/document/d/1x0A1nUz1WWtMCZb5oVzF0SVMY7a_58KQulqQVT8LaVA/edit?tab=t.0#heading=h.4ykexgmwxgcw

[26] S.K. Zhou, H. Greenspan, C. Davatzikos, J.S. Duncan, B. Van Ginneken, A. Madabhushi, J.L. Prince, D. Rueckert, R.M. Summers, A Review of Deep Learning in Medical Imaging: Imaging Traits, Technology Trends, Case Studies with Progress Highlights, and Future Promises. Proceedings of the IEEE, IEEE, 109(5), (2021) 820–838. https://doi.org/10.1109/JPROC.2021.3054390

[27] I. Gupta, M. Bajaj, V. Malhotra, V. Sharma, A. Abraham, Heart Disease Prediction Using a Hybrid Feature Selection and Ensemble Learning Approach. IEEE Access, 13, (2025) 111926–111937. https://doi.org/10.1109/ACCESS.2025.3583757

Authors Contribution Statement

Shamal Salunkhe: Conceptualization, Methodology, Investigation, Project Administration, Writing – Original Draft. Smita Bharne: Methodology, Investigation, Data Curation, Validation, Writing – Review & Editing. Poonam Patil: Formal Analysis, Visualization, Supervision. Pragati Akre: Formal Analysis, Validation, Writing – Review & Editing. Kairavi Patra: Investigation, Validation, Formal Analysis. Anmol Thakur: Investigation, Validation, Formal Analysis. Anya Thakur: Investigation, Validation, Formal Analysis. All the authors have read and agreed to the published version of the manuscript.

Data Availability Statement

The dataset utilized in this study is publicly accessible on Kaggle at the following link: https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset/data.

Funding

The authors declare that no funds, grants or any other support were received during the preparation of this manuscript.

Competing Interests

The authors declare that there are no conflicts of interest regarding the publication of this manuscript.

Data Availability

The data supporting the findings of this study can be obtained from the corresponding author upon reasonable request.

Has this article screened for similarity?

Yes

About the License