http://jhmi.sums.ac.ir
Health Man & Info Sci
2026, 13(2), 98–111

HMIS
Health Management and Information Science

Original Article

Enhancing Clinical Reliability in Coronary Artery Disease (CAD) Diagnosis: A Hybrid Machine Learning Approach Utilizing Genetic Algorithm-Optimized AdaBoost Decision Trees

Parisa Eslami[1], Narges Mahmoudi[2,3]
[1]Health Human Resources Research Center, Shiraz University of Medical Sciences, Shiraz, Iran
[2]Department of Management and Health Information Technology, School of Management and Medical Information Sciences, Isfahan University of Medical Sciences, Isfahan, Iran
[3]Health Information Technology Research Center, Isfahan University of Medical Sciences, Isfahan, Iran

Abstract:

Introduction: Accurate and reliable diagnosis of coronary artery disease (CAD) is critical, requiring predictive models that minimize classification errors, particularly False Positives (FPs), given the high clinical and ethical costs associated with them. This study aimed to identify an optimized classifier with maximal predictive reliability using the Z-Alizadeh Sani dataset.

Methods: The methodology employed a dual-stage feature selection process comparing the Genetic Algorithm (GA) and Particle Swarm Optimization (PSO). The Synthetic Minority Over-Sampling Technique (SMOTE) was subsequently applied to mitigate class imbalance. Models were evaluated using a strict performance criterion that considered a comprehensive set of metrics, including Sensitivity, Accuracy, and the Area Under the Receiver Operating Characteristic curve (AUC).

Results: The proposed hybrid model, GA + AdaBoost DT, demonstrated the highest overall predictive reliability. It achieved an accuracy of 95.88% and an exceptionally high Specificity of 98.08%, significantly surpassing comparable high-accuracy models in the literature. Furthermore, the model achieved a robust AUC of 0.972. The analysis confirmed that the GA effectively identified a superior, minimal feature subset.

Conclusion: The model offers a clinically superior diagnostic tool, with a low false-positive rate, ensuring highly reliable patient screening. This study establishes a new benchmark for reliable CAD diagnosis by prioritizing clinical safety while optimizing specificity.

Keywords: Cardiovascular disease, Coronary Artery Disease, Machine Learning, Genetic Algorithms, Decision Trees

Article History:
Received: 19 April 2026
Accepted: 20 June 2026

Please cite this paper as:
Eslami P, Mahmoudi N. Enhancing Clinical Reliability in Coronary Artery Disease (CAD) Diagnosis: A Hybrid Machine Learning Approach Utilizing Genetic Algorithm-Optimized AdaBoost Decision Trees. Health Man & Info Sci. 2026; 13(2): 98-111. Doi: 10.30476/jhmi. 2026.109799.1341

***Correspondence to:**
Parisa Eslami, Department of Health Information Management, School of Health Management and Information Sciences, Shiraz University of Medical Sciences, Shiraz, Iran. ORCID: 0000-0001-9788-740X. **Email:** parisaeslami@sums.ac.ir

Introduction

Cardiovascular Diseases (CVDs) impose a substantial global health burden, consistently ranking among the leading causes of mortality and disability worldwide(1, 2). Within the spectrum of CVDs, Coronary Artery Disease (CAD), characterized by narrowing or obstruction of the coronary arteries due to the accumulation of atherosclerotic plaque, holds particular clinical significance(3, 4). Consequently, the timely and precise diagnosis of CAD is paramount for

effective patient management and minimizing adverse health outcomes (5).

The current clinical gold standard for accurately quantifying the extent and severity of coronary stenosis is coronary angiography (6, 7). However, its widespread applicability is limited by its invasive nature, high associated costs, and the inherent risk of complications, ranging from minor side effects to serious adverse events (8).

Given the limitations of conventional invasive procedures, the medical and research communities have increasingly prioritized the development of non-invasive, rapid, and cost-effective methodologies for CAD prediction and diagnosis(9-11). In this context, the application of Machine Learning (ML) methodologies on clinical datasets has emerged as a robust and promising alternative(12-14). ML enables the analysis of vast volumes of electronic health records and clinical data, thereby extracting complex patterns and hidden knowledge essential to developing intelligent clinical decision support systems(15, 16).

The availability of reliable data is critical for advancing ML models in medicine. Standardized resources, such as the UCI ML Repository, provide several benchmark datasets relevant to cardiovascular research. Specifically, the Z-alizadehSani dataset has attracted attention in recent studies due to its comprehensive feature set and minimal missing values (17). Prior research using this dataset has successfully employed various classification and optimization algorithms to evaluate CAD prediction models, consistently demonstrating high accuracy and validating the efficacy of this non-invasive approach (18-22).

The primary objective of this study is to develop and rigorously evaluate an optimized ML model for the non-invasive detection of CAD utilizing the Z-alizadehSani dataset. This research aims to improve diagnostic accuracy beyond current standards and, significantly, to identify the most clinically and laboratory-relevant characteristics associated with CAD. The successful identification of these key predictive factors will make a critical contribution by providing cardiologists with valuable non-invasive tools for effective CAD risk stratification and timely diagnosis.

Methods

The protocol of this study was approved by the Ethics Committee of Shiraz University of Medical Sciences (SUMS) under the approval code (IR.SUMS.NUMIMG.REC.1405.003). The study was conducted in accordance with the institutional and Iranian government ethical guidelines for clinical research. This study proposes the development of an optimized machine learning (ML) model for automated CAD diagnosis. The methodology employed in this study is structured around six critical stages: data acquisition, data preprocessing, feature selection, classification and parameter tuning, model evaluation, and rule extraction and decision analysis. Each stage is described in detail in the following subsections.

Data Description Dataset

The empirical foundation for this investigation is the Z-alizadehSani dataset, an authentic, publicly accessible resource documented within the UCI ML Repository. This dataset is specifically structured for supervised learning and contains an explicit target variable for classification (17).

The data collection involved a meticulous effort by Iranian researchers over 8 months, encompassing 303 patient records obtained from individuals who presented with chest pain at the Shahid Rajaee Medical and Cardiovascular Research Center in Tehran, Iran.

This comprehensive dataset comprises a total of 54 features, broadly categorized into four groups: 16 features on Demographic Information, 14 features detailing Clinical Examination Results and Symptoms, 17 features covering Laboratory and Echocardiographic Data, and seven features related to Electrocardiography (ECG), Depression, and Exercise Test data (detailed in Table 1).

The diagnostic outcome, designated as the target variable Cath, is determined based on the

definitive results of coronary angiography and is classified into two binary categories:
The CAD class, representing 50% or greater stenosis in at least one of the three major coronary arteries, accounts for 216 samples.
The Normal class, characterized by less than 50% stenosis, includes 87 samples.

Table 1. Description of the selected features used in Alizadeh Sani's dataset

Demographic	Age	30–86
	Weight	48–120
	Sex	Male, female
	BMI (Body Mass Index) kg/m^2	18–41
	DM (Diabetes Mellitus)	Yes, no
	HTN (hypertension)	Yes, no
	Current smoker	Yes, no
	Ex-Smoker	Yes, no
	FH (family history)	Yes, no
	Obesity	Yes, if BMI > 25, no otherwise
	CRF (chronic renal failure)	Yes, no
	CVA (Cerebrovascular Accident)	Yes, no
	Airway disease	Yes, no
	Thyroid Disease	Yes, no
	CHF (congestive heart failure)	Yes, no
	DLP (Dyslipidemia)	Yes, no
Symptom and examination	BP (blood pressure: mmHg)	90–190
	PR (pulse rate) (ppm)	50–110
	Edema	Yes, no
	Weak peripheral pulse	Yes, no
	Lung rales	Yes, no
	Systolic murmur	Yes, no
	Diastolic murmur	Yes, no
	Typical Chest Pain	Yes, no
	Dyspnea	Yes, no
	Function class	1, 2, 3, 4
	Atypical	Yes, no
	Nonanginal CP	Yes, no
	Exertional CP (Exertional Chest Pain)	Yes, no
	Low Th Ang (low Threshold angina)	Yes, no
ECG	Rhythm	Sin, AF
	Q Wave	Yes, no
	ST Elevation	Yes, no
	ST Depression	Yes, no
	T inversion	Yes, no
	LVH (left ventricular hypertrophy)	Yes, no
	Poor R progression (poor R wave progression)	Yes, no
Laboratory and Echo	FBS (fasting blood sugar) (mg/dl)	62–400
	Cr (creatine) (mg/dl)	0.5–2.2
	TG (triglyceride) (mg/dl)	37–1050
	LDL (low density lipoprotein) (mg/dl)	18–232
	HDL (high density lipoprotein) (mg/dl)	15–111
	BUN (blood urea nitrogen) (mg/dl)	6–52
	ESR (erythrocyte sedimentation rate) (mm/h)	1–90
	HB (hemoglobin) (g/dl)	8.9–17.6
	K (potassium) (mEq/lit)	3.0–6.6
	Na (sodium) (mEq/lit)	128–156
	WBC (white blood cell) (cells/ml)	3700–18,000
	Lymph (Lymphocyte) (%)	7–60
	Neut (neutrophil) (%)	32–89
	PLT (platelet) (1000/ml)	25–742
	EF (ejection fraction) (%)	15–60
	Region with RWMA (regional wall motion abnormality)	0, 1, 2, 3, 4
	VHD (valvular heart disease)	Normal, mild, moderate, severe

Data Preprocessing

Before implementing the ML algorithms, a series of essential preprocessing steps was applied to the dataset to ensure data quality and optimize its structure.

Data cleaning and integrity assessment

The initial phase of data preprocessing focused on ensuring the completeness and integrity of the dataset, a critical prerequisite for the validity of all subsequent modeling and analysis. This step involved detecting and addressing common data quality issues, specifically missing values and duplicate records.

Class Imbalance Mitigation

Subsequently, the issue of Class Imbalance was addressed. The original dataset exhibits a significant imbalance: 216 samples in the CAD class and only 87 in the Normal class. Direct application of learning algorithms to such an imbalanced dataset can lead to model bias toward the majority class and reduced predictive accuracy for the minority class. To mitigate this imbalance and enhance performance metrics, rather than relying solely on overall accuracy, the Oversampling technique was applied to the minority class.

Specifically, the Synthetic Minority Over-sampling Technique (SMOTE) was implemented in Weka. SMOTE overcomes the limitations of simple random oversampling, which only duplicates existing samples, by generating synthetic instances in the feature space. The SMOTE procedure involves selecting a minority class sample (xi € Smin), identifying its K nearest neighbors based on the Euclidean distance, and then randomly choosing one of these neighbors ($\widetilde{x_i}$). A new synthetic sample (X_{new}) is generated along the line segment connecting x_i and $\widetilde{x_i}$ Using the following formula (Equation 1):

$$X_{new} = x_i + (\widetilde{x_i} - x_i) \times \sigma \qquad (1)$$

where $\widetilde{x_i}$ represents one of the K nearest neighbors of x_i, and x_i It is a random number in the range [0, 1]. This technique ensures that the generated synthetic data maintains spatial similarity with the actual minority class observations(23).

Table 2 illustrates the adjusted parameters for the SMOTE technique.

Parameter	Setting	Rational
Class Value	1 (Minority Class)	Designated the 'Normal' class for oversampling.
Nearest Neighbors (*k*)	5	Determined the local neighborhood for synthesizing new data points.
Percentage	100%	Ensured the minority class count was doubled to achieve a near 1:1 balance with the majority class.
Random Seed	1	Fixed the random state to guarantee the reproducibility of the generated synthetic dataset.

Feature Selection

In this phase, two metaheuristic feature selection algorithms—the Genetic Algorithm (GA) and Particle Swarm Optimization (PSO)—were chosen to evaluate and select the most influential variables for the CAD prediction model. This section details the operational mechanism of each technique and its approach to feature weighting and selection.

Genetic Algorithm (GA)

The GA algorithm is an optimization technique based on the principles of natural genetics and evolution, making it well-suited to solving complex optimization problems.

The overall process begins with the random initialization of a population of candidate solutions, where each solution represents a potential subset of features. These candidate solutions then undergo successive generations of evolution, with the fitness of each being rigorously evaluated against a predefined criterion. Subsequent generations are created by applying core genetic operators, such as Crossover and Mutation, and the process continues until a predetermined stopping condition is satisfied. In the context of feature subset selection using GA, mutation corresponds to the probabilistic decision to include or exclude individual features in a given iteration. At the same time, crossover involves swapping features between two parent solutions. Selection methods, such as the roulette wheel or tournament selection, ensure that fitter individuals contribute to the next generation. The process is parametrically defined: an initial set of P variables is selected with an initial probability (P initial); subsequently, mutation is applied with a specific probability (P mutation) to modify feature inclusion; and finally, crossover between two selected variables occurs with a probability of P crossover. Critically, the fitness function that quantifies the optimality of each feature subset is implemented as a classification model. In every generation, the model is trained, and its performance (typically accuracy) is evaluated using the current feature subset, thereby directing the evolutionary search toward the most effective combination of input variables (24).

Particle Swarm Optimization (PSO)

Particle Swarm Optimization (PSO) is a social search algorithm inspired by the collective behavior of bird flocking. Initially, this algorithm was developed to uncover the patterns governing the synchronous flight of birds, their sudden trajectory changes, and the mechanism for achieving an optimal swarm shape. The method operates on the principle that each particle (analogous to a bird in the flock) traverses the search space in pursuit of the optimal solution, actively exchanging positional information within the group. Critically, every particle maintains a memory of its own best local position ($P_{best,\ i}$) achieved thus far, as well as the best global position (G_{best}) found by any particle in the swarm up to that point. This collective memory guides the particles toward the most optimal path and position.

Three core factors dynamically influence the movement of each particle: 1) the particle's current position ($x_i(t)$); 2) the best position the particle has individually reached (the cognitive, or memory, effect, $P_{best,\ i}$); and 3) the best position achieved by the entire collective (the social effect, G_{best}).

The new position of each particle ($x_i(t+1)$) is calculated using its updated velocity ($v_i\ (t+1)$) as follows (Equation 2):

$$x_i\ (t+1) = x_i\ (t) + v_i(t+1) \qquad (2)$$

The velocity, which represents the rate of displacement, is updated by the following relation (Equation 3):

$$v_i(t+1) = wv_i(t) + c_1r_1(P_{best,i} - x_i\ (t)) + c_2r_2(G_{bes} - x_i\ (t)) \qquad (3)$$

In this equation, the first term accounts for the particle's current momentum, and the second term accounts for its memory. It guides it toward its P_{best}, and the final term represents the social influence that directs the particle toward the swarm's G_{best}.

For feature selection, the PSO algorithm searches for the optimal subset of features from the available set. The position of each particle is represented by d binary variables, where d is the total number of features. A value of 0 signifies the non-selection of a feature, and 1 denotes its selection by the swarm. The fitness function used to evaluate and guide the swarm's movement is defined by the accuracy of a classification model, which serves as a wrapper. Specifically, the Decision Tree classifier is embedded within this

operator to serve as the primary evaluation function for feature subset optimality(25).

Classification Algorithms and Parameter Tuning

In this phase of the research, CAD modeling and prediction are performed using a curated selection of classification algorithms that have demonstrated robust performance in prior studies using this dataset.

To construct the predictive model and derive actionable diagnostic rules, the following classifiers were selected: Decision Tree, Naïve Bayes, Random Forest, Bagging Decision Tree, and Boosting Decision Tree.

Each of these classifiers has key parameters for fine-tuning and optimization. The final parameter values for each classifier were determined empirically following multiple simulation and testing stages. This optimization process aimed to achieve the highest possible accuracy, with the chosen parameters documented in Table 3.

Table 3 Optimal Hyperparameters for Classification Models

Model Name	Hyperparameter Settings
Decision Tree	Split Criterion: Gini Index (for splitting nodes) Maximum Depth: 10 Minimum Leaf Size: 3 Confidence Factor for Pruning: 0.25 Minimum Number of Records for Termination: 3
Bagging Decision Tree	Number of Iterations: 10 Sample Selection Ratio (Bootstrap Sample Size): 0.4
Random Forest	Split Criterion: Gini Index (for splitting nodes) Maximum Depth: 15 Minimum Leaf Size: 3 Confidence Factor for Pruning: 0.25 Minimum Number of Records for Termination: 3 Number of Trees (Ensembles): 10
AdaBoost	Number of Iterations (Weak Learners): 10

Model Evaluation

In supervised learning problems, such as classification, two distinct datasets are required: training and test. The hold-out method was employed for model assessment. Across all execution stages, a portion of the dataset is used for training the model, with the remainder reserved for testing and calculating performance metrics(26). In this research, after multiple simulation stages in the software, the data were partitioned into 75% for training and 25% for testing.

Before detailing the classification criteria, it is essential to define the Confusion Matrix, the foundational tool for evaluating model performance. The Confusion Matrix illustrates how the classification algorithm performs on the input data by explicitly separating results into the actual and predicted classes of the classification problem(27).

The four key concepts in this matrix are:

True Positive (TP): The count of records whose actual class is positive and the classifier correctly predicted them as positive.

False Positive (FP): The count of records whose actual class is negative, but the classifier incorrectly predicted them as positive.

True Negative (TN): The count of records whose actual class is negative and the classifier correctly predicted them as negative.

False Negative (FN): The count of records whose actual class is positive, but the classifier incorrectly predicted them as negative.

We now turn to the formal definition of the critical evaluation metrics for classification algorithms. The first metric is Accuracy, which represents the percentage of test records correctly classified by the model. Accuracy calculated as follows (Equation 4):

$$\text{Accuracy} = \frac{\text{Tp} + \text{TN}}{\text{Tp} + \text{TN} + \text{Fp} + \text{FN}} \quad (4)$$

The most critical metric in this research, especially for disease diagnosis, is Sensitivity. This metric represents the number of correct disease diagnoses made by the classifier relative to the total number of truly diseased cases (the sum of TPs and FNs). Sensitivity calculated as follows (Equation 5):

$$Sensitivity = \frac{TP}{FN + TP} \quad (5)$$

Another evaluation metric is Specificity. In the context of disease diagnosis, Specificity represents the classifier's ability to identify records that are non-diseased. Specificity calculated as follows (Equation 6):

$$Specificity = \frac{TN}{TN + FP} \quad (6)$$

A further crucial metric for determining a classifier's overall efficacy is the Area Under the Receiver Operating Characteristic Curve (AUC-ROC). These curves are two-dimensional plots in which the horizontal axis represents the FP Rate (FPR) and the vertical axis represents the TP Rate (TPR). These curves illustrate the classification behavior independent of class distribution or misclassification costs, thereby isolating the classifiers' intrinsic effectiveness from these external factors(27). However, in imbalanced datasets, AUC alone is not always a sufficient metric. The calculated AUC value for each classifier will be presented in the subsequent results section.

Note that, in the software environment used in this study, the disease class was treated as the negative class, and the Normal class as the positive class. While this alters the software's default class assignment, the fundamental concepts derived from the Confusion Matrix remain valid, and the goal is to achieve high performance across all defined criteria.

Rule Extraction and Decision Analysis

Following the determination of the optimal feature subset by the GA algorithm, where the most influential variables were designated with an influence weight of 1, the process of decision rule extraction commenced. The analysis of the generated rules focuses on the structure of the best-performing model to provide transparent, interpretable diagnostic criteria.

This methodological approach allows the hierarchical and logical structure of the decision tree to be transformed into direct "If-Then" rules. The primary goal of this stage is to establish the methodology for converting the identified influential variables into straightforward diagnostic patterns, which will be presented and analyzed in the Results section.

Results

Results of Data Cleaning

During this comprehensive assessment, the dataset was analyzed, and it was conclusively determined that no duplicate entries or missing values were present across the entire sample. This high level of initial data integrity simplified the preprocessing pipeline, allowing a direct transition to feature engineering and model development without imputation or record deduplication.

Results of Class Imbalance Mitigation

To address the significant class imbalance in the dataset, which initially comprised 216 CAD samples (the majority class) and only 87 Normal samples (the minority class), the SMOTE method was implemented in Weka.

The minority class, designated 'Normal' (87 samples), was oversampled at 100%. This procedure effectively doubled the representation of the minority class by generating 87 synthetic instances (87 original + 87 synthetic = 174). This strategic level of oversampling was chosen to mitigate model bias towards the majority class without introducing the potential noise and

complexity associated with higher oversampling rates.

The final dataset, used for subsequent feature selection and modeling, comprised 390 samples (216 CAD and 174 Normal), achieving a near-balanced class distribution.

Results of Feature Selection

The feature selection process, executed using two metaheuristic approaches within the RapidMiner software environment, aimed to identify the optimal, most influential subset from the original 54 variables for the CAD prediction model.

GA Feature Weighting

The GA was implemented using the Optimize Selection (Evolutionary) operator to systematically evaluate each variable's predictive contribution. The algorithm first assigned a weight to every feature. Subsequently, the Select by Weights operator was applied to filter the results based on a binary classification:

A weight value of 1 signified a highly influential variable.

A weight value of 0 indicated that a variable was non-influential or low-impact.

The feature weights assigned by the GA are detailed in Table 4. The final optimal number of influential variables to include in the subsequent classification models was determined by maximizing the classifiers' overall accuracy.

Table 2.Feature Selection Outcomes from the GA Algorithm

Feature Category	Selected Features (weight=1)	Rejected Features (weight=1)
Demographic	Age, DM, HTN, CRF, Airway disease, Thyroid disease	Weight, Sex, BMI, Current smoker, Ex-Smoker, FH, Obesity, CVA, CHF, DLP
Symptom and examination	PR, Typical Chest Pain, Edema, Atypical, Weak peripheral pulse, Exertional CP, Low the Ang	BP, Function class, Systolic murmur, Diastolic murmur, Nonanginal CP, Dyspnea
ECG	T inversion, ST Depression, Q Wave, ST Elevation	Poor R progression, LVH
Laboratory and echo	Na, TG, K, HB, LDL, HDL	CR, Neut, Lymph, FBS, PLT, EF, BUN, ESR, WBC, Region with RWMA, VHD

PSO Feature Weighting

The PSO algorithm was also employed to evaluate the 54 variables, following a comparable workflow. PSO first calculated and assigned a weight to all variables. The Select by Weights operator was then utilized to isolate variables exhibiting higher assigned weights. The resulting dataset, containing the selected feature weights determined by PSO, is also detailed in Table 5. The assessment of the optimal number of features identified by the PSO was deferred to the final modeling phase, where the performance of the classifiers using this subset was rigorously assessed.

Table 3.Feature Selection Outcomes from the PSO Algorithm

Feature Category	Influential Features (Weight >0)	Non-Influential Features (Weight =0)
Demographic	DM (1), Thyroid Disease (1), CVA (1), CHF (1), Current Smoker (1), Weight (0.798), BMI (0.756), Airway disease (0.250)	HTN, Age, Sex, FH, Obesity, CRF, DLP

Symptom and examination	Typical chest pain (1), Atypical (1), Nonanginal CP (1), Diastolic murmur (1), Function class (1), Weak peripheral pulse (0972), Systolic murmur (0.849), Dyspnea (0.812), Low the Ang (0.416), Exertional CP (0.263)	Edema, BP, PR, Lung rales,
ECG	Poor R progression (1), ST Depression (1), LVH (1), ST Elevation (0.374)	T inversion, Q Wave, Rhythm
Laboratory and echo	TG (1), BUN (1), WBC (1), K (1), HDL (1), VHD (0.709), PLT (0.665), ESR (0.151), NA (0.048), HB (0.011)	CR, FBS, Lymph, Neut, EF, Region with RWMA, LDL

Results of model evaluation

Table 6 summarizes the performance metrics for all ten hybrid models, grouped by feature selection method. These results allow a direct comparative assessment of the chosen classifiers and the influence of the feature subsets selected by the GA and PSO algorithms.

This table clearly establishes the GA + AdaBoost DT model as the superior performer, with Specificity (98.08%) and overall Accuracy (95.88%).

Based on our results, the GA proved highly effective at selecting a non-redundant, highly predictive feature subset. The AdaBoost ensemble technique conferred a crucial advantage over bagging methods by adaptively focusing on difficult misclassified instances, and the application of SMOTE was essential for mitigating class imbalance, resulting in significant performance gains.

Table 4. Performance of classifiers on Z-Alizadeh Sani dataset

Model Name	Accuracy	Sensitivity	Specificity	AUC
GA + Adaboost DT	95.88%	93.33%	98.08%	0.972
PSO + Adaboost DT	91.75%	93.33%	90.38%	0.919
GA + B- DT	91.75%	95.56%	88.46%	0.931
PSO + B-DT	90.72%	95.56%	86.54%	0.906
GA + RF	89.69%	95.56%	84.62%	0.912
PSO + RF	91. 75%	91.11%	92.13%	0.978
GA + NB	88.66%	88.89%	88.46%	0.946
PSO + NB	86.60%	88.89%	84.62%	0.944
GA + DT	94.85	93.33%	96.15%	0.960
PSO+ DT	91.75%	93.33%	90.38%	0.931

DT: Decision Tree, B-DT: Bagging Decision Tree, RF: Random Forest, NB: Naive Bayes

Results of Rule Extraction and Decision Analysis

To identify the most influential features for the final model, variables assigned a predictive weight of 1 by the Genetic Algorithm (GA) were selected, as this feature set demonstrated superior performance. To analyze the decision structure and extract actionable knowledge, the rules generated by the best-performing model, the GA + AdaBoost DT, were utilized. A segment of the resulting decision process is visualized in Figure 1.

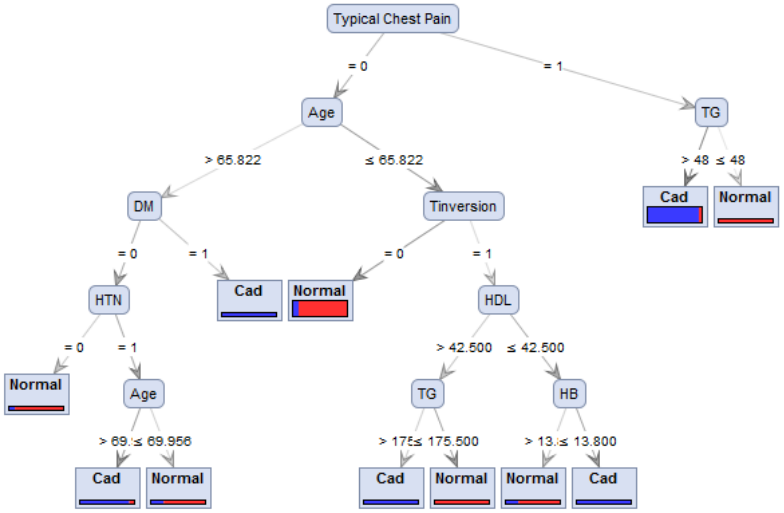

Figure 1 .Visualization of the Decision Tree Structure for the Optimal GA AdaBoost DT Classifier

The decision logic of the optimized GA + AdaBoost DT ensemble model was analyzed by extracting specific IF-THEN rules from the learned trees. This analysis highlights the critical interplay of clinical and biochemical markers in classifying CAD. The GA-based feature selection ensured that these rules were based only on the most influential variables. The complete set of extracted decision rules is provided in Appendix A.

A representative sample of the most clinically significant rules is presented in Tables 7 and 8. The certainty of classification is indicated by the ratio of CAD to Normal instances in the terminal leaf ({CAD, Normal}).

Table 7 illustrates rules that identify patients with a high probability of having CAD, often driven by the presence of key symptoms or adverse lab values.

Table 8 illustrates the rules used for Low-Risk/Normal classification. These rules directly demonstrate the model's high Specificity—its ability to confidently rule out CAD—by identifying cases in which the absence of critical features leads to a negative diagnosis.

Based on the extracted rules, the following clinical insights highlight the model's alignment with established medical knowledge:

Triglycerides (TG) and Chest Pain: The presence of Typical Chest Pain combined with a high TG value (> 48 mg/dL) forms a highly predictive, low-depth rule for CAD (Rule R1), validating the synergistic risk posed by symptoms and dyslipidemia.

Age and Hypertension: Even in the absence of Typical Chest Pain, advanced age combined with established risk factors like Hypertension (HTN) or Diabetes Mellitus (DM) immediately places the patient in a high-risk category (Rules R3, R4), accurately reflecting cumulative cardiovascular risk.

T-Inversion as a Safety Check: The absence of T-inversion (a standard ECG marker of ischemia) serves as a robust 'safe' condition, particularly in younger, non-diabetic patients (Rules R5, R6). This finding validates the model's emphasis on Specificity for reliably ruling out disease.

It is essential to note that the extracted decision rules must be carefully reviewed and validated by a clinical Domain Expert before clinical deployment.

Discussion

The primary objective of this evaluation was to select the classifier with the highest predictive reliability for diagnosing CAD. The GA + AdaBoost DT model was selected as the optimal performer. It demonstrated the highest overall performance, achieving the highest Accuracy (95.88%) and an exceptionally high Specificity (98.08%) among all models tested. This level of Specificity is the highest recorded, confirming its superior ability to reliably identify non-diseased cases. The GA + DT model ranked second, with a Specificity of 96.15% and an overall Accuracy of 94.85%. Conversely, the NB classifier consistently performed worst, indicating its limited capacity to capture the necessary non-linear feature interactions.

The critical finding is the Specificity of 98.08%. This figure significantly exceeds that achieved by many similar studies in the literature using this dataset.

For instance, while an RF model achieved 94.74% accuracy (20) and an Optimized DT achieved 91.66% Specificity (28) Our model's Specificity is demonstrably superior.

This directly translates into a minimal FPR. This makes the GA + AdaBoost DT model highly reliable for patient screening, surpassing the clinical viability of even high-accuracy competitors such as Optimized Deep Neural Networks with 95.36% Accuracy (29). Furthermore, another study (30) achieved a high accuracy of 96.0% using a sophisticated hybrid model, but this was offset by a lower Specificity of 94.7%, confirming that the GA + AdaBoost DT model maintains the highest standard for safe non-disease identification.

The results strongly indicate the effectiveness of the feature selection strategy. Models using features selected by the GA generally achieved significantly higher accuracy and specificity than those chosen by the PSO method. This confirms the GA's robust ability to identify a more predictive and less redundant feature subset. This success aligns with findings by prior research (31), which utilized the Gini Index for feature ranking and found that models built on their reduced set achieved high accuracy, further validating the critical necessity of an optimal feature selection step over using the complete feature set.

Furthermore, the Boosting techniques consistently produced superior results compared to the Bagging methods. This outcome is attributed to AdaBoost's adaptive learning process, which iteratively focuses on misclassified samples by increasing their weight. The implementation of the SMOTE technique for class-imbalance mitigation also proved vital, yielding an average accuracy improvement of up to 40% compared to baseline models, thereby validating the necessity of this preprocessing step. Although the literature reports slightly higher AUCs from complex models (32)The combination of a high AUC and the highest reported Specificity in our model is crucial. This outcome confirms the optimal balance necessary for a reliable diagnostic setting, prioritizing clinical safety alongside strong discriminative power.

Conclusion

This research successfully developed and validated a robust hybrid ML model for the non-invasive diagnosis of CAD using the Z-Alizadeh Sani dataset. Our primary objective was to identify a classifier with maximal predictive reliability, prioritizing high Specificity to ensure minimal FP in a clinical screening context. The GA + AdaBoost DT model was selected as the optimal performer, demonstrating its superiority across all key metrics. The model achieved a Specificity of 98.08%, the highest reported result in the comparative literature using this dataset, alongside a strong Accuracy of 95.88% and an AUC of 0.972. This high specificity confirms its superior ability to reliably rule out disease,

directly enhancing its clinical viability for patient screening. The principal contribution of this work is the development of an optimized hybrid model that balances exceptional discriminatory power with unparalleled safety in the non-diseased class, thereby establishing a new benchmark for reliable CAD diagnosis. Future work could focus on developing a transparent clinical decision support system that integrates the extracted decision rules to facilitate adoption among healthcare professionals.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data Availability

The datasets analyzed during the current study are available in the UCI Machine Learning Repository. The specific feature subset generated by the GA algorithm and the complete set of extracted rules are included in the manuscript and its supplementary files (Appendix A).

Ethics Approval

As this study involves the analysis of clinical data retrieved from the Z-Alizadeh Sani dataset, all procedures were conducted in accordance with the Declaration of Helsinki. The dataset utilized consists of anonymized patient records, and ethical approval was obtained by the original data collectors. No further administrative permissions were required for the secondary analysis of this publicly available/de-identified data.

Conflict of Interest

The authors declare that they have no competing interests that could have influenced the work reported in this paper.

References

1. Bao X, Zhang Z, Shen Y, Tang P, Si G, Kang L, et al. Cardiovascular Information and Health Engineering Medicine. 2025;8:0956.
2. Di Cesare M, Perel P, Taylor S, Kabudula C, Bixby H, Gaziano TA, et al. The heart of the world. 2024;19(1):11.
3. Mir MA, Dar MA, Qadir AJAJBP. Exploring the Landscape of Coronary Artery Disease: A Comprehensive Review. 2024;1:9-22.
4. Sanyaolu SE, Wilson DO, Oguntolu BA, Salako MS, Adeyemi ROJP. Current Evidence on the Clinical Management of Coronary Artery Disease and Acute Coronary Syndrome: A Narrative Review.2:3.
5. Netala VR, Hou T, Wang Y, Zhang Z, Teertam SKJIJoMS. Cardiovascular biomarkers: tools for precision diagnosis and prognosis. 2025;26(7):3218.
6. Jiangping S, Zhe Z, Wei W, Yunhu S, Jie H, Hongyue W, et al. Assessment of coronary artery stenosis by coronary angiography: a head-to-head comparison with pathological coronary artery anatomy. 2013;6(3):262-8.
7. Suzuki N, Asano T, Nakazawa G, Aoki J, Tanabe K, Hibi K, et al. Clinical expert consensus document on quantitative coronary angiography from the Japanese Association of Cardiovascular Intervention and Therapeutics. 2020;35(2):105-16.
8. Tavakol M, Ashraf S, Brener SJJGjohs. Risks and complications of coronary angiography: a comprehensive review. 2012;4(1):65.
9. Kluge B, Harris J, Sánchez-Collado I, Pérez-Román I, Paffett M, Harz C, et al. Cost savings from prioritization of non-invasive modalities within CAD diagnostic protocols: a systematic review. 2025;28(1):1388-404.
10. Wang LWJIMJ. Non-invasive screening for coronary artery disease: current perspectives, patient, public health and ethical considerations in evaluating symptomatic and asymptomatic individuals. 2025;55(4):555-63.
11. Apostolopoulos ID, Groumpos PPJCMiB, Engineering B. Non-invasive modelling methodology for the diagnosis of coronary artery disease using fuzzy cognitive maps. 2020;23(12):879-87.
12. Tasmurzayev N, Imanbek B, Boltaboyeva A, Dikhanbayeva G, Zhussupbekov S, Saparbayeva Q, et al. Explainable AI for Coronary Artery Disease Stratification Using Routine Clinical Data. 2025;18(11):693.
13. Alizadehsani R, Hosseini MJ, Khosravi A, Khozeimeh F, Roshanzamir M, Sarrafzadegan N, et al. Non-invasive detection of coronary artery disease in high-risk patients based on the stenosis prediction of separate coronary arteries. 2018;162:119-27.

14. Ezekwueme F, Tolu-Akinnawo O, Smith Z, Ogunniyi KEJC. Non-invasive Assessment of Coronary Artery Disease: The Role of AI in the Current Status and Future Directions. 2025;17(2).
15. Ahmed Z, Mohamed K, Zeeshan S, Dong XJD. Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine. 2020;2020:baaa010.
16. Vasamsetty CJIJoME, Engineering C. Clinical decision support systems and advanced data mining techniques for cardiovascular care: Unveiling patterns and trends. 2020;8(2).
17. Alizadehsani R, Habibi J, Hosseini MJ, Mashayekhi H, Boghrati R, Ghandeharioun A, et al. A data mining approach for diagnosis of coronary artery disease. 2013;111(1):52-61.
18. Arabasadi Z, Alizadehsani R, Roshanzamir M, Moosaei H, Yarifard AAJCm, biomedicine pi. Computer aided decision making for heart disease detection using hybrid neural network-Genetic algorithm. 2017;141:19-26.
19. Cüvitoğlu A, Işik Z, editors. Classification of CAD dataset by using principal component analysis and machine learning approaches. 2018 5th International Conference on Electrical and Electronic Engineering (ICEEE); 2018: IEEE.
20. Gupta A, Arora HS, Kumar R, Raman B, editors. DMHZ: a decision support system based on machine computational design for heart disease diagnosis using z-alizadeh sani dataset. 2021 International Conference on Information Networking (ICOIN); 2021: IEEE.
21. Hashemi M, Komamardakhi SSS, Maftoun M, Zare O, Joloudari JH, Nematollahi MA, et al., editors. Enhancing Coronary Artery Disease Classification Using Optimized MLP Based on Genetic Algorithm. International Work-Conference on the Interplay Between Natural and Artificial Computation; 2024: Springer.
22. Jalali SMJ, Karimi M, Khosravi A, Nahavandi S, editors. An efficient neuroevolution approach for heart disease detection. 2019 IEEE international conference on Systems, Man and Cybernetics (SMC); 2019: IEEE.
23. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WPJJoair. SMOTE: synthetic minority over-sampling technique. 2002;16:321-57.
24. Lambora A, Gupta K, Chopra K, editors. Genetic algorithm-A literature review. 2019 international conference on machine learning, big data, cloud and parallel computing (COMITCon); 2019: IEEE.
25. Wang D, Tan D, Liu LJSc. Particle swarm optimization algorithm: an overview. 2018;22(2):387-408.
26. Raschka SJapa. Model evaluation, model selection, and algorithm selection in machine learning. 2018.
27. Rashidi HH, Albahra S, Robertson S, Tran NK, Hu BJFiO. Common statistical concepts in the supervised Machine Learning arena. 2023;13:1130229.
28. Sayadi M, Varadarajan V, Sadoughi F, Chopannejad S, Langarizadeh MJL. A machine learning model for detection of coronary artery disease using noninvasive clinical parameters. 2022;12(11):1933.
29. Nasarian E, Sharifrazi D, Mohsenirad S, Tsui K, Alizadehsani RJapa. AI Framework for Early Diagnosis of Coronary Artery Disease: An Integration of Borderline SMOTE, Autoencoders and Convolutional Neural Networks Approach. 2023.
30. Mansoor C, Chettri SK, Naleer H, editors. Predicting Coronary Artery Disease using an Ensemble Voting Model and Machine Learning Techniques. 2023 9th International Conference on Smart Structures and Systems (ICSSS); 2023: IEEE.
31. Vijayaraj AR, Pasupathi S, editors. Enhancing Cardiac Health: Machine Learning in Coronary Artery Disease Prediction. International Conference on Computational Intelligence in Pattern Recognition; 2024: Springer.
32. Zhang M, Wang H, Zhao JJPo. Use machine learning models to identify and assess risk factors for coronary artery disease. 2024;19(9):e0307952.

Appendix

Full Set of Extracted IF-THEN Logic Hierarchies and Decision Paths of the GA + AdaBoost Ensemble

This appendix provides the complete decision logic for the ensemble model. The paths represent the hierarchical decision-making process, where features selected by the Genetic Algorithm (GA) are used to stratify patients into CAD and Normal categories. Classification certainty is indicated by the {CAD, Normal} distribution at each terminal leaf.

Typical Chest Pain = 0
| Age > 65.822

```
| | DM = 0
| | | HTN = 0: Normal {Cad=1, Normal=7}
| | | HTN = 1
| | | | Age > 69.956: Cad {Cad=11, Normal=1}
| | | | Age ≤ 69.956: Normal {Cad=1, Normal=3}
| | DM = 1: Cad {Cad=15, Normal=0}
| Age ≤ 65.822
| | Tinversion = 0: Normal {Cad=20, Normal=144}
| | Tinversion = 1
| | | HDL > 42.500
| | | | TG > 175.500: Cad {Cad=2, Normal=0}
| | | | TG ≤ 175.500: Normal {Cad=0, Normal=6}
| | | HDL ≤ 42.500
| | | | HB > 13.800: Normal {Cad=1, Normal=3}
| | | | HB ≤ 13.800: Cad {Cad=11, Normal=0}
Typical Chest Pain = 1
| TG > 48: Cad {Cad=154, Normal=8}
| TG ≤ 48: Normal {Cad=0, Normal=2}
  ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
DM = 0
| Age > 46.986: Cad {Cad=132, Normal=130}
| Age ≤ 46.986
| | Tinversion = 0: Normal {Cad=0, Normal=32}
| | Tinversion = 1
| | | Typical Chest Pain = 0: Normal {Cad=0, Normal=2}
| | | Typical Chest Pain = 1: Cad {Cad=4, Normal=0}
DM = 1
| PR > 62: Cad {Cad=79, Normal=8}
| PR ≤ 62: Normal {Cad=1, Normal=2}
DM = 0
| K > 4.650
| | HDL > 29.218
| | | Age > 76.500: Cad {Cad=2, Normal=1}
| | | Age ≤ 76.500
| | | | Age > 51.500: Cad {Cad=20, Normal=0}
| | | | Age ≤ 51.500
| | | | | LDL > 138.104: Normal {Cad=0, Normal=2}
| | | | | LDL ≤ 138.104: Cad {Cad=3, Normal=0}
| | HDL ≤ 29.218: Normal {Cad=1, Normal=2}
| K ≤ 4.650
| | St Depression = 0: Normal {Cad=79, Normal=143}
| | St Depression = 1
| | | HTN = 0
| | | | HB > 13.950
| | | | | HB > 14.953: Normal {Cad=0, Normal=4}
| | | | | HB ≤ 14.953: Cad {Cad=6, Normal=0}
| | | | HB ≤ 13.950
| | | | | Typical Chest Pain = 0
| | | | | | Age > 49.500: Normal {Cad=0, Normal=8}
| | | | | | Age ≤ 49.500: Cad {Cad=1, Normal=1}
| | | | | Typical Chest Pain = 1: Cad {Cad=5, Normal=1}
| | | HTN = 1: Cad {Cad=19, Normal=2}
DM = 1
| TG > 156: Cad {Cad=39, Normal=0}
| TG ≤ 156
| | Age > 47.500
| | | Typical Chest Pain = 0
| | | | Age > 59
| | | | | LDL > 93: Cad {Cad=8, Normal=0}
| | | | | LDL ≤ 93
| | | | | | HDL > 26.500: Normal {Cad=1, Normal=2}
| | | | | | HDL ≤ 26.500: Cad {Cad=2, Normal=0}
| | | | Age ≤ 59: Normal {Cad=0, Normal=3}
| | | Typical Chest Pain = 1: Cad {Cad=28, Normal=0}
| | Age ≤ 47.500
| | | Tinversion = 0: Normal {Cad=0, Normal=5}
| | | Tinversion = 1: Cad {Cad=2, Normal=0}
  ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
```