

Article

An Optimized PSO-PNN Hybrid Model for Enhancing Diagnostic Accuracy in Cardiovascular Disease Prediction

Norah Altimyat ^{1,*} , Maali Alshammari ¹, Khawlah Alshammari ¹, Aljawharah Alshammari ¹ and Jihane Ben Slimane ²

¹ Department of Information Systems, Faculty of Computing and Information Technology, Northern Border University, Rafha 91911, Saudi Arabia; st202506014@stu.nbu.edu.sa (M.A.); st202506013@stu.nbu.edu.sa (K.A.); st202506006@stu.nbu.edu.sa (A.A.)

² Department of Computer Sciences, Faculty of Computing and Information Technology, Northern Border University, Rafha 91911, Saudi Arabia; jehan.saleh@nbu.edu.sa

* Correspondence: st202506027@stu.nbu.edu.sa

Abstract

Cardiovascular Disease (CVD) is a prevailing issue across the world. It requires an accurate method for diagnosing in order to treat the disease effectively. The use of machine learning (ML) techniques has gained popularity for diagnosing CVDs. Some existing ML models have demonstrated good results using the process of manual hyperparameter tuning. But manual hyperparameter tuning is a very tedious task that may cause fluctuations in diagnosis accuracy. To construct a PSO-tuned classifier that provides stable diagnosis for CVD, this paper proposes a robust hybrid framework named PSO-PNN using a particle swarm optimization algorithm for fine-tuning the value of the Smoothing Parameter (σ) of a probabilistic neural network (PNN). PSO-PNN could be employed as an effective classifier for cardiovascular disease. To validate the effectiveness of PSO-PNN, a dataset obtained from UCI Cleveland database was utilized containing 303 patients' data. The results indicated that the proposed PSO-PNN framework improved the baseline PNN performance; obtained an accuracy of 91.3%, an ROC-AUC value of 94.8%, and a recall rate of 95.24%; and showed stable predictive performance across 30 independent PSO executions. Additionally, SHAP was applied as a post hoc explainability method to interpret the trained PSO-PNN predictions and identify the contribution of input features. The proposed framework may support future decision-support applications for CVD prediction; however, further external validation is required before practical clinical use can be considered.

Keywords: heart disease prediction; medical machine learning; probabilistic neural network (PNN); particle swarm optimization (PSO); explainable AI (XAI); clinical decision support systems



Academic Editor: José Manuel García-Nieto

Received: 31 May 2026

Revised: 31 July 2026

Accepted: 14 August 2026

Published: 21 August 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Cardiovascular diseases (CVDs) remain the leading cause of death worldwide, accounting for an estimated 19.8 million deaths globally according to the World Health Organization (WHO) [1]. This burden is associated with major risk factors, including aging, metabolic conditions, and lifestyle-related risks, which are also reported in recent heart disease and stroke statistics [2]. It should be noted that the issue of CVDs is not restricted to the adult population only but also manifests itself in young people, hence the necessity for a reliable prediction system. Based on the findings in the studies conducted in Saudi Arabia, it was established that the main cause of deaths in the country is cardiovascular

diseases, with 37% to 42% mortality rates [3]. There are also other types of cardiovascular diseases like coronary artery disease, heart valves diseases, and cardiomyopathies, which are life threatening and need to be diagnosed in a timely manner. Another critical category of cardiovascular diseases is congenital heart disease (CHD). In this type, the person suffers from heart abnormalities from childhood, affecting about eight in every one thousand live births [4,5].

However, while it is critical for early predictions in a clinical setting, the use of classical statistical and clinical scoring systems still poses a challenge because these methods cannot represent complicated nonlinear relations inherent to the medical data adequately. Therefore, machine learning techniques became widely used approaches that can effectively process data in predicting heart diseases [6–8]. Specifically, several ML models have been utilized for this task, including support vector machines (SVMs), k-nearest neighbors (KNN), Random Forest, and boosting algorithms such as XGBoost and LightGBM [9]. However, despite these approaches having proved efficient at prediction tasks, there are some shortcomings that should be noted. First, models based on the distance metric (KNN) and on the margin concept (SVM) are affected by the choice of parameter and by the scale of features. Moreover, ensemble algorithms such as Random Forest as well as all boosting techniques tend to lead to overfitting when applied to relatively small or imbalanced clinical samples [10–12]. In order to mitigate such drawbacks, current developments have been focused on hybrid learning paradigms as well as optimization-based solutions. The use of particle swarm optimization (PSO) for tuning the hyperparameters and accelerating the process of model training for medical prediction applications has become popular [13]. Likewise, the use of eXplainable Artificial Intelligence (XAI), represented by methods such as SHAP, has become common among researchers as a tool for enhancing the interpretability and transparency of decisions made using clinical data, which is crucial when it comes to trusting AI [14]. Moreover, research related to the use of signal processing techniques for the purpose of feature extraction from cardiovascular signals has been performed. Such an approach has resulted in increased accuracy of diagnosis in the task of heart disease recognition [15]. Recent intelligent healthcare studies have investigated attention mechanisms for feature refinement and lightweight model architectures in computer vision applications [16,17]. However, the use of these approaches with structured cardiovascular data remains limited. Even though the machine learning algorithms have been found to perform very effectively, it is notable that most of the existing algorithms focus on some of the particular issues such as classification and optimization but not on the combination of all of these issues at once. Additionally, the effectiveness of the probabilistic neural network is influenced by its smoothing parameter and the selected input features [18,19]. Furthermore, irrelevant features in the clinical data can impact adversely on the stability and prediction accuracy [13,20]. Therefore, there is a need for a unified and adaptive framework that integrates feature selection, PSO-based parameter tuning and probabilistic learning in a single pipeline. Recent studies have explored hybrid metaheuristic frameworks for CVD prediction. Notably, Reddy and Murthy [21] proposed a PSO-NN framework that integrates particle swarm optimization with a Multi-Layer Perceptron Neural Network on the UCI heart disease dataset and compares several classifier/metaheuristic combinations. Their study highlights the role of PSO in improving neural-network-based prediction through hyperparameter tuning.

In contrast, the proposed PSO-PNN framework focuses on a probabilistic neural network (PNN), where PSO is used to tune the smoothing parameter σ . Unlike standard feedforward neural networks that rely on iterative weight learning, the PNN uses a probabilistic classification structure based on Parzen-window density estimation. Therefore, the proposed framework extends related PSO-based hybrid approaches by focusing on

smoothing-parameter tuning in the PNN, combined with mutual-information-based feature selection for heart disease prediction.

A comparative summary of representative studies and their reported performance values is presented in Table 1.

Table 1. Summary of previous studies using machine learning for heart disease and related cardiovascular prediction tasks and their reported performance.

Reference	Algorithms Used	Dataset and Validation Method	Performance
Naser et al., 2024 [7]	Support vector machine (SVM), K-Nearest Neighbors (KNN), Decision Tree (DT)	Review study; varied datasets and validation methods	85–89.3% accuracy
Almutairi & Dardouri, 2025 [3]	Random Forest (RF), support vector machine (SVM), Logistic Regression (LR), Hybrid Ensemble	Kaggle heart disease dataset; stratified 80:20 holdout	Acc. 89.3%, Prec. 90%, Rec. 91%
Naskinova et al., 2026 [8]	Random Forest (RF), Multiple Linear Regression (MLR), Artificial Neural Network (ANN)	Kaggle cardiovascular dataset; stratified 64/16/20 train/validation/test split	88.0–90.1%
Villoth et al., 2026 [10]	Two-Tier Classification, Adaptive Variable Neighborhood Search (VNS), Machine Learning Classifiers (RF, SVM, NB, etc.)	Kaggle CVD risk dataset; stratified 70:30 holdout	90.0%
Waqar et al., 2025 [22]	Support vector machine (SVM), Random Forest (RF), XGBoost, Logistic Regression (LR)	UCI heart attack dataset; 10-fold CV	86.0–89.5%
Tompra et al., 2024 [11]	Support vector machine (SVM), particle swarm optimization (PSO) for Hyperparameter Tuning	BRFSS 2021 dataset; holdout and cross-validation	88.4%
Ullah et al., 2022 [19]	Support vector machine (SVM), Random Forest (RF), K-Nearest Neighbors (KNN), feature selection (MI, ReliefF)	Cleveland dataset; 70:30 holdout	87.0–89.0%
Gupta et al., 2026 [23]	Stacked LR, SVM, RF, XGBoost, and MLP	Cleveland dataset; 5-fold CV	Acc. 91%, Prec. 92%, Rec. 90%, F1 91%
Chandrasekhar & Peddakrishna, 2023 [24]	LR, RF, KNN, NB, boosting, and soft voting	UCI Cleveland dataset; Grid-SearchCV with 5-fold CV	Acc. 90.16% (LR) and 93.44% (soft voting)
Reddy & Murthy, 2025 [21]	Particle swarm optimization (PSO) + Multi-Layer Perceptron Neural Network (PSO-NN); comparison of 13 classifier/metaheuristic combinations	UCI cardiovascular disease dataset; comparative experimental evaluation	10–30% cost/score improvement over baseline classifiers (UCI dataset)
Our Work	MI-based feature selection + PSO-tuned probabilistic neural network (PSO-PNN) for smoothing-parameter σ tuning in heart disease prediction	UCI Cleveland dataset; stratified 70/15/15 holdout split and stratified 10-fold cross-validation	Holdout: Acc. 91.30%, Recall 95.24%, ROC-AUC 94.86%; 10-fold CV: Acc. $84.45 \pm 4.56\%$, Recall $76.92 \pm 7.60\%$, ROC-AUC $89.74 \pm 4.03\%$

Note: Reported performance values are presented as stated in the cited studies; direct comparisons should be interpreted cautiously because datasets, validation methods, and reported metrics may differ.

For the current study, a novel hybrid model known as PSO-PNN, which combines the power of particle swarm optimization and probabilistic neural network, is introduced as a predictive solution for early detection of cardiovascular diseases. Specifically, a particle swarm optimization algorithm is utilized to optimize the smoothing factor of the PNN model, which will reduce human intervention in selecting the value of the smoothing factor [13]. Mutual Information (MI) was adopted as a feature selection method to select statistically informative features with respect to the target outcome. Feature selection can

help retain informative features and eliminate irrelevant features [19]. In addition, SHAP was applied as a post hoc explainability method to interpret the predictions of the proposed model that builds on the traditional SHAP method by tailoring the concept of feature attributions towards medicine [14].

The UCI Heart Disease dataset is a real clinical dataset of cardiovascular patients' records and is often used in medicine as a benchmark dataset. The goal of this research study is to develop a better and more reliable prediction model for cardiovascular disease identification through the use of mutual-information based feature selection and probabilistic learning. This study presents an approach that combines MI-based feature selection with PSO-based tuning of the PNN smoothing parameter leading to enhanced performance of the model. Moreover, SHAP was used to provide post hoc explanations of the model predictions.

2. Materials and Methods

All computational experiments were conducted using Python 3.12.13 in the Google Colaboratory environment. The main software libraries used in the experiments included scikit-learn 1.6.1, XGBoost 3.3.0, LightGBM 4.6.0, and SHAP 0.52.0.

2.1. Problem Formulation

Let $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}$ be a given dataset of N clinical samples, where each feature vector $\mathbf{x}_i$ represents the clinical features of the i -th patient, with a corresponding binary label y_i indicating the presence or absence of cardiovascular disease. The data were partitioned in a stratified manner into three non-overlapping subsets: a training set $\mathcal{D}_{\text{tr}}$ (70%), a validation set $\mathcal{D}_{\text{val}}$ (15%), and a test set $\mathcal{D}_{\text{te}}$ (15%). A probabilistic neural network (PNN) classifier $\mathcal{M}_\sigma$ was trained on $\mathcal{D}_{\text{train}}$. Besides adapting the number of neurons in the input layer, a PNN also allows for an essential trade-off that is governed by a single key hyperparameter, the smoothing parameter $\sigma > 0$, which determines the width of the Gaussian kernels.

For a query point $\mathbf{x}$, the PNN estimates the class-conditional probability density of class c using a Parzen–Rosenblatt Gaussian kernel estimator:

$$\hat{f}_c(\mathbf{x}; \sigma) = \frac{1}{n_c (2\pi)^{d/2} \sigma^d} \sum_{i: y_i=c} \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}_i\|^2}{2\sigma^2}\right), \quad (1)$$

where $n_c = |\{i : y_i = c\}|$ is the number of training samples in class c . Classification was performed using the Bayes maximum-a posteriori rule:

$$\hat{y} = \arg \max_{c \in \{0,1\}} \pi_c \hat{f}_c(\mathbf{x}; \sigma), \quad (2)$$

where $\pi_c = n_c / N_{\text{train}}$ is the empirical class prior.

The predictive performance of the model is a non-trivial function of σ :

- Small σ : narrow kernels lead to memorization of training samples, resulting in overfitting and poor generalization.
- Large σ : overly smooth density estimates reduce class separability, leading to underfitting.

Accordingly, the optimization problem is defined as

$$\sigma^* = \arg \min_{\sigma \in [\sigma_{\min}, \sigma_{\max}]} \mathcal{F}(\sigma), \quad \mathcal{F}(\sigma) = 1 - \text{Accuracy}_{\text{val}}(\sigma), \quad (3)$$

where $\mathcal{F}(\sigma)$ is the validation misclassification rate evaluated on $\mathcal{D}_{\text{val}}$, and the search bounds are set to $\sigma_{\min} = 0.01$ and $\sigma_{\max} = 5.0$, covering the full range of practically meaningful kernel widths.

Since $\mathcal{F}(\sigma)$ is non-convex and non-differentiable, gradient-free optimization is required. This work employed particle swarm optimization (PSO), which is detailed in Section 2.7.

2.2. Dataset

In this study, the UCI Heart Disease dataset [25] was utilized to develop and evaluate the proposed framework. Specifically, the Cleveland Clinic Foundation subset was selected due to its widespread use in benchmarking cardiovascular disease prediction models. The used source of this database is the Cleveland Clinic Foundation. This database is made up of 303 medical records (which are records of real patients) that include 14 attributes such as age, sex, chest pain type, ST depression, slope of ST segment, and number of major vessels colored by fluoroscopy. It categorizes the five types of heart disease into two types of diseases: having disease or not having disease. Many records with missing attribute values were handled using mean and mode imputation, resulting in a final dataset of 303 records. Following data preprocessing and missing value imputation, Mutual Information (MI) was employed to rank the predictive importance of the available attributes. Based on this ranking, the eight highest-ranked attributes were selected for training the proposed PSO-PNN framework. This feature reduction aims to retain the most informative clinical variables while reducing redundant information and limiting the effect of the curse of dimensionality on the PNN density estimation process. To verify the effectiveness of the selected subset, an additional experiment was conducted using the full feature configuration with the same experimental settings. As presented in Table 2, the MI-selected subset achieved higher accuracy, recall, F1-score, and ROC-AUC than the full feature configuration while maintaining the same specificity. These results demonstrate that the selected subset provides a more informative representation of the data, leading to improved predictive performance of the proposed PSO-PNN framework.

Table 2. Performance comparison between the full feature configuration and the proposed MI-selected feature subset.

Configuration	Accuracy	Precision	Recall	Specificity	F1-Score	ROC-AUC
Full feature set	0.8696	0.8571	0.8571	0.8800	0.8571	0.9162
MI-selected subset	0.9130	0.8696	0.9524	0.8800	0.9091	0.9486

2.3. Preprocessing Pipeline

To ensure data quality and suitability for our model, a rigorous five-step preprocessing pipeline was applied to the dataset. The raw data is publicly available at the UCI Machine Learning Repository (<https://archive.ics.uci.edu/dataset/45/heart+disease> (accessed on 15 November 2024)).

1. **Missing Value Treatment:** missing values were filled in. For numbers, we used the average value, and for categories, we used the most common one.
2. **Categorical Encoding:** One-hot encoding was applied to categorical variables *cp*, *restecg*, *slope*, and *thal*, which were converted into numbers that computers can understand.
3. **Binary Target Transformation:** The original multi-class targets were transformed into binary labels (0 = no disease, 1 = disease).
4. **Feature Scaling:** Continuous numerical input (age, *thalach*, *oldpeak*, *ca*) was standardized to zero mean and unit variance using a *StandardScaler* in order to prevent features with larger scales from dominating in distance-based learning methods.

5. Stratified Data Splitting: The data was split and reserved for training (70%), validation (15%) and testing (15%) using stratified (balanced class distribution) random split with a fixed seed of 42.

2.4. Feature Selection via Mutual Information

To enhance model performance and interpretability within the proposed framework, Mutual Information (MI) was used to rank and select the most relevant features from the 14 available attributes based on their dependency with the target variable. MI was chosen because it handles both continuous and categorical variables and captures non-linear relationships. After selecting the the MI-selected feature subset, the selected features were validated to improve the reliability and interpretability of the model. MI is defined as

$$I(X;Y) = \int_{\mathcal{X}} \int_{\mathcal{Y}} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} dx dy, \quad (4)$$

and is estimated non-parametrically using k -nearest-neighbor density estimation.

For Ranking Results, Mutual Information (MI) scores were computed for all 14 attributes, and the resulting rankings are illustrated in Figure 1. Based on both the MI scores and the clinical relevance of the variables, eight features were selected for inclusion in the predictive model: age, cp, thalach, exang, oldpeak, slope, ca, and thal.

The selected subset comprises a combination of categorical and continuous variables associated with cardiovascular health. Specifically, cp (chest pain type), thal (thalassemia), exang (exercise-induced angina), and slope (slope of the ST segment) represent categorical indicators of cardiac condition, while age, thalach (maximum heart rate achieved), oldpeak (ST depression), and ca (number of major vessels colored by fluoroscopy) correspond to continuous clinical measurements.

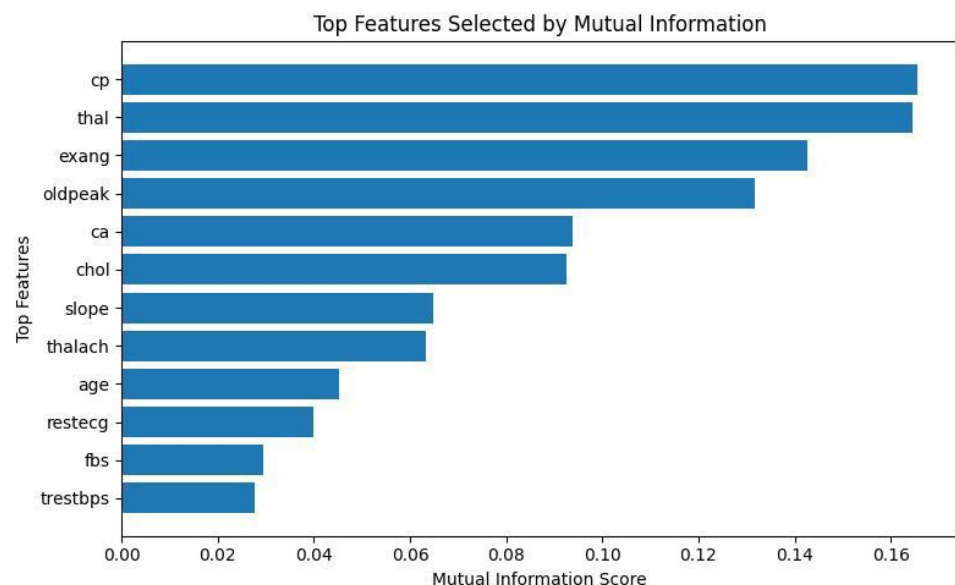


Figure 1. Mutual Information ranking and selected features for the UCI Heart Disease dataset.

The selection is consistent with established cardiovascular risk factors validated extensively in the literature and summarized in Table 3.

Table 3. Final selected features, their MI scores, types, and clinical relevance.

Feature	Type	MI Score	Clinical Meaning
cp	Categorical	0.166	Chest pain type (from typical angina to asymptomatic)
thal	Categorical	0.164	Thalassemia type (normal/fixed defect/reversible defect)
oldpeak	Continuous	0.132	ST depression induced by exercise relative to rest
ca	Ordinal	0.094	Number of major vessels (0–3) via fluoroscopy
exang	Binary	0.076	Exercise-induced angina (yes/no)
slope	Ordinal	0.071	Slope of the peak exercise ST segment (upsloping/flat/downsloping)
thalach	Continuous	0.063	Maximum heart rate achieved during exercise
age	Continuous	0.046	Patient age in years

MI scores computed via k -nearest-neighbor density estimation on the training split. Features are ordered by descending MI score.

2.5. Baseline Probabilistic Neural Network (PNN)

Within the proposed framework, the probabilistic neural network (PNN), originally introduced by [18], was adopted as the baseline classification model for heart disease prediction. The PNN is a feedforward network grounded in kernel-based non-parametric classification and Bayesian decision theory, making it well-suited for medical datasets with complex and potentially non-linear relationships. The PNN was selected in this study because it is suitable for the relatively small size of the UCI Heart Disease dataset and provides a direct parameter for tuning. Since the dataset contains 303 patient records, a classifier with a simple structure and a limited number of tunable parameters is appropriate for this setting. Unlike conventional neural networks that require iterative weight learning, the PNN estimates the class probability densities using Gaussian kernels without backpropagation.

In the PNN, the smoothing parameter σ controls the width of the Gaussian kernel and affects the model generalization. If σ is too small, the model may become highly sensitive to individual training samples. If it is too large, the estimated class distributions may become too smooth and reduce class separation. Therefore, tuning σ is an important step for improving the baseline PNN. In this work, PSO was applied to search for a suitable value of σ using the validation set. The architecture of the PNN, illustrated in Figure 2, comprises four layers: (1) the input layer, which receives the n -dimensional feature vectors; (2) the pattern layer, where a Gaussian kernel neuron is generated for each training sample; (3) the summation layer, which aggregates the kernel activations for each class to estimate the class-conditional probability density functions (pdfs); and (4) the output layer, which assigns each input instance to the class with the highest estimated posterior probability.

$$\hat{f}_c(\mathbf{x}) = \frac{1}{n_c (2\pi)^{d/2} \sigma^d} \sum_{i=1}^{n_c} \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}_{c,i}\|^2}{2\sigma^2}\right) \quad (5)$$

$$y^* = \arg \max_{c=1,2} \frac{1}{n_c} \sum_{i=1}^{n_c} e^{\frac{\mathbf{x}}{\sigma} \mathbf{w}^T \mathbf{x}_{c,i}} \quad (6)$$

where n_c is the number of training samples in class c (with a total training size $N = 212$ distributed between the two classes), $d = 8$ is the number of features, σ is the smoothing parameter (bandwidth) optimized via PSO, and $\mathbf{x}_{c,i}$ are the 212 individual training vectors of class c .

$$\hat{y} = \arg \max_c \pi_c \cdot \hat{f}_c(\mathbf{x}) \quad (7)$$

where π_c is the estimated prior probability of class c from the training data.

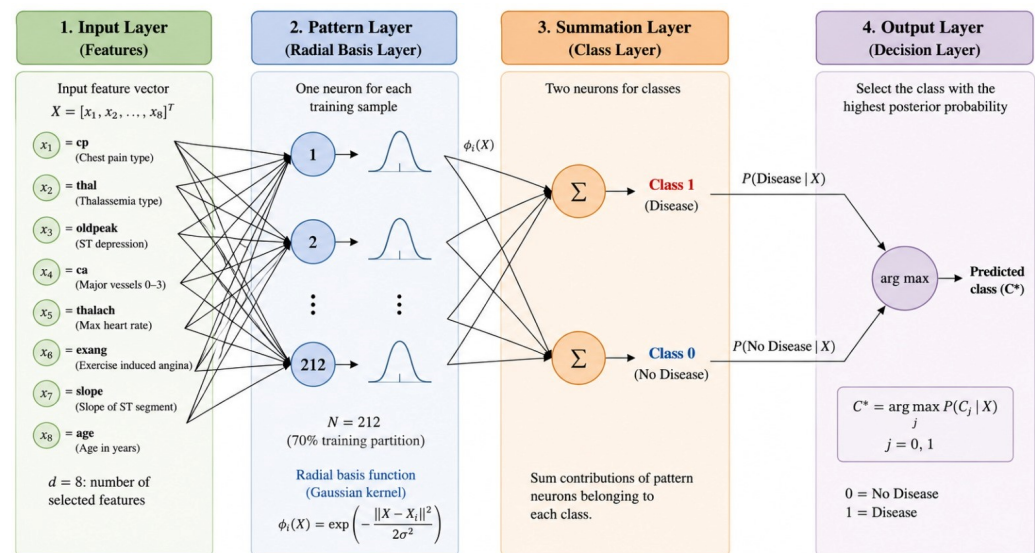


Figure 2. Architecture of the probabilistic neural network (PNN).

2.6. Sigma Sensitivity Analysis

To evaluate the impact of the smoothing parameter σ on the classification performance of the PNN, a manual sensitivity analysis was first conducted. The parameter σ was varied across an empirical range from 0.1 to 3.0, and the corresponding classification accuracy and Area Under the ROC Curve (ROC-AUC) were recorded. As illustrated in Figure 3, the model performance changes across different σ values, indicating that the choice of this parameter has a direct effect on the behavior of the PNN classifier.

The value $\sigma = 1.0$ was selected as the manual baseline because it provided a strong ROC-AUC of approximately 0.92 while maintaining competitive classification accuracy. However, this manual selection was based on empirical observation and did not involve an automated search procedure. Therefore, PSO was subsequently applied as a single-objective tuning method to search for a best-found value of σ based on validation accuracy. ROC-AUC was retained as an important threshold-independent evaluation metric and was reported during final model evaluation, but it was not used as the PSO fitness objective in the current implementation.

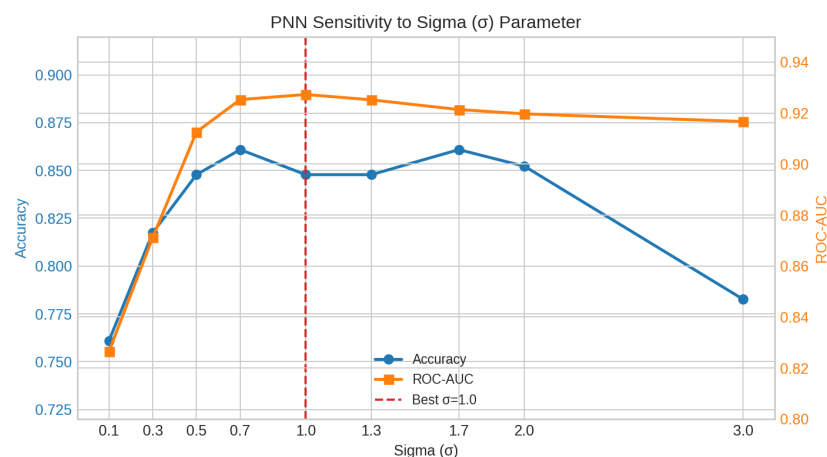


Figure 3. Sensitivity of PNN performance to the smoothing parameter σ , showing accuracy and ROC-AUC across different σ values.

2.7. Particle Swarm Optimization (PSO) for Sigma Tuning

Building on the baseline PNN model, particle swarm optimization (PSO) [26] is introduced as a heuristic stochastic search strategy for tuning the smoothing parameter σ .

In this study, we used PSO to iteratively search over the range $\sigma \in [0.01, 5.0]$. The goal of the search process was to find the value of σ that produces the minimum PNN classification error on the validation set. We utilized a swarm of $N = 20$ particles, where each particle corresponds to a candidate value of the smoothing parameter σ . This swarm size ensures sufficiently dense exploration of the one-dimensional search space, reducing the likelihood of convergence to local minima while maintaining rapid convergence during the validation process. The position and velocity of each particle at iteration t are updated according to the following equations:

$$v_k^{(t+1)} = w v_k^{(t)} + c_1 r_1 (p_k^{\text{best}} - \sigma_k^{(t)}) + c_2 r_2 (g^{\text{best}} - \sigma_k^{(t)}) \quad (8)$$

$$\sigma_k^{(t+1)} = \sigma_k^{(t)} + v_k^{(t+1)} \quad (9)$$

where $w = 0.9$ is the inertia weight controlling the exploration momentum of the particles, $c_1 = 0.5$ is the cognitive acceleration coefficient governing the particle's reliance on its personal historical best, and $c_2 = 0.3$ is the social acceleration coefficient dictating the influence of the global swarm best. The random variables r_1 and r_2 are drawn.

$$\mathcal{F}(\sigma) = 1 - \text{Accuracy}_{\text{val}}(\sigma) \quad (10)$$

This objective was selected to guide the PSO search toward a value of σ that improves validation accuracy. Therefore, the current PSO formulation is a single-objective tuning procedure based on validation accuracy, while ROC-AUC, recall, specificity, and F1-score were used only for final evaluation on the held-out test set. Although the tuning problem involves only one parameter, PSO was used to automate the search over a continuous bounded interval and reduce reliance on manual trial-and-error or a predefined grid resolution. For the optimization 30 iterations of search were performed. The parameters for the PSO search are summarized in Table 4. The convergence performance of the PSO search is shown in Figure 4. The fitness function decreased during the early iterations and then stabilized around iteration 15, where the best-found value of σ was obtained.

Table 4. PSO hyperparameter configuration.

Parameter	Symbol	Value
Number of particles	N	20
Number of iterations	T	30
Inertia weight	w	0.9
Cognitive coefficient	c_1	0.5
Social coefficient	c_2	0.3
Lower bound	$\sigma_{\min}$	0.01
Upper bound	$\sigma_{\max}$	5.00
Optimal sigma	σ^*	1.7714

The selected PSO hyperparameters were chosen to provide sufficient exploration of the one-dimensional search space while maintaining low computational cost. A swarm size of 20 particles was used to evaluate multiple candidate σ values at each iteration, and 30 iterations were adopted because the convergence curve showed stabilization before the final iteration. The inertia weight was set to $w = 0.9$ to encourage exploration, while the cognitive and social coefficients were set to $c_1 = 0.5$ and $c_2 = 0.3$ to control the influence of the personal and global best positions. The search bounds $\sigma \in [0.01, 5.0]$ were selected to cover both narrow and wide Gaussian kernel settings. The robustness of the selected PSO settings was further examined through 30 independent PSO executions using different random seeds, as discussed later.

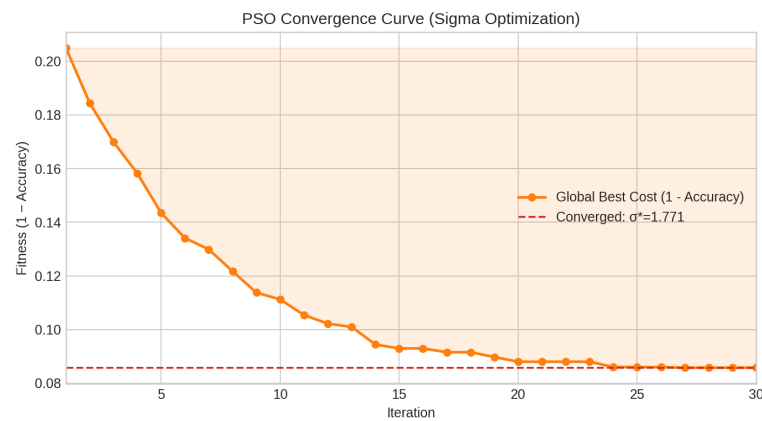


Figure 4. PSO performance during sigma optimization, showing the continued decrease in global best fitness to find $\sigma^* = 1.7714$ at iteration 15.

2.8. Overall System Workflow

The complete end-to-end workflow of the proposed framework is illustrated in Figure 5 and is organized into seven stages, namely, data collection, data preprocessing, feature selection, PNN modeling, PSO-based tuning of the smoothing parameter σ , comparative model evaluation, and performance assessment.

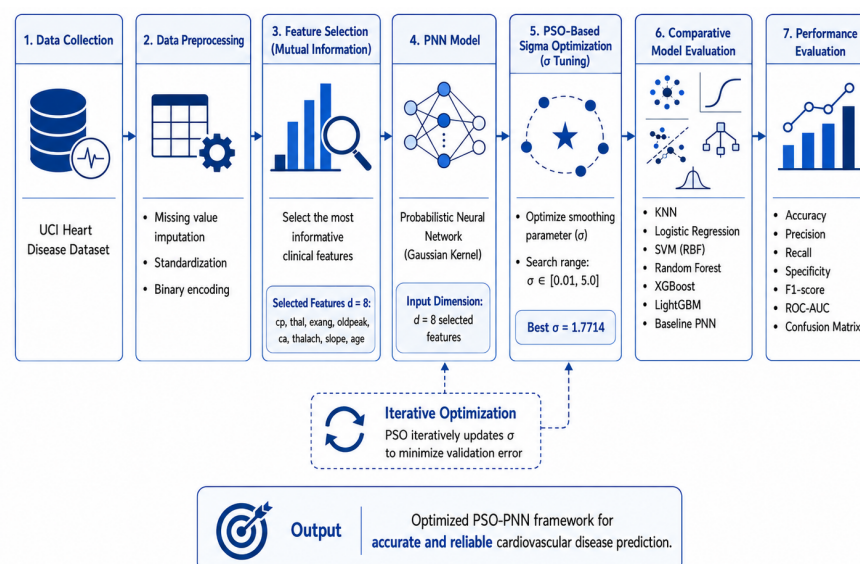


Figure 5. PSO-PNN workflow for heart disease prediction.

1. **Data Collection:** We used the publicly available UCI Heart Disease dataset for this study.
2. **Data Preprocessing:** We cleaned the data and did the necessary encoding on the columns, changing the target into binary and then performing feature scaling.
3. **Feature Selection:** Mutual Information analysis was performed in order to select the most relevant clinical features ($d = 8$): cp, thal, exang, oldpeak, ca, thalach, slope, and age.
4. **PNN Model:** A probabilistic neural network with a Gaussian Kernel was constructed using the 8 selected features as input dimensions.
5. **PSO-Based Sigma Optimization:** For a given training/validation set, we optimized the smoothing parameter σ by searching over $\sigma \in [0.01, 5.0]$, using PSO to iteratively minimize the validation error. The best value found is $\sigma = 1.7714$.

6. Comparative Model Evaluation: KNN, Logistic Regression, the SVM (RBF), Random Forest, XGBoost, LightGBM, and the standard PNN were evaluated against the optimized PSO-PNN with the same experimental setting.
7. Performance Evaluation: The PSO-optimized PNN was evaluated on the held-out test set using accuracy, precision, recall, specificity, F1-score, ROC-AUC, and a confusion matrix.

This section presents the complete PSO methodology to structured optimization problems by detailing the highest level workflow in Figure 5 and the resulting PSO iterations in Section 3, where for each iterative run of PSO, the methodology to update σ until convergence is also described.

2.9. Evaluation Metrics

To ensure a comprehensive and clinically meaningful assessment of model performance, six widely adopted evaluation metrics were employed. Let TP , TN , FP , and FN denote true positives, true negatives, false positives, and false negatives, respectively:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (11)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (12)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (13)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (14)$$

$$\text{F1-score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (15)$$

In addition to the threshold-dependent metrics above, the ROC-AUC was computed from model probability outputs and provides a threshold-independent measure of discriminative performance, which is particularly important in medical prediction tasks for reliable clinical decision support [9].

3. Results

The experimental results of the proposed PSO-PNN framework are presented in this section, following the methodological pipeline described in Section 2. The evaluation is structured into five components: (1) evaluation of baseline model performance on the UCI Heart Disease dataset; (2) analysis of the PSO-tuned PNN model; (3) confusion matrix analysis; (4) comprehensive multi-model comparison to assess the contribution of PSO-based tuning of the smoothing parameter σ ; and (5) robustness analysis using stratified 10-fold cross-validation. All reported results were obtained using a held-out test set that was not involved in either model training or hyperparameter optimization, ensuring an unbiased evaluation of generalization performance.

3.1. Baseline Model Comparison on UCI Heart Disease Dataset

To establish a performance benchmark, several standard machine learning classifiers were evaluated on the UCI Heart Disease dataset consisting of 303 patient records. The evaluated baseline models were Logistic Regression, KNN, the SVM (RBF), Random Forest [27], XGBoost [28], LightGBM [29], and a PNN. As presented in Table 5, Logistic Regression, the SVM (RBF), Random Forest, and the PNN achieved an accuracy of 0.8913, while XGBoost achieved 0.8478 and LightGBM achieved 0.7826. Among the baseline models, Logistic Regression and the SVM (RBF) obtained the highest ROC-AUC of 0.9371, followed by

Random Forest, XGBoost, the PNN, and LightGBM. The baseline PNN achieved the highest precision and specificity, but its recall remained lower than several classifiers, including Logistic Regression, the SVM (RBF), Random Forest, and XGBoost. This result supports the need for tuning the PNN smoothing parameter to improve its sensitivity while maintaining its probabilistic structure.

Table 5. Baseline model performance on the UCI Heart Disease dataset.

Model	Accuracy	Precision	Recall	Specificity	F1-Score	ROC-AUC
Logistic Regression	0.8913	0.8333	0.9524	0.8400	0.8889	0.9371
KNN	0.8043	0.7500	0.8571	0.7600	0.8000	0.8476
SVM (RBF)	0.8913	0.8636	0.9048	0.8800	0.8837	0.9371
Random Forest	0.8913	0.8333	0.9524	0.8400	0.8889	0.9200
XGBoost	0.8478	0.7917	0.9048	0.8000	0.8444	0.9181
LightGBM	0.7826	0.7200	0.8571	0.7200	0.7826	0.9010
PNN	0.8913	0.9000	0.8571	0.9200	0.8780	0.9124

Baseline models were evaluated under the same experimental setting before applying PSO-based tuning to the PNN.

3.2. PSO-PNN Hybrid Model Results

The particle swarm optimization (PSO) procedure applied to the PNN smoothing parameter identified a best-found value of $\sigma^* = 1.7714$ on the UCI Heart Disease dataset. The definitive evaluation metrics resulting from this PSO-tuned configuration are summarized in Table 6. By implementing the PSO algorithm, the hybrid model missed only one patient compared to three with the standard PNN. This is an a promising result from a sensitivity perspective because a false negative where a sick patient is misclassified as healthy is extremely dangerous in medical diagnostics.

Table 6. PSO-PNN results on UCI Heart Disease test data.

Metric	PSO-PNN Value
Accuracy	0.9130
Precision	0.8696
Recall	0.9524
Specificity	0.8800
F1-Score	0.9091
ROC-AUC	0.9486
Best-Found σ	1.7714

The results reported in Table 6 correspond to the best-found PSO-PNN configuration used for the final model comparison, whereas the following stability analysis summarizes the variability of PSO-PNN performance across repeated independent PSO executions.

To examine the robustness of the PSO-based tuning procedure, 30 independent PSO runs were performed using different random seeds. This analysis was conducted to evaluate whether the obtained performance was dependent on a single stochastic execution of PSO.

As shown in Table 7, the best-found σ values varied across runs, with a mean of 1.6327 and a standard deviation of 0.2904. Despite this variation, the predictive performance remained stable. The recall remained constant at 0.9524 across all runs, while the ROC-AUC achieved a mean of 0.9413 with a standard deviation of 0.0149. The mean accuracy was 0.9043, with a standard deviation of 0.0177. These results indicate that the performance of PSO-PNN was not dependent on a single stochastic execution of PSO. As a reference, a fine grid search over the same σ interval was also examined and achieved the same

validation accuracy of 0.8222. The repeated PSO runs maintained stable test performance, with a mean accuracy of 0.9043 and a mean ROC-AUC of 0.9413.

Table 7. Stability analysis of PSO-PNN across 30 independent PSO runs.

Metric	Mean	Std	Min	Max
Best-Found σ	1.6327	0.2904	1.0573	1.8715
Validation Accuracy	0.8222	0.0000	0.8222	0.8222
Accuracy	0.9043	0.0177	0.8696	0.9130
Precision	0.8557	0.0283	0.8000	0.8696
Recall	0.9524	0.0000	0.9524	0.9524
Specificity	0.8640	0.0325	0.8000	0.8800
F1-Score	0.9012	0.0161	0.8696	0.9091
ROC-AUC	0.9413	0.0149	0.9105	0.9486

In Figure 6 the corresponding confusion matrices for the baseline PNN and PSO-PNN are presented. The PSO-PNN model shows improved sensitivity, where the false negatives were reduced from 3 to 1, while the false positives remained nearly the same. The corresponding improvement in recall is discussed later in Section 3.4. By implementing the PSO algorithm, the hybrid model missed only one patient compared to three with the standard PNN. This is an excellent clinical result because a false negative, where a patient with cardiovascular disease is misclassified as healthy, is one of the most dangerous errors in medical diagnostics.

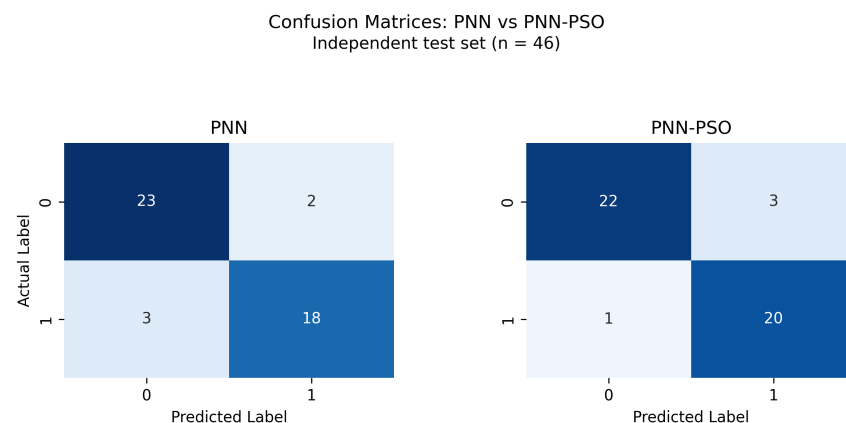


Figure 6. Confusion matrices of the baseline PNN and optimized PSO-PNN models evaluated on the independent test set ($n = 46$) of the UCI Heart Disease dataset.

3.3. Comprehensive Model Comparison on UCI Dataset

Table 8 presents the performance comparison of all models on the UCI Heart Disease dataset, namely, KNN, Logistic Regression, the SVM (RBF), Random Forest, XGBoost, LightGBM, the baseline PNN, and the proposed PSO-PNN. The proposed PSO-PNN model achieved the highest accuracy, F1-score, and ROC-AUC, while also attaining the top recall value among the evaluated models. Meanwhile, the baseline PNN achieved the highest precision and specificity. Figures 7–9 show the model rankings based on accuracy, ROC-AUC, and overall metric performance.

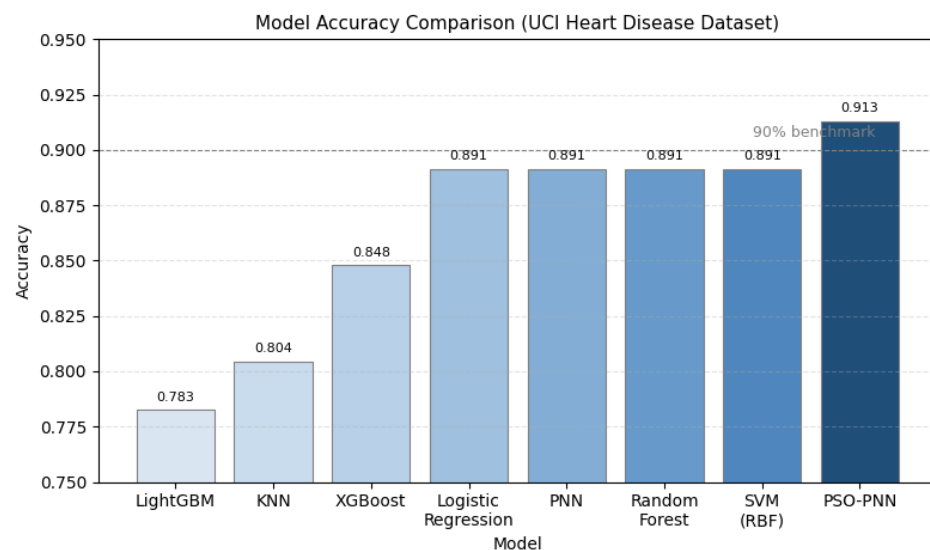


Figure 7. Accuracy comparison of tested models on the UCI Heart Disease dataset, highlighting the competitive performance of the proposed PSO-PNN.

Table 8. Comprehensive performance comparison of all models on the UCI Heart Disease dataset (test split $n = 46$). Best values in each column are highlighted in bold.

Model	Accuracy	Precision	Recall	Specificity	F1-Score	ROC-AUC
KNN	0.8043	0.7500	0.8571	0.7600	0.8000	0.8476
Logistic Regression	0.8913	0.8333	0.9524	0.8400	0.8889	0.9371
SVM (RBF)	0.8913	0.8636	0.9048	0.8800	0.8837	0.9371
Random Forest	0.8913	0.8333	0.9524	0.8400	0.8889	0.9200
XGBoost	0.8478	0.7917	0.9048	0.8000	0.8444	0.9181
LightGBM	0.7826	0.7200	0.8571	0.7200	0.7826	0.9010
PNN (Baseline)	0.8913	0.9000	0.8571	0.9200	0.8780	0.9124
PSO-PNN (Proposed)	0.9130	0.8696	0.9524	0.8800	0.9091	0.9486

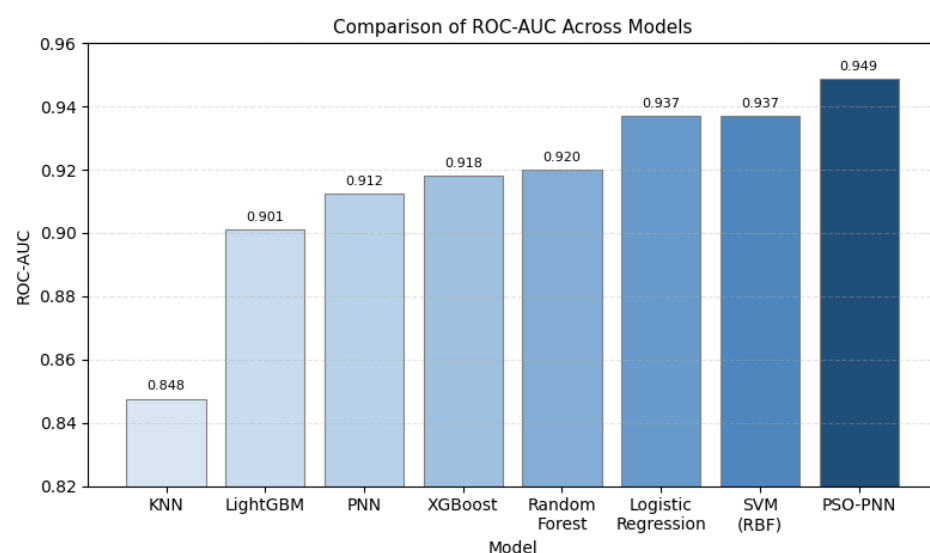


Figure 8. ROC-AUC comparison of all models on the UCI Heart Disease dataset, showing the highest score achieved by PSO-PNN.

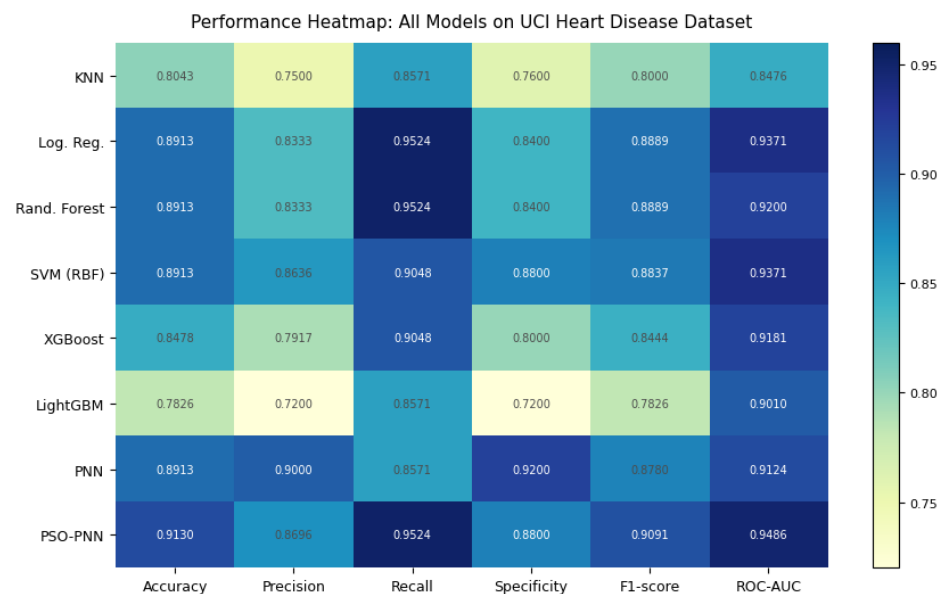


Figure 9. Performance heatmap of all models across six evaluation metrics on the UCI Heart Disease dataset, highlighting the overall performance of PSO-PNN.

3.4. Impact of PSO Optimization: PNN vs. PSO-PNN

To quantify the effect of PSO-based σ tuning of the smoothing parameter σ on classification performance, the proposed PSO-PNN model was compared against the standard PNN, as reported in Table 9. Figure 10 illustrates the comparative performance trends of both models across the evaluation metrics, highlighting consistent improvements achieved by the PSO-tuned variant.

Table 9. Improvement using PSO-based tuning of the smoothing parameter σ on the UCI Heart Disease test set.

Metric	Baseline PNN	PSO-PNN	Gain (Δ)
Accuracy	0.8913	0.9130	+0.0217
Precision	0.9000	0.8696	−0.0304
Recall	0.8571	0.9524	+0.0953
Specificity	0.9200	0.8800	−0.0400
F1-score	0.8780	0.9091	+0.0311
ROC-AUC	0.9124	0.9486	+0.0362

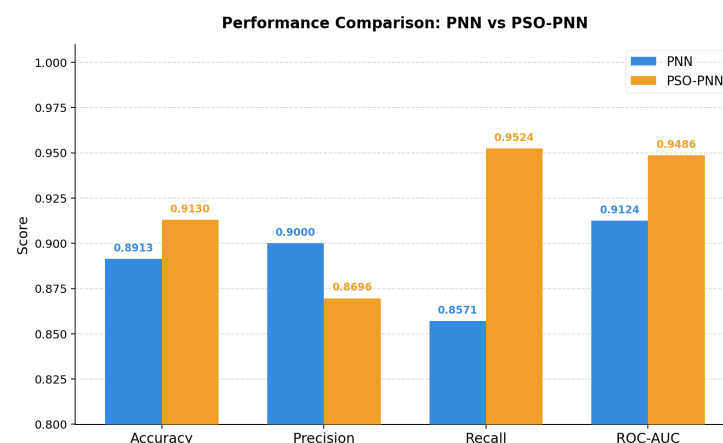


Figure 10. Comparison of PNN and PSO-PNN performance across accuracy, precision, recall, and ROC-AUC metrics.

3.5. Performance Evaluation and Model Robustness Analysis

Model performance was evaluated using two complementary settings. First, the baseline models and the proposed PSO-PNN were evaluated on the held-out test set using the same held-out split. Then, stratified 10-fold cross-validation was conducted for all models to examine performance consistency across different data partitions, as reported in Table 10.

Table 10. Comparative performance of models using stratified 10-fold cross-validation on the UCI Heart Disease dataset. Values are reported as mean $\pm$ standard deviation.

Model	Accuracy	Precision	Recall	Specificity	F1-Score	ROC-AUC
Logistic Regression	0.8381 $\pm$ 0.0602	0.8665 $\pm$ 0.0990	0.7764 $\pm$ 0.0735	0.8897 $\pm$ 0.0925	0.8154 $\pm$ 0.0654	0.9019 $\pm$ 0.0447
KNN	0.8049 $\pm$ 0.0664	0.8280 $\pm$ 0.1041	0.7484 $\pm$ 0.1073	0.8533 $\pm$ 0.1151	0.7780 $\pm$ 0.0742	0.8470 $\pm$ 0.0565
SVM (RBF)	0.8411 $\pm$ 0.0524	0.8730 $\pm$ 0.0930	0.7764 $\pm$ 0.0654	0.8956 $\pm$ 0.0885	0.8184 $\pm$ 0.0561	0.9017 $\pm$ 0.0340
Random Forest	0.8049 $\pm$ 0.0585	0.8152 $\pm$ 0.0893	0.7549 $\pm$ 0.1087	0.8478 $\pm$ 0.0831	0.7785 $\pm$ 0.0711	0.9025 $\pm$ 0.0279
XGBoost	0.8114 $\pm$ 0.0557	0.8315 $\pm$ 0.0792	0.7484 $\pm$ 0.0901	0.8654 $\pm$ 0.0768	0.7839 $\pm$ 0.0648	0.8888 $\pm$ 0.0314
LightGBM	0.7854 $\pm$ 0.0397	0.8031 $\pm$ 0.0770	0.7198 $\pm$ 0.0908	0.8423 $\pm$ 0.0806	0.7533 $\pm$ 0.0502	0.8902 $\pm$ 0.0266
PNN (Baseline)	0.8113 $\pm$ 0.0657	0.8407 $\pm$ 0.0958	0.7407 $\pm$ 0.1135	0.8717 $\pm$ 0.0950	0.7806 $\pm$ 0.0805	0.8638 $\pm$ 0.0467
PSO-PNN	0.8445 $\pm$ 0.0456	0.8851 $\pm$ 0.0853	0.7692 $\pm$ 0.0760	0.9081 $\pm$ 0.0795	0.8193 $\pm$ 0.0521	0.8974 $\pm$ 0.0403

As shown in Table 10, PSO-PNN achieved the highest mean accuracy, precision, specificity, and F1-score among the evaluated models in stratified 10-fold cross-validation. Its ROC-AUC remained competitive with the strongest baseline models, although Logistic Regression, the SVM, and Random Forest achieved slightly higher mean ROC-AUC values. These results provide additional support that PSO-PNN maintains competitive performance across different data partitions. The higher performance observed on the holdout test set (91.30% accuracy and 95.24% recall) compared with the mean results from stratified 10-fold cross-validation (84.45% accuracy and 76.92% recall) can be attributed to the different evaluation strategies. The holdout test reflects performance on a single train–test split, whereas stratified 10-fold cross-validation evaluates the model over multiple data partitions, providing a more conservative estimate of its generalization performance.

3.6. Computational Efficiency and Scalability

The computational efficiency of the proposed PSO-PNN was evaluated in terms of optimization time, prediction time, model memory, and scalability with increasing training samples. Across 30 independent runs, PSO required an average of 3.87 ± 1.69 s to optimize the PNN smoothing parameter. The average prediction time was 0.0949 ± 0.0074 ms per patient, and the approximate memory required to store the PNN training patterns, labels, and optimized smoothing parameter was 23.20 KB.

As shown in Table 11, the model memory increased gradually with the number of training samples, while the prediction time remained below 0.12 ms per patient. The PSO optimization time also remained low across the evaluated training sizes. Overall, the proposed framework maintained low computational requirements as the number of training samples increased. The consistently short prediction time and low memory requirement demonstrate that PSO-PNN is computationally efficient and suitable for rapid prediction in settings with limited computational resources.

Table 11. Computational scalability of the proposed PSO-PNN.

Training Samples	PSO Time (s)	Prediction Time (ms)	Model Memory (KB)
53	2.56	0.0699	5.80
106	2.88	0.1062	11.60
159	2.64	0.0841	17.40
212	3.19	0.1148	23.20

The measurements were obtained using Python 3.12.13 in a Linux environment with an Intel Xeon 2.20 GHz processor, two logical CPU cores, and 12.67 GB of RAM.

4. Discussion

The results from the study are presented in this section and discussed in relation to the proposed framework. The contributions of the work and implications of the results are also presented. These are considered from a theoretical, methodological and practical perspective and are specific to a machine learning approach to cardiovascular disease prediction.

4.1. Interpretation of Key Results

The experimental findings indicate that PSO-based tuning of the probabilistic neural network (PNN) leads to improvements in predictive performance and sensitivity-related evaluation on the UCI Heart Disease dataset. The proposed PSO-PNN model achieves a test accuracy of 91.30%, representing the highest performance among all evaluated methods. Given the limited test set size, the reported improvement is interpreted as a promising performance gain with the adopted experimental setting rather than evidence of a statistically significant performance advantage. This result demonstrates competitive performance compared with the evaluated baseline classifiers, including Random Forest, Logistic Regression, and the SVM with the RBF kernel, which achieve accuracies of 89.13%, 88.95%, and 89.01%, respectively. Given that Random Forest and the SVM are widely regarded as strong and robust classifiers for structured medical data, the observed improvement highlights the effectiveness of PSO-based hyperparameter tuning in enhancing the performance of probabilistic models such as the PNN.

The most important result of this study is the improvement in the measure of recall from 85.71% in the baseline PNN to 95.24% in PSO-PNN (an absolute improvement of 9.53%). Recall is a very important measure in cardiovascular disease screening as it measures the ability of a screening test to correctly identify patients with the disease. Missing a case of cardiovascular disease is one of the most serious errors in screening as it may result in delay of treatment for a potentially life threatening condition [30].

From Table 9 we observe that the most important improvement by PSO-PNN is in recall (+9.53 pp), followed by ROC-AUC (+3.62 pp), F1-score (+3.11 pp), and accuracy (+2.17 pp). Despite this trade-off, the number of true negatives remained largely stable, maintaining a specificity of 88.00%. Consequently, PSO optimization achieved a clinically meaningful improvement by increasing recall from 85.71% to 95.24% (a gain of 9.53 percentage points) while accepting only a modest reduction in Precision, resulting in a better balance between sensitivity and specificity.

The 10-fold cross-validation results provide additional support for the observed improvement. As shown in Table 10, PSO-PNN achieved the highest mean accuracy, precision, specificity, and F1-score among the evaluated models. Compared with the baseline PNN, PSO-PNN also produced lower standard deviation values across the main evaluation metrics, which indicates more consistent performance across different data partitions. While Table 9 reports the improvement on the held-out test set, Table 10 provides an additional evaluation using stratified 10-fold cross-validation. These results further support the benefit of PSO-based tuning for improving the performance consistency of the PNN model.

To further support these findings, SHAP (SHapley Additive exPlanations) was applied to explain the contribution of each feature to the model's predictions. As shown in Figure 11, the two most influential features are *tha1* and *cp*, since for high values of these features the model predicts the presence of cardiovascular disease. Also, it can be noticed that the features *exang* and *ca* present large values as well, meaning that they have also

a considerable contribution into the model's decisions. Moreover, as it can be observed from the figure, low values of *thalach* positively contribute to the prediction of the model regarding the presence of cardiovascular disease, which is consistent with clinical understanding. Finally, features such as *oldpeak*, *slope* and *age* seem to influence the model's output with values of moderate size, and therefore they showed a smaller influence on the model output. Thus, the SHAP analysis indicates that the model predictions are influenced by clinically relevant features, which improves the interpretability of the proposed model.

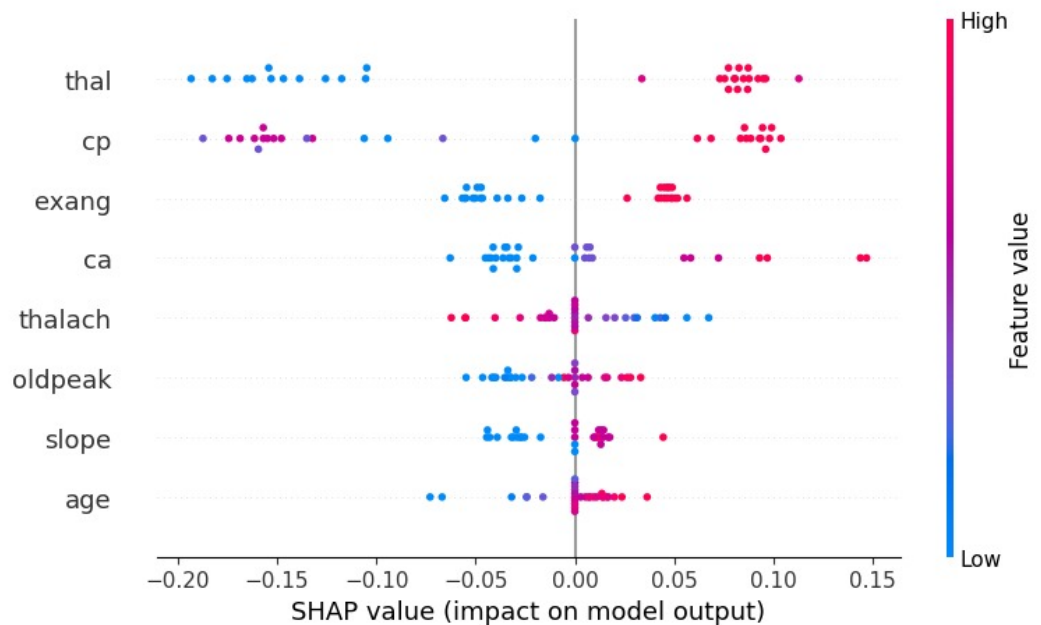


Figure 11. SHAP beeswarm plot showing the impact of each feature on the PSO-PNN model output.

4.2. Role of PSO in Parameter Optimization

Convergence of the PSO minimization of the fitness function occurs smoothly and monotonously over 30 iterations, with quick convergence to an optimum reached by iteration 15, as shown in Figure 4. The best-found value of σ for the UCI dataset was approximately 1.7714, obtained through a continuous PSO-based search over the predefined range, given the smooth but non-trivial shape of the performance over the parameter σ . This illustrates the advantage of PSO-based tuning over manual trial-and-error selection of the smoothing parameter.

The consistency between optimization method (validation-based optimization during PSO) and performance on the test set suggests that PSO successfully found a generalizable parameter that maintained competitive performance on unseen test data. Additionally, the strong consistency between improvements in accuracy and in ROC-AUC values suggests that there is a more cautious statement about improved sensitivity and competitive overall performance.

4.3. Comparison with the State of the Art

The results achieved by the proposed PSO-PNN model, specifically an accuracy of 91.30% and an ROC-AUC of 94.86%, demonstrate competitive performance compared to existing studies reported in the literature. The best accuracy achieved by Stonier et al. [31] was 88.52%. Other studies by Bhatt et al. [32] on the UCI dataset with ensemble methods also reported competitive results. More recent studies on the Cleveland Heart Disease dataset have reported predictive accuracies exceeding 90% using different machine learning strategies. Chandrasekhar and Peddakrishna [24] achieved an accuracy of 93.44% using a soft voting ensemble classifier optimized with GridSearchCV and five-fold cross-validation,

whereas Shrestha [33] and Mushtaq et al. [34] reported competitive results using XGBoost-based and explainable machine learning frameworks, respectively. These studies show that predictive accuracies above 90% have been reported on the Cleveland Heart Disease dataset with different experimental settings.

Compared with these studies, the proposed PSO-PNN framework achieves competitive predictive performance while combining mutual information-based feature selection, PSO-based parameter optimization, and SHAP-based explainability within a probabilistic neural network. Nevertheless, direct numerical comparisons across studies should be interpreted with caution because differences in preprocessing strategies, feature selection methods, and validation protocols may influence the reported performance.

4.4. Limitations and Future Directions

Future work will explore the application of the PSO-PNN framework on much larger multicenter clinical datasets to evaluate its generalizability. Moreover, as discussed in Section 1, the current simple formulation for the PSO tunes only the smoothing parameter σ . It is possible that further improvements in performance may be realized by extending the current formulation to a multi-objective tuning problem in which multiple model parameters, as well as relevant features, are simultaneously optimized. Future work may also compare MI with other feature selection methods, such as Recursive Feature Elimination (RFE) and the Least Absolute Shrinkage and Selection Operator (LASSO) [35]. This comparison may help determine whether different feature subsets can further improve predictive performance or reduce the number of input features. Although SHAP improves the interpretability of the proposed framework by providing explanations for individual predictions, future work may also investigate complementary explainability techniques to provide additional perspectives on model predictions [36].

In addition, several practical challenges remain before clinical deployment. In particular, the memory requirements of the PNN increase with the size of the training dataset, which may affect scalability in large-scale clinical environments. Efficient integration into real-time clinical decision support systems also requires careful optimization to ensure low inference latency. Furthermore, additional validation using larger and multi-institutional datasets is necessary to evaluate the robustness and generalizability of the proposed framework under diverse clinical conditions and before practical clinical use can be considered, as recommended for clinical prediction models [37]. The proposed pipeline also has the potential to be applied to a variety of other medical prediction tasks, including diabetes, sepsis, and cancer risk prediction. From a broader healthcare perspective, cardiovascular prediction models may also be integrated into patient monitoring and clinical decision support systems [16]. They may complement medical robotic systems by providing patient risk information during care [17]. Future work may investigate this integration using external clinical data and real-time evaluation.

5. Conclusions

This study designed a hybrid framework specifically to improve the predictive performance of cardiovascular disease prediction. We combined the particle swarm optimization (PSO) algorithm with probabilistic neural networks (PNNs). The proposed model addresses the sensitivity of PNN performance to the smoothing parameter through PSO-based tuning. The main contribution of the proposed PSO-PNN model lies in the dynamic optimization of the smoothing parameter using PSO, which improved predictive performance and recall with the adopted experimental setting when handling nonlinear and complex cardiac data. Experimental results showed that the proposed model achieved an accuracy of 91.3%, an ROC-AUC of 94.8%, and a recall of 95.24%. To further support these findings, the mod-

els were also evaluated using stratified 10-fold cross-validation. Under this evaluation setting, the proposed PSO-PNN achieved a mean accuracy of 84.45% and a mean recall of 76.92%. Although these values are lower than the held-out test results, they provide a more conservative assessment of the model performance across different data partitions. The 10-fold cross-validation results further showed that PSO-PNN maintained competitive mean performance across accuracy, precision, specificity, and F1-score relative to the evaluated baseline models. Taken together, the results from both evaluation settings indicate that the proposed PSO-PNN offers highly competitive performance relative to the evaluated baseline models. In particular, the improvement in recall suggests that the model was able to detect more positive heart disease cases while reducing missed diagnoses.

The study also considered model interpretability through the use of post hoc SHAP-based interpretability analysis. SHAP was used to explain the prediction results and to identify the contribution of each feature to the final prediction. It also provided a clearer view of how the model reached its decisions.

Future studies may also explore the use of other types of medical data and examine the contribution of the model to the prediction of the results.

Author Contributions: Conceptualization, J.B.S. and N.A.; methodology, N.A.; software, N.A.; validation, N.A., M.A., K.A. and A.A.; formal analysis, N.A.; investigation, N.A.; data curation, N.A., M.A.; writing—original draft preparation, N.A., M.A., K.A. and A.A.; writing—review and editing, N.A. and J.B.S.; visualization, N.A.; supervision, J.B.S. All authors have read and agreed to the published version of the manuscript.

Funding: The authors extend their appreciation to the Deanship of Scientific Research at Northern Border University, Arar, KSA, for funding this research work through the project number “NBU-FFRGS-2026”.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The UCI Heart Disease dataset is publicly available at the UCI Machine Learning Repository: <https://archive.ics.uci.edu/ml/datasets/heart+disease> (accessed on 15 November 2024). The Python implementation is available from the corresponding author upon reasonable request.

Acknowledgments: We acknowledge the developers and maintainers of the public dataset referenced in this work.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CVD	Cardiovascular Disease
ML	Machine Learning
PNN	Probabilistic Neural Network
PSO	Particle Swarm Optimization
AI	Artificial Intelligence
KNN	K-Nearest Neighbors
RF	Random Forest
SVM	Support Vector Machine
RBF	Radial Basis Function

LR	Logistic Regression
MI	Mutual Information
ROC	Receiver Operating Characteristic
AUC	Area Under the Curve
ROC-AUC	Area Under the Receiver Operating Characteristic Curve
UCI	University of California, Irvine
XAI	Explainable Artificial Intelligence
SHAP	SHapley Additive Explanations
LIME	Local Interpretable Model-Agnostic Explanations
ANN	Artificial Neural Network
DT	Decision Tree
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
pp	Percentage Points

References

1. World Health Organization. Cardiovascular Diseases (CVDs). 2025. Available online: <https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-%28cvds%29> (accessed on 16 July 2026).
2. Martin, S.S.; Aday, A.W.; Almarzooq, Z.I.; Anderson, C.A.M.; Arora, P.; Avery, C.L.; Baker-Smith, C.M.; Barone Gibbs, B.; Beaton, A.Z.; Boehme, A.K.; et al. 2025 Heart Disease and Stroke Statistics: A Report of US and Global Data From the American Heart Association. *Circulation* **2025**, *151*, e41–e660. [CrossRef] [PubMed]
3. Almutairi, M.; Dardouri, S. Intelligent hybrid modeling for heart disease prediction. *Information* **2025**, *16*, 869. [CrossRef]
4. Sharifi, S.; Donyadadi, A. Detection and diagnosis of congenital heart disease from chest X-rays with deep learning models. *Int. J. Appl. Data Sci. Eng. Health* **2025**, *1*, 1–9.
5. Hao, K.; Williams, C.A.; Barker, A.R.; Wadey, C.A.; Müller, J.; Oberhoffer, R.; Willinger, L. Predictive value of musculoskeletal fitness for cardiovascular risk factors in adolescents with congenital heart disease: A cross-sectional study. *J. Clin. Med.* **2026**, *15*, 2863. [CrossRef] [PubMed]
6. Gul, G.; Korejo, I.A.; Hakro, D.N.; Alqahtani, H.; Abbasi, A.; Babar, M.; Al Rahbi, O.; Ali, N.I. Machine learning and ensemble methods for cardiovascular disease prediction: A systematic review of approaches, performance trends, and research challenges. *Computers* **2026**, *15*, 25. [CrossRef]
7. Naser, M.A.; Majeed, A.A.; Alsabah, M.; Al-Shaikhli, T.R.; Kaky, K.M. A review of machine learning's role in cardiovascular disease prediction: Recent advances and future challenges. *Algorithms* **2024**, *17*, 78. [CrossRef]
8. Naskinova, I.; Kolev, M.; Karova, D.; Milev, M. Machine learning-based blood pressure prediction using cardiovascular disease data: A comprehensive comparative study. *Electronics* **2026**, *15*, 312. [CrossRef]
9. Dritsas, E.; Trigka, M. Efficient Data-Driven Machine Learning Models for Cardiovascular Diseases Risk Prediction. *Sensors* **2023**, *23*, 1161. [CrossRef] [PubMed]
10. Villoth, S.J.; Dabic, P.; Zivkovic, T.; Zivkovic, M.; Andjelic, S.; Mravik, M.; Simic, V.; Abdel-Salam, M.; Bacanin, N. Cardiovascular disease risk prediction utilizing two-tier classification framework optimized with adapted variable neighborhood search algorithm. *Algorithms* **2026**, *19*, 130. [CrossRef]
11. Tompra, K.V.; Papageorgiou, G.; Tjortjis, C. Strategic machine learning optimization for cardiovascular disease prediction and high-risk patient identification. *Algorithms* **2024**, *17*, 178. [CrossRef]
12. Ran, G.; Wang, Y.; Zhang, H.; Cheng, J.; Lai, D. Balancing energy consumption and detection accuracy in cardiovascular disease diagnosis: A spiking neural network-based approach with ECG and PCG signals. *Sensors* **2025**, *25*, 5263. [CrossRef] [PubMed]
13. Liu, F.; Fan, Z.; Hu, W.; Xu, D.; Peng, M.; He, J.; He, Y. Vision Transformer-based overlay processor for edge computing. *Appl. Soft Comput.* **2024**, *156*, 111421. [CrossRef]
14. Baklola, M.; Elmahdi, R.R.; Ali, S.; Elshenawy, M.; Mossad, A.M.; Al-Bawah, N.; Mansour, R.M. Artificial intelligence in disease diagnostics: A comprehensive narrative review of current advances, applications, and future challenges in healthcare. *Ann. Med. Surg.* **2025**, *87*, 4237–4245. [CrossRef] [PubMed]
15. Zheng, D.; Chen, F.; Li, P.; Peng, S.Y. Advanced signal processing for cardiovascular and neurological diseases. *Comput. Math. Methods Med.* **2018**, *2018*, 3416540. [CrossRef] [PubMed]
16. Riaz, W.; Ji, J.; Ullah, A. TriViT-Lite: A Compact Vision Transformer–MobileNet Model with Texture-Aware Attention for Real-Time Facial Emotion Recognition in Healthcare. *Electronics* **2025**, *14*, 3256. [CrossRef]

17. Riaz, W.; Ullah, A.; Ji, J. Multi-Scale Attention Networks with Feature Refinement for Medical Item Classification in Intelligent Healthcare Systems. *Sensors* **2025**, *25*, 5305. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Specht, D.F. Probabilistic neural networks. *Neural Netw.* **1990**, *3*, 109–118. [\[CrossRef\]](#)
19. Ullah, F.; Chen, X.; Rajab, K.; Al Reshan, M.S.; Shaikh, A.; Hassan, M.A.; Rizwan, M.; Davidekova, M. An Efficient Machine Learning Model Based on Improved Features Selections for Early and Accurate Heart Disease Predication. *Comput. Intell. Neurosci.* **2022**, *2022*, 1906466. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Myśliwiec, A.; Bartusik-Aebisher, D.; Xavierselvan, M.; Paul, A.; Aebisher, D. Deep learning and cardiovascular diseases: An updated narrative review. *J. Clin. Med.* **2026**, *15*, 3053. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Reddy, S.R.; Murthy, G.V. Cardiovascular Disease Prediction Using Particle Swarm Optimization and Neural Network Based an Integrated Framework. *SN Comput. Sci.* **2025**, *6*, 186. [\[CrossRef\]](#)
22. Waqar, M.; Shah Nawaz, M.B.; Saleem, S.; Dawood, H.; Muhammad, U.; Dawood, H. Enhancing heart attack prediction: Feature identification from multiparametric cardiac data using explainable AI. *Algorithms* **2025**, *18*, 333. [\[CrossRef\]](#)
23. Gupta, S.; Gupta, M.; Kumar, R.; Singh, J.; Sharama, M. A Hybrid Machine Learning and Deep Learning Framework for Heart Disease Prediction. In Proceedings of the 2026 World Conference on Computational Science and Technology (WcCST), Gharuan, India, 26–27 March 2026; pp. 23–28. [\[CrossRef\]](#)
24. Chandrasekhar, N.; Peddakrishna, S. Enhancing Heart Disease Prediction Accuracy through Machine Learning Techniques and Optimization. *Processes* **2023**, *11*, 1210. [\[CrossRef\]](#)
25. Asuncion, A.; Newman, D. *UCI Machine Learning Repository*; University of California: Irvine, CA, USA, 2007. Available online: <https://archive.ics.uci.edu/ml/datasets/heart+disease> (accessed on 15 November 2024).
26. Kennedy, J.; Eberhart, R. Particle swarm optimization. In Proceedings of the ICNN'95—International Conference on Neural Networks, Perth, Australia, 27 November–1 December 1995; Volume 4, pp. 1942–1948.
27. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [\[CrossRef\]](#)
28. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794. [\[CrossRef\]](#)
29. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.Y. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In Proceedings of the Advances in Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; Volume 30.
30. Saqib, M.; Perswani, P.; Muneem, A.; Mumtaz, H.; Neha, F.; Ali, S.; Tabassum, S. Machine learning in heart failure diagnosis, prediction, and prognosis. *Ann. Med. Surg.* **2024**, *86*, 3615–3623. [\[CrossRef\]](#) [\[PubMed\]](#)
31. Stonier, A.A.; Gorantla, R.K.; Manoj, K. Cardiac disease risk prediction using machine learning algorithms. *Healthc. Technol. Lett.* **2024**, *11*, 213–217. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Bhatt, C.M.; Patel, P.; Ghetia, T.; Mazzeo, P.L. Effective heart disease prediction using machine learning techniques. *Algorithms* **2023**, *16*, 88. [\[CrossRef\]](#)
33. Shrestha, D. Advanced Machine Learning Techniques for Predicting Heart Disease: A Comparative Analysis Using the Cleveland Heart Disease Dataset. *Appl. Med. Inform.* **2024**, *46*, 91–102.
34. Mushtaq, M.; Mehmood, Z.; Butt, A.A.; Shabir, M.Y.; Khan, M.N.U. Causal and Explainable Machine Learning Framework for Heart Disease Prediction using XGBoost and SHAP. *J. Comput. Biomed. Inform.* **2025**, *10*, 1–12. [\[CrossRef\]](#)
35. Mamun, M.M.R.K.; Elfouly, T. Detection of Cardiovascular Disease from Clinical Parameters Using a One-Dimensional Convolutional Neural Network. *Bioengineering* **2023**, *10*, 796. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Rezk, N.G.; Alshathri, S.; Sayed, A.; El-Din Hemdan, E.; El-Behery, H. XAI-augmented voting ensemble models for heart disease prediction: A SHAP and LIME-based approach. *Bioengineering* **2024**, *11*, 1016. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Riley, R.D.; Archer, L.; Snell, K.I.E.; Ensor, J.; Dhiman, P.; Martin, G.P.; Bonnett, L.J.; Collins, G.S. Evaluation of clinical prediction models (part 2): How to undertake an external validation study. *BMJ* **2024**, *384*, e074820. [\[CrossRef\]](#) [\[PubMed\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.