

Article

XAI Framework for Cardiovascular Disease Prediction Using Classification Techniques

Pratiyush Guleria ¹, Parvathaneni Naga Srinivasu ², Shakeel Ahmed ^{3,*}, Naif Almusallam ³ and Fawaz Khaled Alarfaj ³

¹ National Institute of Electronics and Information Technology, Shimla 171001, India

² Department of Computer Science and Engineering, Prasad V Potluri Siddhartha Institute of Technology, Vijayawada 520007, India

³ Department of Computer Science, College of Computer Sciences and Information Technology, King Faisal University, Al-Ahsa 31982, Saudi Arabia

* Correspondence: shakeel@kfu.edu.sa

Abstract: Machine intelligence models are robust in classifying the datasets for data analytics and for predicting the insights that would assist in making clinical decisions. The models would assist in the disease prognosis and preliminary disease investigation, which is crucial for effective treatment. There is a massive demand for the interpretability and explainability of decision models in the present day. The models' trustworthiness can be attained through deploying the ensemble classification models in the eXplainable Artificial Intelligence (XAI) framework. In the current study, the role of ensemble classifiers over the XAI framework for predicting heart disease from the cardiovascular datasets is carried out. There are 303 instances and 14 attributes in the cardiovascular dataset taken for the proposed work. The attribute characteristics in the dataset are categorical, integer, and real type and the associated task related to the dataset is classification. The classification techniques, such as the support vector machine (SVM), AdaBoost, K-nearest neighbor (KNN), bagging, logistic regression (LR), and naive Bayes, are considered for classification purposes. The experimental outcome of each of those algorithms is compared to each other and with the conventional way of implementing the classification models. The efficiency of the XAI-based classification models is reasonably fair, compared to the other state-of-the-art models, which are assessed using the various evaluation metrics, such as area under curve (AUC), receiver operating characteristic (ROC), sensitivity, specificity, and the F1-score. The performances of the XAI-driven SVM, LR, and naive Bayes are robust, with an accuracy of 89%, which is assumed to be reasonably fair, compared to the existing models.



Citation: Guleria, P.; Naga Srinivasu, P.; Ahmed, S.; Almusallam, N.; Alarfaj, F.K. XAI Framework for Cardiovascular Disease Prediction Using Classification Techniques. *Electronics* **2022**, *11*, 4086. <https://doi.org/10.3390/electronics11244086>

Academic Editor:

Senthilkumar Mohan

Received: 14 November 2022

Accepted: 6 December 2022

Published: 8 December 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Machine learning (ML) is a subfield of artificial intelligence (AI), followed by deep learning techniques. The analytical model will rely on the training data for making future predictions. The model extracts helpful information, based on previous learning, when provided with new data. Machine learning and AI are emerging technologies and play leading roles in healthcare and personalized clinical support. In healthcare, the clinical data consists of electronic health data and sensor data from the Internet of Things (IoT) devices. Due to the data availability in unstructured and structured forms, it becomes difficult for humans to derive meaningful and decision-making information [1]. With the help of IoT and ML techniques, medical aid can easily reach people in remote areas and those who want preliminary medical assistance. ML and AI are already working on image-processing techniques related to diseases and prognoses. The cloud-enabled technologies collaborate closely with IoT technologies and provide medical aid, and maintaining patients' electronic health records [2]. Apart from that, mobile edge computing turns up as an engaging

criterion to provide time-sensitive computing services for the industrial Internet of Things (IIoT) [3]. Felkey, B. G. and Fox, B. I. [4] have discussed the utility of mobile apps, which regularly take the status of patient medication behavior and caretakers. It helps maintain patients' electronic medical records and facilitates the effective communication between patients and physicians. IBM Watson systems use AI-enabled technologies in healthcare to extract information related to drug discovery, find cancerous cells, and research areas which show how the immune system can fight cancer. On the Internet of Medical Things platform, AI and ML use deep learning assisted convolutional neural networks for heart disease prediction [5] and algorithms, such as neural networks, which help analyze the voluminous data available to physicians, investigate the diagnostic patterns, streamline them, and integrate the patient's records to avoid mistakes [6]. The other areas where AI and ML technologies in healthcare can be explored include biopharmaceuticals, the digitization of health records, and converting them into the form of electronic medical records, drug discovery, and development. The AI-enabled platforms help in the early detection of the disease, clinical trials, etc. AI provides radiologists with assistance in interpreting the X-ray images and CT scan reports correctly and with more precision [7,8]. According to the [9], 10% of mammography reports cannot predict breast cancer correctly, resulting in further biopsies for a clear diagnosis, which can cause discomfort to patients. Similar problems are faced by patients, in the case of thyroid reports [10]. AI integrated with ML algorithms, and health applications provide significant heart disease-related consultations to rural patients who do not have expert cardiologists. AI expert systems can be devised to provide an expert opinion with the help of knowledge-based systems and digital medical devices [11].

The current study is motivated by the advancements in the AI and IoT technology, that assist in the better livelihood of individuals. AI algorithms examine IoT data from smartwatches, medicines, wearable monitoring, etc. The data helps patients, doctors, and pharmaceuticals evaluate medical conditions, offer feedback on treatments, medication therapy, patient outcomes, etc. [12–14]. AI and ML-assisted mobile applications and smart gadgets enable remote patient–doctor contact. These gadgets give remote clinical help to patients with heart attacks, sleep difficulties, and epilepsy [15]. Researchers may benefit from electronic health data [16]. ML algorithms on big datasets may find data patterns. ML may affect healthcare epidemiology if utilized successfully. Health risk assessments and other predictive healthcare applications use supervised machine learning [17,18]. Model interoperability and data interpretation are healthcare's significant challenges. Distributed, heterogeneous data storage is another difficulty. The effective use of computer resources, EHRs, and data complexity are challenges in healthcare. Massive amounts of organized, semi-structured, and unstructured data present several issues. A non-uniform dataset structure might hinder data scientists and reduce the prediction accuracy.

Along with it, outliers and missing values in the dataset result in the poor performance of ML trained models. There are challenges from wearable devices after health effects, accuracy, and reliability issues of the data [19]. There are also chances of hacking patients' web IoT devices and the inappropriate use of associated data, which cannot be denied. Privacy and ethics related to data, its protection [20], and the reliability of the predictive results obtained from the ML models are major issues that need to be resolved. AI and ML-enabled techniques assist technicians and physicians in personalized healthcare for patients. With the help of AI and ML techniques, one can identify patients at risk, using pattern recognition features. The regular feedback of the end-user obtained using AI agents, helps improve the machine learning models to respond to and solve the queries. The robotic field is already being explored using AI to assist doctors in surgical operations and perform repetitive tasks performed in hospitals and labs, which saves a lot of time. Another major advantage of AI in the healthcare sector is the immediate assistance from medical staff in a virtual mode, saving their time and expenditure by reducing unnecessary visits to the hospital. These emerging fields help streamline health operations. ML techniques applied to electronic healthcare records can predict the disease outbreak and improve the patient's risk of disease [21]. ML-trained models can predict the disease's diagnosis and

prognosis [22]. ML unsupervised learning techniques help in identifying complicated diseases with more precision. Deep learning applications, for example, aid in predicting cardiac arrest from one-dimensional heartbeat signals to a smart diagnosis [23]. AI and ML mobile integrated applications with an infusion of the IoT, help in the remote diagnosis of patients, providing telemedicine services to patients, remotely [24]. There is an automated risk prediction system for clinical care with the help of AI, ML, and IoT-enabled devices. Wearable and connectivity devices with AI assistance, benefit senior citizens and their fitness tracking [25].

The objectives of the current study are presented below in a point-wise manner.

- XAI ML techniques are implemented, showing the model's details and working for the desired predictions. It will help the doctors to decide the predictive results and the diagnoses of heart disease;
- Explainable feature weight initialization and normalization for the uniform distribution of the training dataset are carried out;
- The heart disease classification is set out, and the predictions of the patients suffering from cardiovascular disease are performed by taking the attributes of 'age' and 'sex', and 'cholesterol'. However, more attributes can predict the disease in the future scope of work;
- The statistical values of all the attributes related to heart diseases in a sample dataset are monitored;
- The performance of the XAI-based prediction of cardiovascular disease using the ensemble classifiers, is being assessed over various evaluation metrics, the AUC, accuracy, true positive rate (TPR), recall, and precision.

The current study is organized over six sections. Section 1 presents the introduction to the XAI and ensemble classifiers in healthcare, followed by a literature review in Section 2. The concept of explainable artificial intelligence is discussed in Section 3. Section 4 covers the experimental approach using machine learning techniques. The experimental outcome of the proposed study is discussed in Section 5, and the conclusion with future perspectives is presented in Section 6.

2. Literature Review

AI undoubtedly plays a crucial role in the medical sector when integrated with health-care applications. Still, the results generated from machine learning models need to explain the machine's decisions and justify the predictions' reliability [26]. AI integrated with smart wearable devices shows a lot of potential in converting healthcare to smart healthcare. The inclusion of healthcare applications into wearable devices, smartwatches helps capture, predict, and analyze health data. An explainable AI system is required to enhance the trust and reliability of the results. XAI is necessary because the results generated from black-box ML models have led to a lack of accountability, transparency, and trust. In such a scenario, those techniques need to be developed using the XAI system, where predictions are well explained by AI systems and helpful for medical practitioners in diagnosing patients, with trust and reliability on the ML model predictions [27]. Vincent Paul et al. [28] have proposed an intelligent framework for the prediction of heart disease using deep learning.

The purpose of this study is to forecast cardiac sickness using computational intelligence methods, such as the K-nearest neighbor, random forest (RF), decision tree (DT), and AdaBoost, through the internet and mobile applications. The KNN was considered for its ability to cope with the most accurate prediction models and offers incredibly exact predictions. The distance measurement determines the forecast's accuracy. As a consequence, the KNN technique may be used in situations requiring great precision. The RF is an ensemble learning approach that focuses on the bagging technique. The DT excels with data management and actually works with a linear pattern. It can analyze massive volumes of information within a short amount of time. It creates as many trees as possible on a portion of the data, then it combines all of the trees' results. Boosting, in contrast, minimizes bias rather than variance. Models are weighted in boosting according to their

performance. As a result, boosting is preferred over bagging [29]. Authors have performed the machine learning models for coronary artery disease in Nigeria and the RF model emerged as the best model for performing the predictions, in terms of accuracy followed by the best sensitivity results for the support-vector machine learning model [30]. The RF classifier achieved remarkable results for the AUC value for predicting the coronary artery stenosis in Taiwanese patients with coronary artery disease [31].

AI can classify data and make predictions with the help of ML models. Still, due to the non-availability of a proper explanation of how to input, the output is related to each other, and the predictive results lack accountability, explainability, and transparency. XAI can simply be stated as intending to make AI systems more coherent to humans [32]. The terms transparency, interpretability, and explainability are often used interchangeably. Interpretability is related to how much a model can be understood [33], although it is also used instead of the term explainability [34]. Transparency either refers to a comprehensive distinction for providing collaborators with relevant information about how the model works, this includes the testimony of the training practices, analysis of the training data distribution, the code releases, and trait-level details [35], or an algorithm-specific clarity on how the model works, as opposed to opacity [32,35]. Fellous et al. [36] have discussed the XAI approach followed for neurosciences. In contrast, Dave, D. et al. [37] have deliberated the XAI techniques taking heart disease datasets and intending to create trust in medical practitioners in AI systems in healthcare and on the outcome generated from the black-box model for the diagnosis of patients. The purpose of XAI is to build an AI system that is understandable, accountable, trustable, observable, interpretable, explainable, etc. [38]. The interpretation of the results is essential when the results support decision-making, especially in the healthcare domain and is related to human lives. The machine learning model improvement with a proper explanation of the results is the focus of the XAI domain. Porto, R. et al. [39] have predicted cardiovascular disease using the heart dataset. Their study aims to define a simple mechanism for the straightforward interpretation, based on principles (a) the reduction of attributes in such a way that it does not degrade the performance of the prediction system and (b) the interpretation methods for the final prediction model. Aggarwal et al. implemented XAI techniques for the ECG classification and analysis for the heart disease prediction [40]. XAI may upgrade the analytic outcome by making the prognosis clear to medical specialists [41].

XAI techniques can be grouped by scope into those providing global details of the entire system and those providing local details of a single prognosis. The global explanations ease the understanding of the entire model behavior and hypothesis leading to expected outcomes. For local explanations, the reasons for a single prediction are provided to justify why the model made a specific decision for that instance. Further, these techniques can be grouped by whether they are model agnostic (they can be applied to any ML algorithm), or model specific (they can be only applied to a specific ML algorithm) [42]. The explainable frameworks utilizes the Shapley additive explanation (SHAP) and the local interpretable model-agnostic explanation (LIME) techniques for analyzing each feature weight, for estimating the model's predictive ability [43,44].

ML algorithms offer supportive decision-making for clinical predictions, and with the integration of XAI, patient treatment and diagnostics can be improved. Moreno-Sanchez, P. A. [45] have developed a ML model using ensemble trees ML techniques to predict heart failure survival among patients. In healthcare, AI utilizes computer algorithms to examine complex healthcare data and employs ML to evaluate patients [46]. Peng, J. et al. [47] have used a XAI based framework for interpreting the auxiliary diagnosis of hepatitis with trust in the prediction performance. In most cases, the complex ML models avoid explaining how they come to a decision, resulting in the distrust of clinicians. Therefore, the authors in their study have implemented the XAI framework and selected the transparent black-box ML models, such as the eXtreme Gradient Boosting (XGBoost), SVM, random forests, etc. for forecasting hepatitis deterioration.

These AI models, including deep learning neural networks [48] and others, perform well on various biomedical natural language processing tasks. The authors, in their study, have discussed the post-hoc explanation method of the black-box AI models and found the relationship between the feature values and their outcomes. Similarly, Muddamsetty, S. M. et al. [49] have carried out an expert-level evaluation to understand the results of complex ML models, their accuracy, and, if necessary, their debugging. A XAI Framework helps generate an understandable human explanation of the black-box models [50]. Black-box AI systems that give predictions without any explanation, are uncertain for various reasons, not only because of their lack of clarity, but also because they hide probable biases within the system [51]. The XAI systems are inevitable in deep learning applications, especially in medicine. In the absence of proper explainability transparency in AI systems, the integration of AI with ML in healthcare becomes uncommon and unworthy. Yang, G. et al. [52] surveyed the current XAI progress and its efficacy in the healthcare sector. The authors have also proposed a solution for XAI via the multi-model and multi-center data fusion. All of the models mentioned above lack the capability to explain the decision-making strategy in classifying the instances. It is desired to have a model that could classify the data with an adequate explainability of the outcome of the decision made. Explainability is provided across various aspects of the model, such as the framework comprehensibility and understandability of the system's structure, understanding of the objective functions, ease of understanding of the training dataset, and the parameters in the training set, the comprehensibility of the validation sets, the comprehensibility of the parameters for analysis and the comprehensible model.

3. XAI-Based Classification Model for Predicting Cardiovascular Disease

AI methods help solve complex tasks and achieve a high level of accuracy by learning to address complex computational assignments [53]. To increase the effectiveness of AI, XAI is introduced to enhance the machine learning model's capability. It increases the prediction accuracy of machine learning models and produces explanations about the functioning of ML models, especially in deep learning and neural network models. CNN-based models have been implemented to determine and ascertain childhood pneumonia, using chest X-rays [54]. Deep-learning-based models help in CT scans, chest X-ray image analyses [55], heart disease prediction [56], and the reliable detection of lung diseases from medical images [57]. They can also investigate the predictive factor of RhoB in patients with rectal cancer [58].

In XAI, the audience is the target for the research, and the explainability includes details and reasons. A ML model makes its functioning clear and easy to understand. According to Gunning D. [59], "Explainable AI will develop a set of machine learning algorithms which will allow individual users to comprehend, adequately trust, and manage the next era of artificially intelligent companions". In XAI, the training data has a comprehensible model with an explanation interface, when bringing a new machine Learning process. The explainable model comprehends and handles the input assignment.

In the XAI model, the discussion or recommendations from the ML model have a proper explanation. The output is accepted with trust, with an appropriate explanation of the results of the ML models, mainly the black-box ML models. The black-box ML model follows the explainability approaches for the model, which includes (a) visualization techniques, (b) model explanation, (c) feature selection, (d) feature relevance, (e) model simplification, (f) model explanation by example, (g) text explanations, etc. The XAI model is integrated with the ML model to answer three keywords mainly, Why, When, and How? These keywords mainly focus on the predictions performed by the ML model, trust in the predictions shown by the machine learning model, and finally, how errors in the ML model can be corrected. XAI integration to ML is important, especially when the results generated from the ML model impact human lives or society. Such areas include healthcare, education, the financial sector, weather forecast, medicine, etc. Predicting a patient's disease generated from the ML model needs the XAI model to decide whether to accept the ML model. XAI

helps explain the predictions generated from the ML model, the features selected for the training of the ML algorithms, etc.

The XAI system integration with the ML model can be easily understood from Figure 1a,b for the ML model for predicting cardiovascular disease. From Figure 1a, it is concluded that the doctor is not confident in accepting the results obtained from the ML model, as the doctor seeks an explanation for the results obtained. The doctor is not sure of the predictions of the ML model, and she will trust the results if the workflow and algorithmic techniques adopted by the ML model are explained well.

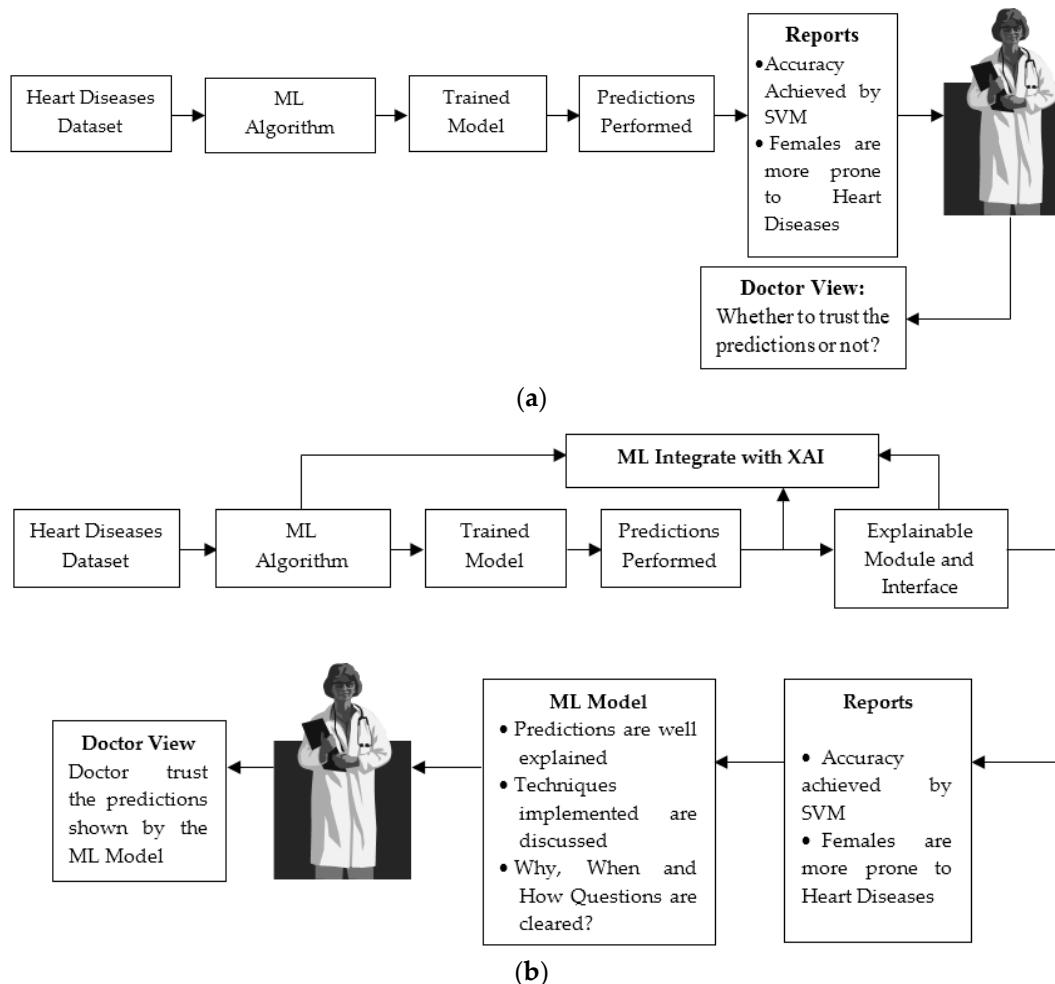


Figure 1. (a) ML approach followed without explainable AI. (b) improvised ML approach followed with explainable AI.

The improvised machine learning model integrated with XAI is shown in Figure 1b. The answers to three keywords, why, when, and how, are delivered using the explainable AI module with the explainable interface. The XAI system integration with the black-box ML models is inevitable to explain the results shown by ML and interpret how it comes to a decision. With proper explainability and transparency, medical practitioners will trust the ML models' predictions. Most of the conclusions by the ML models, especially the deep neural network models, are very complex and not easy to understand. Therefore, the doctors and staff require proper explanations of the predictive decisions, providing clinical support.

3.1. Feature Selection

In the current study on heart disease prediction, the feature selection is a key step in developing a model. When creating a predictive model, the feature selection minimizes the

number of associated parameters in the dataset used, to evaluate the model. Limiting the sum of features to reduce the computational modeling costs is preferable, and in certain situations, it also increases the model's performance. The feature selection varies from the supervised to unsupervised learning models. In supervised learning, the target variable is used, and insignificant predictors are removed. In unsupervised learning, superfluous inputs are removed, and the target variable is not required.

In the proposed work, the numerical features are present. Therefore, the correlation coefficient and the embedded methods are used as filters for the data handling. The correlation coefficient measures the degree of the linear statistical relationship between two numerical variables, whereas, in the embedded method, the feature selection is made during the model training [60]. The predictor variables are chosen to enhance the model's prediction performance. The three types of feature algorithms in MATLAB explored for the proposed study, are mentioned below:

Filter Type: In this selection process, emphasis is placed on the feature qualities, such as the feature variance and relevance to the answer variable. This algorithm selects the features during the preprocessing phase and trains the model.

Wrapper Type: The features are added and removed using the selection criteria in this selection algorithm. The performance of the model is measured by adding or removing the features.

Embedded Type: In this type of feature selection, the algorithm learns the importance of the features during the model's training. The embedded feature selection approach selects the feature considered important during the learning process.

The initial ranking of the features is carried out using the T-test ranking method. Following the comparison, the supervised ranking method is implemented, and the features are ranked by the minimum attainable classification error. The normalization scheme is used, as shown in Equation (1).

$$t = (m_1 - m_2) / \left((\sigma(\text{diff})) / \sqrt{n} \right) \quad (1)$$

m_1, m_2 are the average values of each of the sample datasets. $\sigma(\text{diff})$ is the standard deviation of the differences of the data value. Here, t represents the t-test and n is the sample size of the dataset.

The neighborhood component analysis (NCA) method for the feature extraction is simulated in the MATLAB platform for the classification task in the current study and is used to assist in dealing with the continuous features. In this method, the feature weights are decisively used in a diagonal transformation of the NCA. The feature significance is evaluated for the distance-based supervised models that link the distance among the data, to anticipate feedback. The principal component analysis is implemented to reduce the magnitude of the large datasets and remodel a larger set of variables into a smaller one, without compromising the information in the large set. In the PCA, standardization is achieved and the variables are scaled using Equation (2).

$$\beta_1 = (v_1 - m_1) / sd_1 \quad (2)$$

where β_1 denotes the standard value, whereas v_1 denotes the value, m_1 denotes the mean, and sd_1 denotes the standard deviation. Each feature vector's weights are represented by the variable ω for all of the n variables represented by v . The input features of the dataset are taken as the variables, and ω represents the weight for each feature. Equation (3) is the linear combination that predicts the value. As a result of the weight, the following Equation (3) expresses how the feature vector changes using the variable $\eta(p)$. Here, η_p means the prediction attained by the input features to a machine learning model.

$$\eta_p = (\omega_1 \times v_1, \omega_2 \times v_2, \dots, \omega_n \times v_n) \quad (3)$$

As the count of the feature vectors in the associated dataset increases, the weights associated with each feature vary, which is expressed in Equation (4) with the number of feature vectors. In Equation (4), the weights associated with each feature are summated for predicting the output value.

$$\Delta c = \frac{1}{n} \sum_{x=1}^n \left(\frac{\partial u}{\partial v_1} \times w_1 + \frac{\partial u}{\partial v_2} \times w_2 + \dots + \frac{\partial u}{\partial v_n} \times w_n \right) \quad (4)$$

From the above equation, the variable u , the feature vector, is another corresponding variable with which variable v is being considered. The SHAP calculates the relevant Shapley values for each feature, to establish its importance [61].

3.2. Explainable Feature Weight Initialization and Normalization

When independent factors have varying effects on the class label, weighing the attributes to boost performance, and the in-class labeling is feasible. The normalized mutual information (NMI) [62] is used as the feature weight among each feature and class label. This dataset contains two parameters, α , and β , over the sample i, j concerning the class label c . Here, the class label of the instances is required to determine the normalized mutual information. It computes the mutual information score to scale the results between 0, it refers to no mutual information, and 1 is the perfect correlation. Equations (5) and (6) will estimate the initial weights for the features.

$$i_w(\alpha) = \frac{mi(\alpha, c)}{avg(e(\alpha), e(c))} \quad (5)$$

$$i_w(\beta) = \frac{mi(\beta, c)}{avg(e(\beta), e(c))} \quad (6)$$

From Equations (5) and (6), the function $i_w()$ is used in assessing the initial weights of the feature, and the variable mi refers to the mutual information. The variable e denotes the entropy associated with the feature. The function $avg()$ in the denominator is used in assessing the mean of the values. The mutual information associated with the feature over the dataset with T_i instances, are evaluated, as denoted in Equations (7) and (8).

$$mi(\alpha, c) = \sum_{x=0}^{i-1} \sum_{y=0}^{j-1} \frac{|\alpha_x \cap c_y|}{T_i} \log \left(\frac{T_i |\alpha_x \cap c_y|}{|\alpha_x| |c_y|} \right) \quad (7)$$

$$mi(\beta, c) = \sum_{x=0}^{i-1} \sum_{y=0}^{j-1} \frac{|\beta_x \cap c_y|}{T_i} \log \left(\frac{T_i |\beta_x \cap c_y|}{|\beta_x| |c_y|} \right) \quad (8)$$

Entropy ' e' refers to the information's degree of unpredictability and irregularity. The variables ' i ' and ' j ' denote the number of samples in each of the clusters. The more information there is, the more challenging it is to conclude [63]. Equation (9) is used to calculate the entropy. It represents the conditional entropy and the mutual information reduces the entropy of the class label v which is similar to the information gain in the decision trees.

$$e(v) = - \sum_{x=0}^{i-1} \rho \left(\frac{|v_x|}{T_i} \right) \log \left(\frac{|v_x|}{T_i} \right) \quad (9)$$

A feature normalization strategy is used for a more uniform distribution of the training dataset. While reducing the loss function, the feature set associated with the program comprises variable values during training. The algorithm scales up and down with each iteration until it finds the global or local best value. The min-max normalization is used to scale the feature values from 0 to 1 in the proposed XAI-based algorithm. Conventional scaling methods are vastly outclassed by the normalizing strategy known as min-max. The non-Gaussian feature distribution may be addressed through min-max scaling. The

min-max normalization is used to overcome the problem of the accuracy loss by optimizing the gradient, as the problem moves in the direction of a global solution. Equation (10) provides goal values ranging from 0 to 1, relying on the minimum and maximum values of the column of the particular feature.

$$\gamma_{new} = \frac{\gamma - \gamma_{min}}{\gamma_{max} - \gamma_{min}} \quad (10)$$

From the above equation, The variable γ_{new} denotes the scaled values of the feature weight in the range of 0 and 1. The variable γ_{min} , denotes the smallest value in the entire column of the corresponding feature, and similarly, the γ_{max} , denotes the maximum value associated with the particular feature in the column.

The change influences the result in weights of the features in the corresponding feature vector, which is represented using the function $\delta(x)$, which is evaluated as shown in Equation (11) with the variable m denoting the sum of feature vectors in the corresponding dataset, with the weight w over the features $[y_1, y_2, \dots, y_m]$.

$$\delta(x) = \frac{1}{i} \sum_{m=0}^{i-1} \left[\frac{\partial x}{\partial y_1} \times w_1 \times y_1 + \frac{\partial x}{\partial y_2} \times w_2 \times y_2 + \dots, \frac{\partial x}{\partial y_m} \times w_m \times y_m \right] \quad (11)$$

3.3. Feature Weight Optimization

The minority samples might be more important if their weights were increased, making them more difficult to disregard. Generally, it might be difficult to make a rational decision without understanding the weight associated with each such class. The frequency of the classes in the training data is the frequent way to determine the weight distribution. The goal is to develop a weight distribution that will allow us to achieve good results in the training and testing sets. The explainable weight optimization model would iterate over s number of rounds over the k samples that are classified as the nearest hits and misses. The weight optimization can be best explained through Equation (12), using the gradient descent strategy with a consistent learning rate of $\frac{1}{k \times s}$.

$$O_w = \sum_{i=0}^k \sum_{j=0}^q w_j \rho^r \left(a_j^s, a_{j_{hit}}^{s_i} \right) - \sum_{i=0}^k \sum_{j=0}^q w_j \rho^r \left(a_j^s, a_{j_{miss}}^{s_i} \right) \quad (12)$$

From the above equation, the function ρ^r determines the similarity among the features and w_j denotes the current weight assigned to the feature j . The variable $a_{j_{hit}}^{s_i}$ is the j^{th} feature value associated with i^{th} , the nearest hit over the instance m . Similarly, the variable $a_{j_{miss}}^{s_i}$ is the j^{th} feature value associated with the i^{th} nearest miss.

Using the feature selection wrapper method, the number of predictors is chosen to take into consideration the importance of the target variable. Using this method, the predictors are added and removed to find the optimal combination, to maximize the model performance. Therefore, the supervised classification techniques are implemented for developing a predictive model in the proposed work. The Shapley values in the summary plot of the features, according to their importance. Therefore, as a result, ‘age’, and ‘sex’ predictors are chosen, and the ‘cholesterol’ feature is related with ‘age’, and ‘sex’, to obtain the output.

4. XAI Approach over Various Classifiers

The proposed work follows an experimental approach using machine learning techniques for the cardiovascular risk prediction. In supervised machine learning, the input data is labeled with the desired output, and the data classification is carried out. The model is trained using the machine learning algorithms under supervision, and the predictions are performed. In contrast, the input data is not labeled with any desired output in unsupervised learning and performs the clustering instead of the classification. The supervised machine learning algorithms implemented over the dataset, shown in Table 1, are the support vector machines, nearest neighbor, logistic regression, Gaussian naive Bayes, and the ensemble methods. In ML, the ensemblers aggregate the predictions of the individ-

ual classifiers that produce predictions, combined to produce an aggregate prediction. Bagging and boosting are the ensembling classifiers implemented in the proposed work. The proposed experiment predicts heart disease using SVM, AdaBoost, KNN, bagging, logistic regression, and the Gaussian naive Bayes. The features selected for performing the predictions are taken as ‘sex’, ‘age’, and ‘cholesterol’. The accuracy and the ROC curve values of these algorithms are also compared. The results show the values generated by the ML model, which predict the patients who will actually have the disease and those who are also identified as having the disease. The results also show the misclassified values where the patient is identified as not having the disease, but has the disease. The block diagram of the proposed model with the XAI framework for the feature processing that includes the feature selection, feature weight initialization, weight normalization, and weight optimization tasks in predicting cardiovascular disease is presented in Figure 2.

Table 1. Sample records associated with the dataset used in the current study.

Age	Sex	CP (Chest Pain)	TrestBps (Resting Blood Pressure)	Chol (Cholesterol)	FBS(Fasting Blood Pressure)	Respect(Resting Electrocardiographic)	Thalach (Heart Rate)	Exang (Exercise Induced Angina)	Oldpeak	Slope	Ca Vessels Colored by Fluoroscopy	Thal Thallium Stress Test	Num
63	1	3	145	233	1	0	150	0	2.3	1	0	1	1
37	1	2	130	250	0	1	187	0	3.5	1	0	2	1
41	0	1	130	204	0	0	172	0	1.4	2	0	2	1
56	1	1	120	236	0	1	178	0	0.8	2	0	2	1
57	0	0	120	354	0	1	163	1	0.6	2	0	2	1
57	1	0	140	192	0	1	148	0	0.4	1	0	1	1
56	0	1	140	294	0	0	153	0	1.3	1	0	2	1
44	1	1	120	263	0	1	173	0	0	2	0	3	1
52	1	2	172	199	1	1	162	0	0.5	2	0	3	1
57	1	2	150	168	0	1	174	0	1.6	2	0	2	1

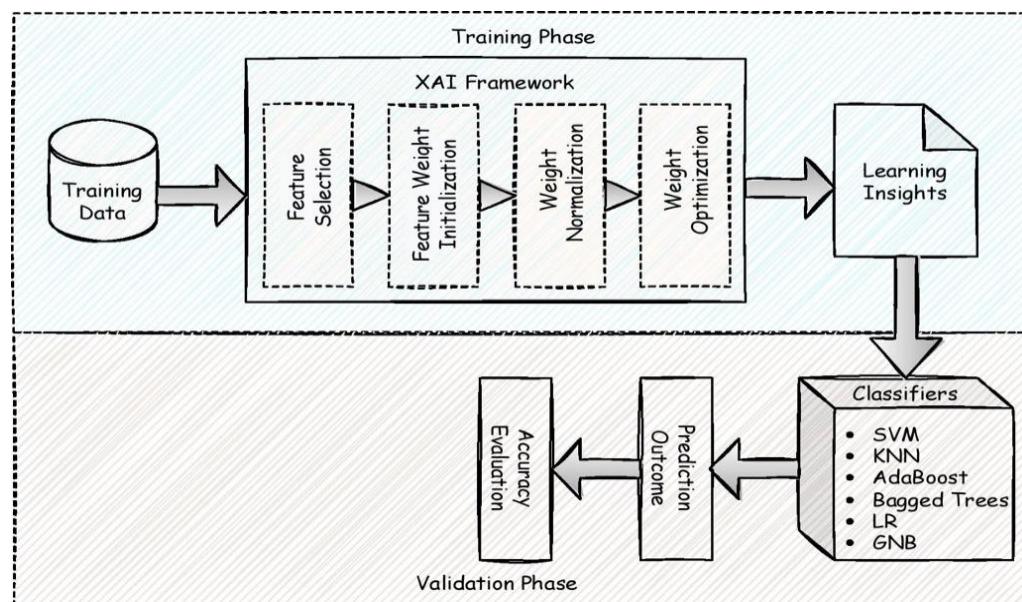


Figure 2. XAI framework in the feature processing for predicting heart disease.

In the supervised learning technique, the input variables are labeled and mapped with the output variables. The ensemble classification model is trained using the pre-existing dataset, and the model parameters may also be chosen. The features in the training dataset can be chosen using the feature extraction techniques to enhance the model's efficiency [64]. Once the model is trained, it can easily classify the input data and perform the desired predictions.

4.1. Support Vector Machines

The support vector machine is a data mining approach for regression and classification. The SVM is familiar with its capacity to deal with multi-dimensional data [65]. It is a discriminant classifier with a separating hyperplane. The SVM creates a hyperplane is utilized for classification and regression [66]. The SVM builds the model by constructing a hyperplane that separates the provided data, by maximizing the distance between the two clusters [67]. It detects the closest data vectors to the decision boundary in the training dataset, known as the support vectors (SVs), then categorizes a specific test vector using only those nearest data vectors [68]. The hyperplane is determined in Equation (13).

$$F(i) = \beta + \omega_i T_i \quad (13)$$

where ω_i denotes the weight vector, β denotes the threshold value, and T_i is the scalar offset. One of the hyperplane's potential representations is $|\beta + \omega_i T_i| = 1$, where i denotes the training set instances closest to the hyperplane, known as the support vectors. The authors used a support vector machine method to predict the drug adherence among the patients associated with heart failure [69].

4.2. AdaBoost

AdaBoost employs ensemble classifiers and boosted trees as classification algorithms. The method employs the decision tree ensembling approach. AdaBoost techniques need less time and memory than bagged trees but may require more ensemble members. The AdaBoost method, also known as adaptive boosting, grows the learners sequentially. It operates on the premise of assigning extra weight to erroneously identified examples received in the first trained model [70]. The output of the weak learners is boosted and blended into a weighted sum, to obtain the final result of the boosted method. Multiple decision trees are created in the AdaBoost algorithm. Every decision tree improves the preceding tree's weak learners. It constantly improves itself in consecutive trees, to enhance the ensemble's performance with all of its predecessors. AdaBoost increases the accuracy of a weak learning algorithm [71] and the performance of other learning algorithms, the weak learners, by merging their output. The boosted classifier's output is a weighted sum of the weak learners' findings. AdaBoost has the advantage of being less prone to overfitting. A boosted classifier in AdaBoost has the form described in Equation (14).

$$L_C(y) = \sum_{j=1}^C l_j(y) \quad (14)$$

Each l_j is a weak learner that accepts an object y as input and returns the object's class. C is the classifier in this case, and each weak learner constructs an output premise $h(y_i)$, for each instance over the training set. From each iteration j , a weak learner is chosen, and a coefficient β_j is given so that the total training error e_j is minimized. Equation (15) indicates that the resultant j stage enhanced classification is optimized [72].

$$e_j = \sum_i e[L_{j-1}(y_i) + \beta_j h(y_i)] \quad (15)$$

$L_{j-1}(y)$ represents the boosted classification that has relied on to the previous training step. $e(L)$ represents the error function, and $L_j(y) = \beta_j h(y)$ represents the weak learner considered for inclusion in the final classifier.

4.3. K-Nearest Neighbor

In KNN, the classification relies on the preponderance of the closest neighbors of a given space. In this context, the data class is allocated, based on which points in the immediate closeness of the point have the most representatives. The value supplied by the KNN is derived from the closest neighbors using uniform weights. Class labels are associated with each of the training dataset vectors. There are positive and negative classes. K-nearest neighbors use the local neighbors to predict, and the distance functions are used to compare the examples of similarity. The distance formula is the rounded magnitude of the numerical variance of their coordinates in one dimension, and it is generalized in higher dimensions, as illustrated in Equations (16) and (17) [73].

$$d(\alpha, \beta) = |\alpha - \beta| \quad (16)$$

$$d(\alpha, \beta) = \sqrt{(\alpha - \beta)^2} \quad (17)$$

In the KNN, the Euclidean distance is calculated, and the distance between p and q having Cartesian coordinates (m1, n1) and (m2, n2) is shown in Equation (18) [74].

$$d(\alpha, \beta) = \sqrt{(\alpha_1 - \beta_1)^2 + (\alpha_2 - \beta_2)^2} \quad (18)$$

4.4. Bagging

Depending on the base learner, bagging conducts classification and regression. It lowers the variance [75]. It comprises the bag size, which is the size of each bag as a proportion of the training set size. Back fitting is used to trim a decision/regression tree constructed using information gain/variance. The numeric attribute values are then sorted, and the corresponding instances are terminated when a value is missing. The total number of folds used is three. It influences how much data is utilized for trimming. The one-fold is used for pruning, while the other is utilized to expand the rules. Equation (19) [76] illustrates the final aggregate of the classifier. The average of f_i for $i = 1, \dots, T$, for the classification.

$$\bar{f}(x) = \arg_{y \in Y} \max \sum_{i=1}^T ||(f_i(x) = y) \quad (19)$$

4.5. Logistic Regression

The logistic regression is a basic classifier that calculates the output probabilities by modeling the total mean as a function of a linear combination of predictors. Logistic regression models are used for predicting the risk of a disease, such as diabetes, heart, etc. [77]. The standard logistic function is a form of sigmoid function and takes any input x, and outputs the value between 0 and 1 [78], as shown in Equation (20).

$$f(x) = \hat{e}x / (\hat{e}x + 1) \quad (20)$$

4.6. Gaussian Naive Bayes Distribution

A naive Bayes classifier relays on the Gaussian distribution in analyzing the numeric predictors and the multivariate, multinomial distribution (MVMN) for the categorical predictors. It calculates the conditional probability for the n independent variables, as shown in Equation (21). p_b is the probability for each of the i possible outcomes, the classes A_i [79] and y is the vector represented as $p_b(A_i | y_1, y_2, \dots, y_n)$.

$$p_b(A_i | y) = \frac{p_b(A_i)p_b(y | A_i)}{p_b(y)} \quad (21)$$

4.7. Dataset Description and the Implementation Environment

The dataset is collected from the UCI ML repository (<https://archive.ics.uci.edu/ml/index.php>, accessed on 10 October 2022). The dataset used in the experiment is related to heart disease. The dataset characteristics are multivariate, and the attributes are categorial, integer, and real. The task associated with the dataset is the classification. There are 303 occurrences in the dataset, and 14 features are included. Table 1 displays a representative sample of the dataset used in the current study. The ‘num’ feature is the response variable, while the other characteristics are predictors. The ‘num’ property which reflects the angiographic disease state, is the diagnosis of heart disease. It refers to the narrowing in any major blood vessel due to cholesterol and plaque deposits, as detected through the use of an angiogram. The response variable has also been reduced to a binary problem, to detect the presence of any vessel narrowing.

The target variable ‘num’ shows the diagnosis of heart disease (angiographic disease status). It divides into two classes and numerically, the ‘num’ feature can only take on two values: 0 or 1. A diameter narrowing of more than 50% in any major blood vessel is represented by a value of 1 and refers to the presence of heart disease, whereas a diameter narrowing of less than 50% in any major vessel is represented by a value of 0 and refers to no presence of heart disease in the patient.

The details about the implementation environment and the simulation framework are shown in Table 2. The details presented are the resources used in the simulation for attaining the reasonable performance of the model. The execution time is also considered one of the performance’s evaluation parameters. The ML model involves a 10-fold cross-validation, splitting the ratio of data into training and test data. The ratio of the training and testing in the current study is 70:30. A visualization of all of the 14 attributes of the dataset shown in Table 1 is shown in Figure 3. The statistical values of all of these attributes are shown in Table 3.

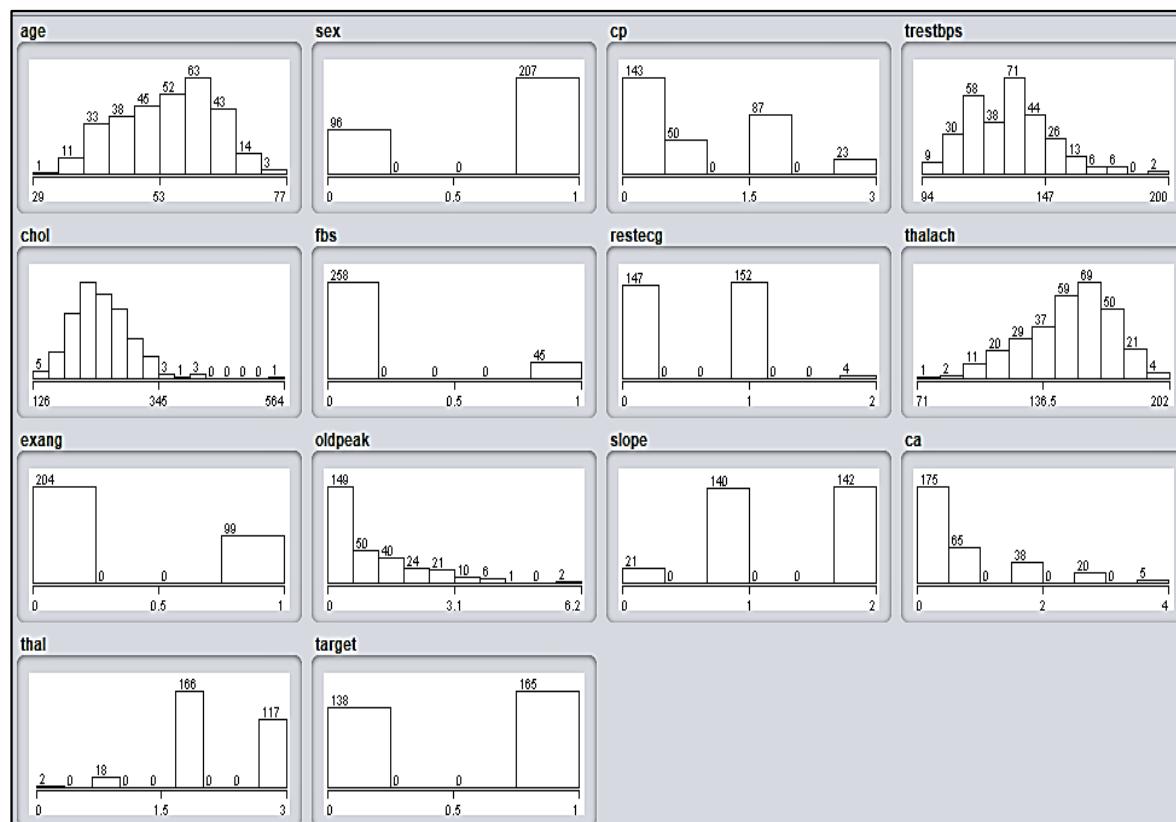


Figure 3. The visualization feature set of the cardiovascular disease dataset.

Table 2. Details of the implementation environment used in the current study.

Environment Details		Specification
Operating Systems		Microsoft Windows 10
System Type		x64-based type
Processor		Intel(R) Core i7-1165 CPU @ 2.80–4.70 GHz, 12 MB Cache
Architecture		64-bit
Processor Graphics in Use		Intel(R) Iris Xe Graphics
Installed Physical (RAM)		8 GB
Software Used		Matlab R2021a, Weka 3.8.5

Table 3. The statistical value of the feature sets and their information.

Attribute	Minimum	Maximum	Mean	Standard Deviation	Attribute Type	Attribute Information
Age	29	77	54.366	9.082	Numeric	-
Sex	0	1	0.683	0.466	Nominal	0: female, 1: male
cp (chest pain type)	0	3	0.967	1.032	Nominal	1: typical angina 2: atypical angina 3: non-anginal pain 4: asymptomatic
trestbps (blood pressure when resting)	94	200	131.624	17.538	Numeric	-
chol (serum cholesterol in mg/dL)	126	564	246.264	51.831	Numeric	-
fbs (fasting blood sugar > 120 mg/dL))	0	1	0.149	0.356	Nominal	0: false 1: true
restecg (resting electrocardiographic results)	0	2	0.528	0.526	Nominal	0: normal 1: having ST-T wave abnormality 2: showing probable or definite left ventricular hypertrophy by Estes' criteria
thalach (max heart rate achieved)	71	202	149.647	22.905	Numeric	-
exang (exercise-induced angina)	0	1	0.327	0.47	Nominal	0: no 1: yes
Old peak (exercise-induced ST depression in comparison to the rest)	0	6.2	1.04	1.161	Numeric	-
slope (the peak exercise ST segment's slope)	0	2	1.399	0.616	Nominal	1: upsloping 2: flat 3: downsloping
ca (count of the major vessels colored by fluoroscopy (0–3))	0	4	0.729	1.023	Nominal	-
thal	0	3	2.314	0.612	Nominal	3 = normal, 6 = fixed defect, 7 = reversible defect
num (predicted attribute)	0	1	0.545	0.499	Nominal	0: no risk 1: risk

5. Results and Discussion

The performance of the proposed XAI-based prediction of cardiovascular disease using the ensemble classifiers is being assessed over various evaluation metrics, such as the AUC, accuracy, true positive rate (TPR), recall, and precision. The XAI-driven feature selection and weight optimization have enhanced the model's performance and made the internal evaluation of the model better interpretable, to make the decisions trustworthy. The values of the confusion matrix that attributes the true positive (tp), false positive (fp),

true negative (tn), and false negative (fn) are used in assessing the performances. The instances of correctly identifying the abnormality as the positive cases are recognized as the true positive, whereas recognizing the normal cases correctly as negative instances are recognized as true negative. The misinterpretation of the normal instances as positive is recognized as a false positive. Similarly, the cases that are recognized as negative, in fact, are positive instances, and are recognized as false negative. The comparative results obtained from the dataset using six ensemble models, i.e., SVM, KNN, AdaBoost, bagging, logistic regression, and the Gaussian naive Bayes algorithms, are shown in Table 4. The accuracy obtained by the SVM after training on the dataset is 82.5%, the highest accuracy among the other algorithms.

Table 4. Comparative analyses of the ensemble classification models.

Model	Training Time	Features	Predictors	Accuracy (Validation)	tpr (True Positive Rate) or SENSITIVITY $tpr = \frac{tp}{tp+fn}$	fnr (False Negative Rate) $fnr = \frac{fn}{fn+tp}$	ppv (Positive Predicted Value) $ppv = \frac{tp}{tp+fp}$	fdr (False Discovery Rate) $fdr = \frac{fp}{fp+tp}$
AdaBoost	8.7015 s	13	X: age	79.5%	Predicted Class 0	Predicted Class 1	Predicted Class 0	Predicted Class 1
					75.4%	83.0%	24.6%	17.0%
KNN	6.605 s			75.9%	69.6%	81.2%	30.4%	18.8%
					75.6%	76.1%	75.6%	76.1%
SVM	7.9932 s			82.5%	71.7%	91.5%	28.3%	8.5%
Bagging	12.039 s	Y: sex		79.2%	74.6%	83.0%	25.4%	17.0%
					78.6%	79.7%	78.6%	79.7%
Logistic Regression	5.3715 s		X: age	81.2%	74.6%	86.7%	25.4%	13.3%
					82.4%	80.3%	82.4%	80.3%
Gaussian Naïve Bayes	2.0986 s		Y: chol	80.2%	76.1%	83.6%	23.9%	16.4%
					79.5%	80.7%	79.5%	80.7%
					20.5%	19.3%	20.5%	19.3%

The training time taken by the SVM is 7.9932 s. The kernel function used by the SVM is linear, which implements the one-vs-one multiclass method. The accuracy obtained by the boosted trees is 79.5%, and the training time taken is 8.7015 s. The boosted trees use the AdaBoost ensemble technique, with the learner type set to the decision tree, a maximum of 20 splits, and a total of 30 learners. The next ML model is KNN and the accuracy obtained is 75.9% which is the least one. However, the training time taken by the KNN algorithm is less than the other two models, i.e., 6.605 s. The distance metric implemented by the KNN is Euclidean distance, and the distance weight is equal. The accuracy achieved by the bagging tree model is 79.2%, whereas the LR and the Gaussian naive Bayes' accuracy is 81.2% and 80.2%.

The predictions for the features 'age' and 'sex' of patients suffering from cardiovascular disease in the form of scatter plots obtained from the SVM, KNN, AdaBoost, bagged tree, LR, and the Gaussian naive Bayes are shown in Figure 4a–f, whereas the scatter plots of the correct predictions of the Class 0 and Class 1 on the attributes 'sex' and 'cholesterol', using the algorithms, such as SVM, KNN, AdaBoost, bagged tree, LR, and the Gaussian naive Bayes are shown in Figure 5a–f.

The scatter plots shown in Figure 4 examine the classifier results and show the model predictions. The correct and incorrect predictions of Class 0 and Class 1 for the features 'age', and 'sex' are visualized. The predictors identified here separate the classes well by plotting on the scatter plot. It also helps us to investigate the features to include or exclude and visualize the training data, and the misclassified points on the scatter plot.

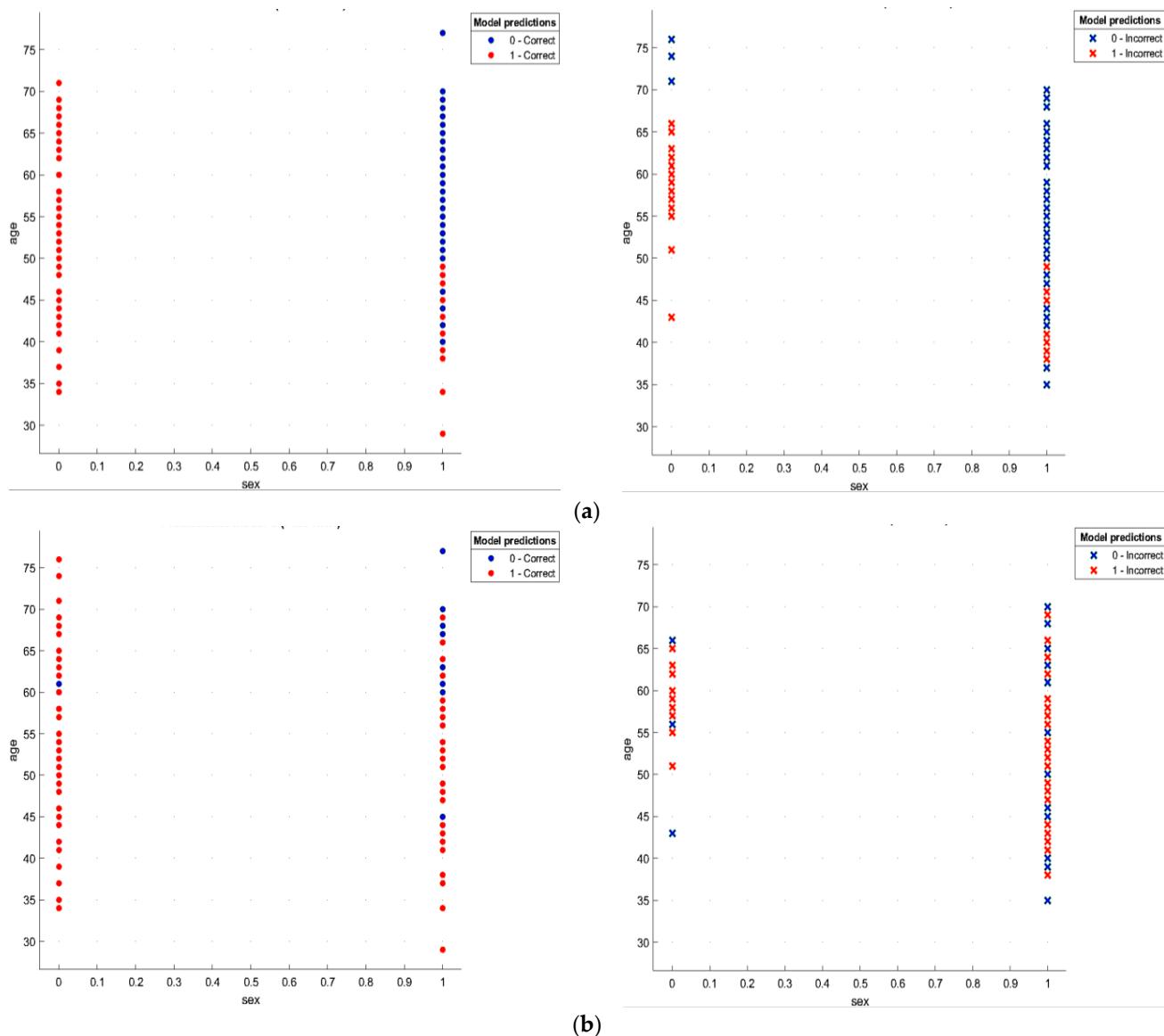


Figure 4. Cont.

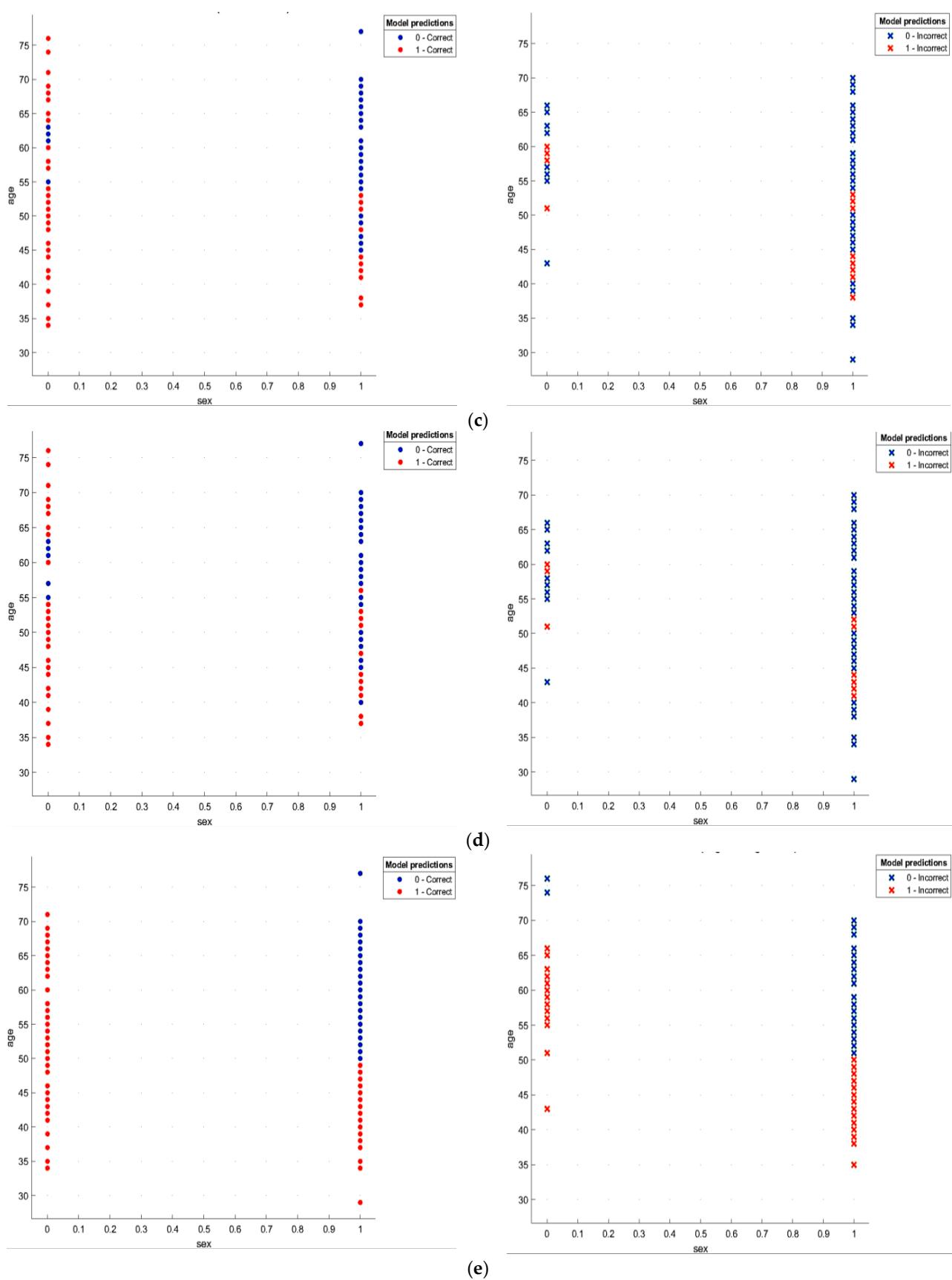
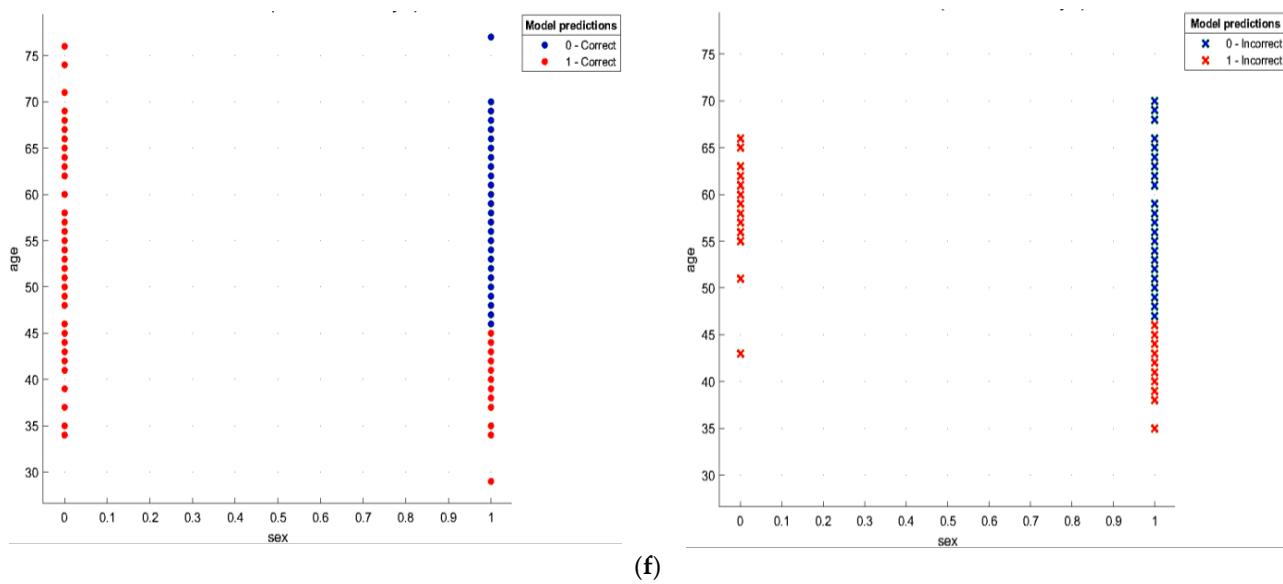
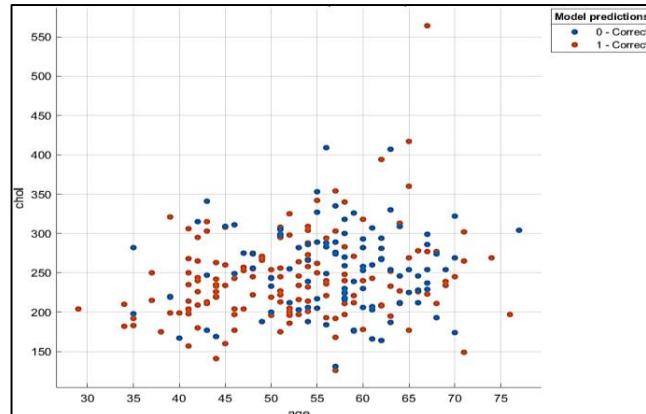


Figure 4. *Cont.*

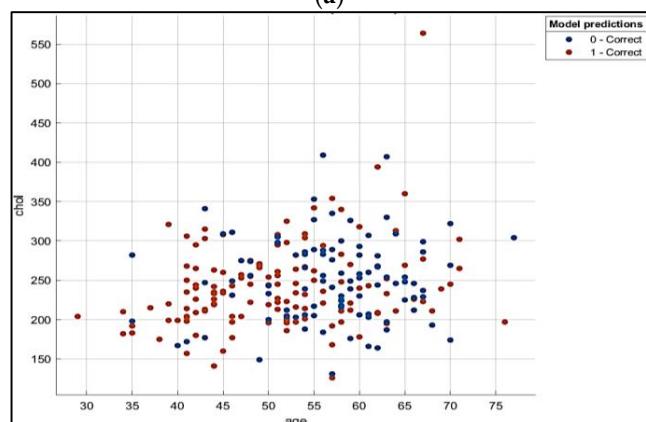


(f)

Figure 4. (a) SVM model with the correct and incorrect predictions concerning the features ‘age’ and ‘sex’. (b) KNN model with the correct and incorrect predictions concerning the features ‘age’ and ‘sex’. (c) AdaBoost model with the correct and incorrect predictions concerning the features ‘age’ and ‘sex’. (d) Bagged tree model with the correct and incorrect predictions concerning the features ‘age’ and ‘sex’. (e) Logistic regression model with the correct and incorrect predictions concerning the features ‘age’ and ‘sex’. (f) Gaussian naive Bayes model with the correct and incorrect predictions concerning the features ‘age’ and ‘sex’.



(a)



(b)

Figure 5. Cont.

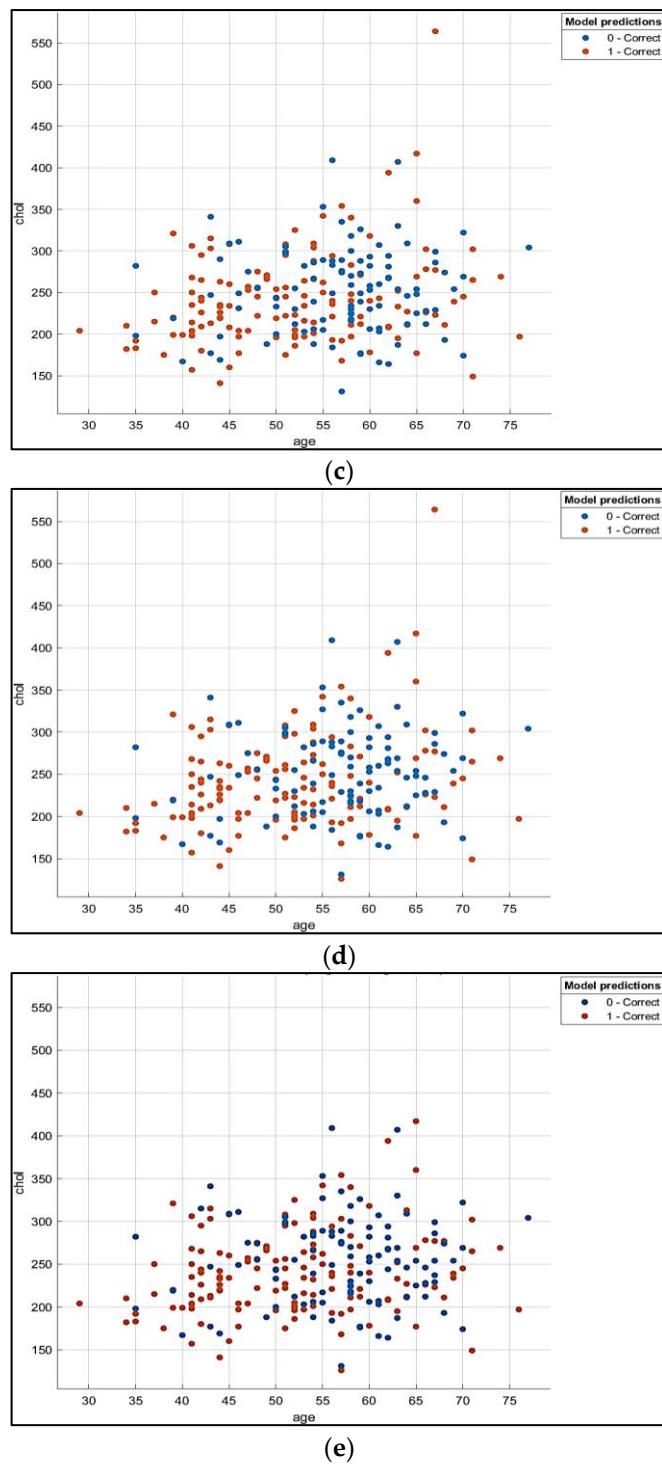
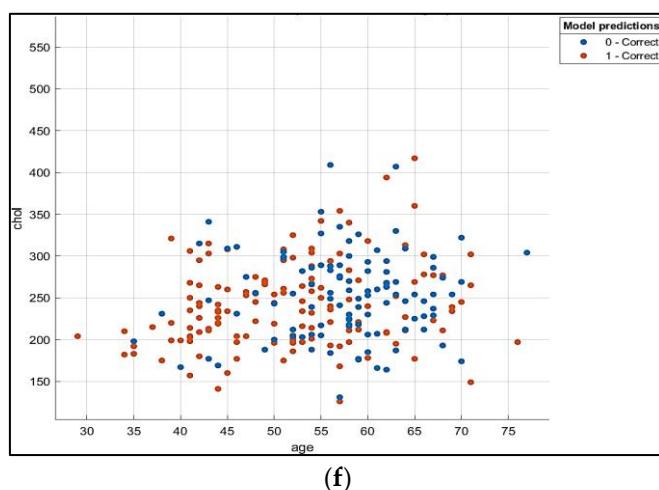


Figure 5. Cont.



(f)

Figure 5. (a) SVM model with the correct predictions concerning the features ‘age’ and ‘cholesterol’. (b) KNN model with the correct predictions concerning the features ‘age’ and ‘cholesterol’. (c) AdaBoost model with the correct predictions concerning the features ‘age’ and ‘cholesterol’. (d) Bagged tree model with the correct predictions between ‘age’ and ‘cholesterol’. (e) Logistic regression model with the correct predictions concerning the features ‘age’ and ‘cholesterol’. (f) Gaussian naive Bayes model with the correct predictions concerning the features ‘age’ and ‘cholesterol’.

Plaque accumulation in the walls of the arteries that deliver blood to the heart causes coronary artery disease, which are also known as coronary arteries. Plaque is formed by cholesterol deposits. Over time, the plaque formation causes the interior of the arteries to constrict, which is called diameter narrowing. The risk analysis is evaluated against the two parameters, such as “age vs. sex”, “age vs. cholesterol”. It is derived from the results of two features, ‘age’ and ‘sex’, that the maximum number of male patients, compared to females in the age group of 35 to 65, are showing the Class 0 positive predictions, the angiographic heart disease status has a less than 50% diameter narrowing and is at a low risk. However, for the Class 1 positive predictions, where the angiographic disease status is greater than 50%, the diameter narrowing in the age group of 35 and 70, the females, more than the males, are at a greater risk. Upon taking the features, ‘age’ and ‘cholesterol’, it is observed that the cholesterol level in the age group of 50 to 65, in both males and females, is above 240. The angiographic disease status, refers to a less than 50% diameter narrowing, is lower than the angiographic disease status that has a greater than 50% diameter narrowing, in the same age group, for cholesterol levels above 240. It shows that a cholesterol level above 240 in both males and females, aged 50 to 65, is risky and needs to be carefully related to heart disease.

The SHAP values determine how each attribute contributes to the model’s prediction. In Figure 6, at the x-axis, the base value is $E [f(x)] = 0.541$, which is the average predicted value of the target class or the outcome in the current work for the first observation, ‘age’ in the sample dataset, whereas, the aggregate values of all of the features of the sample dataset are shown in Figure 7, using a mean SHAP plot. For each attribute, the mean of the absolute SHAP values across all of the observations are calculated, and all of the features are shown using a bar plot. It is derived from the mean plot in Figure 7, that the ‘sex’ feature weight is greater than the other features, followed by ‘age’, ‘thalach’, and ‘ca’, in model’s predictions. The Shapley values show the impact of each feature and a red color in Figure 6 represents high Shapley values, whereas the blue color represents low Shapley values. Therefore, for the feature ‘sex’, if the values are high i.e., in the color red, the Shapley values will be low, otherwise, when the values are low, in the blue color, the Shapley values will be high. The impact of each feature on a specific sample is shown in Figure 8. In Figure 8, on the left-hand side, each feature is ordered according to its importance. The ‘sex’ feature is the

most important and generates the most impact on the predictions, whereas the ‘fbs’ feature is the least important.

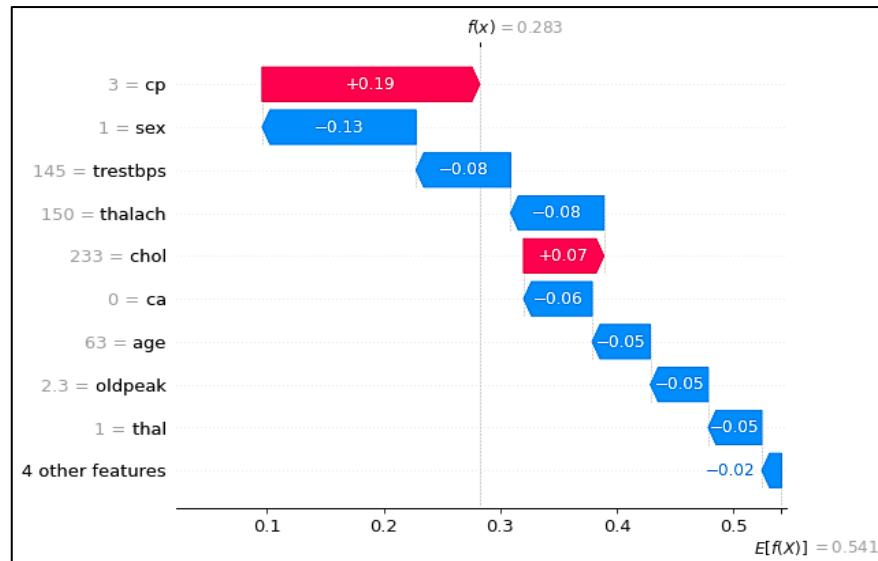


Figure 6. SHAP value for the ‘age’ factor.

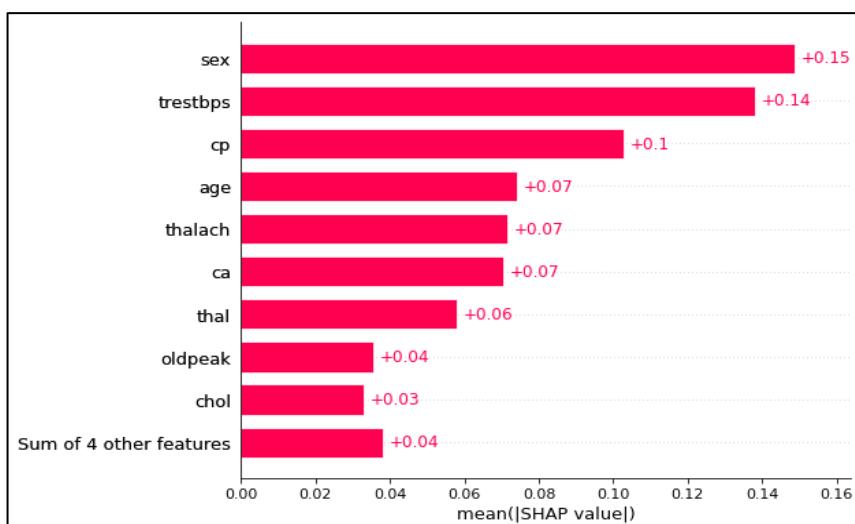


Figure 7. SHAP absolute mean plot.

In the XAI framework, there are three stages which are pre-modelling, explainable modeling, and post-modelling. In the proposed work, all three stages are followed, so that the results obtained are well explained and interpreted clearly. The evaluation of the influence of each feature on the model predictions is obtained and the SHAP results in Figure 8 also show the importance of each feature in the predictive results. In the pre-modelling stage, the dataset is examined and its prior understanding is crucial. The perception of the dataset is important to apply the statistical techniques and to obtain information, such as mean, standard deviation, etc., which are shown in Figure 3.

In the second stage of the explainable modeling, the model achieves the predictive results after observing the dataset, and the feature selection, and optimization are attained to include or exclude in the model’s training. In this stage, in the proposed work, the XAI framework is shown in Figure 2. This stage consists of white-box models which involve the SVM, naive Bayes, logistic regression, KNN, etc., which are easier to explain and interpret, compared to the black-box models, such as neural networks, bagged trees, AdaBoost,

random forest, etc. The results obtained from the white-box models are explainable, transparent, and well trusted by the doctors and clinical staff for diagnosis. The inherent workings of the black-box models are complex to understand and they don't provide an estimate of the importance of each feature on the model predictions, nor is it easy to understand how the different features are related, but they are accurate in results, compared to the white-box models.

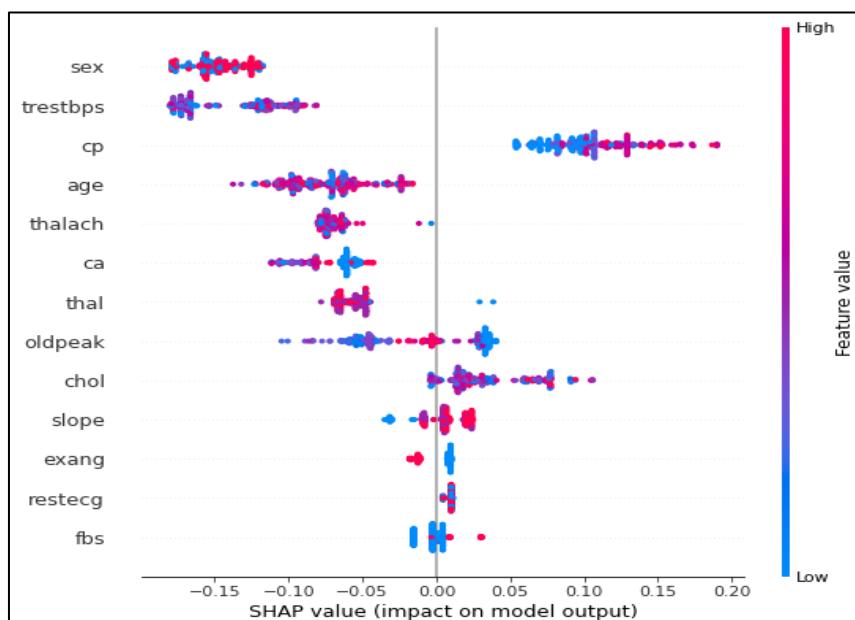


Figure 8. Summary plot of the SHAP values.

In the last stage of the post-modeling stage, the post-justification of the predictive results by the model is given. In this, the results are not self-explanatory and are not understood by the model design. The black-box models are not explainable by themselves and therefore, techniques, such as model properties, local logic, global logic, etc., are adopted to make the black-box model explainable from the internal logic or the output of the model. Therefore, in this stage of the XAI, the results generated from both models are compared and their results are shown in Table 4, whereas the results in the form of a confusion matrix are shown in Table 5.

Table 5. Confusion matrix and the ROC curve values of various classifiers.

Model	AUC	Accuracy	TNR/Specificity	TPR/Recall	F-Score	Precision	Training Time (s)
AdaBoost	0.85	79.5%	75.4%	83%	81.5%	80.1%	8.7
SVM	0.89	82.5%	71.7%	91.5%	84.7%	78.9%	7.9
KNN	0.75	75.9%	69.6%	81.2%	78.5%	76.1%	6.6
Bagging	0.88	79.2%	74.6%	83.0%	81.2%	79.6%	12.0
Logistic Regression	0.89	81.2%	74.6%	86.7%	83.3%	80.3%	5.3
Gaussian Naive Bayes	0.89	80.2%	76.08%	83.6%	82.1%	80.7%	2.0

The ROC curve train categorizes a patient's disease condition as either positive or negative, based on the test results and perceives the ideal cut-off value with the best symptomatic performance [80]. The AdaBoost algorithm has a false positive rate (FPR) of 0.17, suggesting that the present model mistakenly classified 17% of the data as positive. A genuine positive rate of 0.75 means that the present model properly classified 75% of the data as positive. The ensemble ML model has an area under the curve of 0.85. The value of the AUC (area under the curve) measures the trained model's overall quality. A higher AUC value implies that the ML model is doing better. The AUC curve quantifies the potential of

a ML model to determine between the classes and is widely used to measure the accuracy of the diagnostic tests. The higher value of the AUC curve depicts the model's performance at determining between the positive and negative class values, Class 0 and Class 1, in the current study for the classification models. The ROC curve is known as the curve of probability. The TPR value is plotted on the y-axis, whereas the FPR is plotted on the x-axis. The trade-off observed among the true positive rate and the false-positive rate over a prediction software model for heart disease that utilizes varying probability thresholds, is summarized to interpret the performance using the ROC curves for the AdaBoost ensemble, SVM, and the KNN trained models presented in Figure 9a,b, Figure 10a,b and Figure 11a,b. The ROC (receiver operating characteristic) curve shows the TPR (true positive rate) and FPR (false positive rate) values of the trained ML Models.

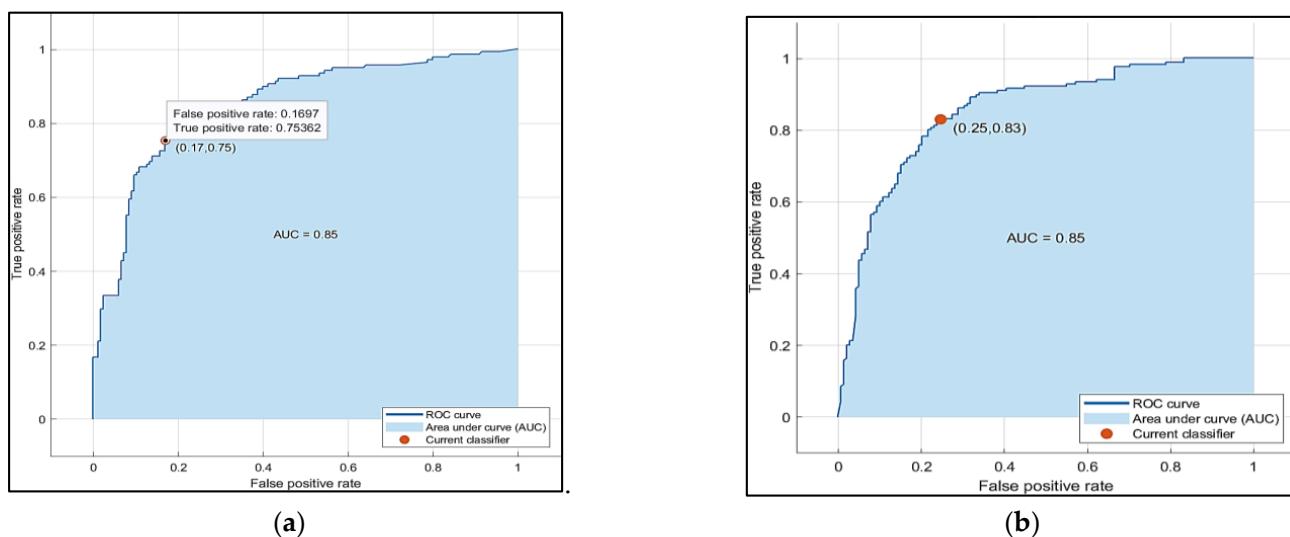


Figure 9. (a) ROC curve corresponding to class 0 (b) ROC curve corresponding to class 1 using AdaBoost.

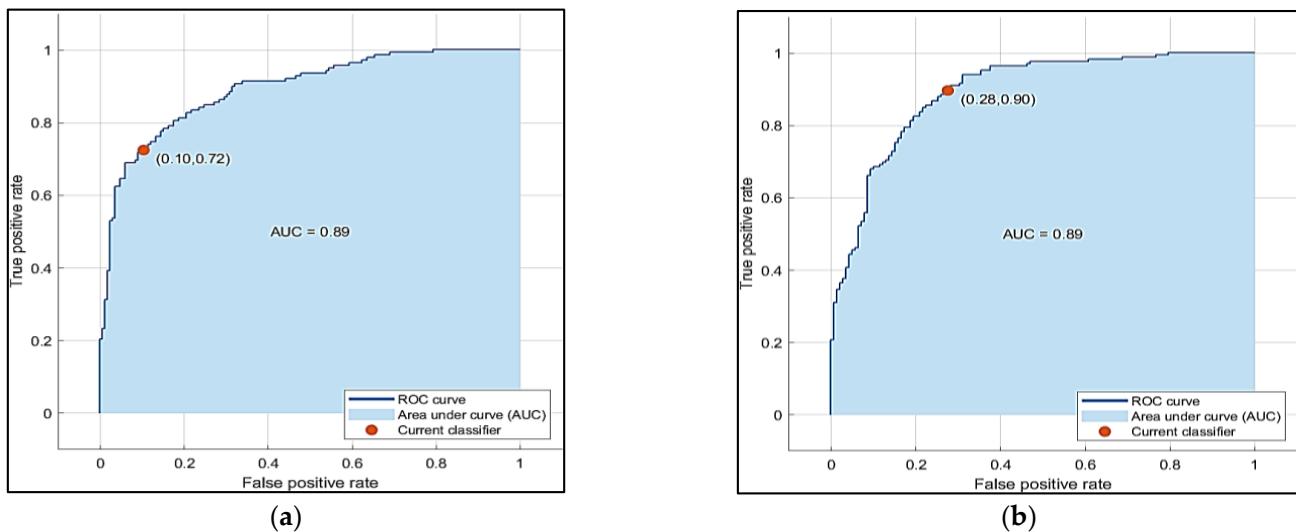


Figure 10. (a) ROC curve corresponding to class 0 using the SVM (b) ROC curve corresponding to class 1 using the SVM.

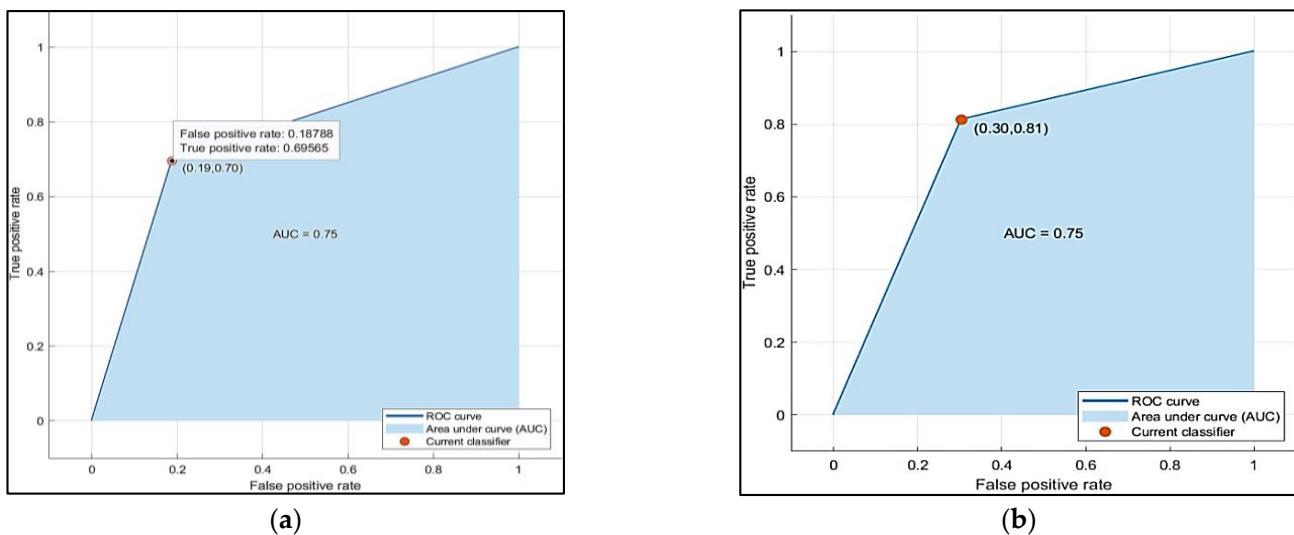


Figure 11. (a) ROC curve corresponding to class 0 using the KNN (b) ROC curve corresponding to class 1 using the KNN.

The SVM classifier has a false positive rate (FPR) of 0.078, indicating that the trained model mistakenly allocated 7.8 percent of the data to the positive class. A genuine positive rate of 0.72 depicts that the model properly allocated 72 percent of the data to the positive Class 0 and 90 percent to the positive Class 1. The SVM model has an AUC of 0.89.

The KNN model has a false positive rate (FPR) of 0.18, indicating that the model mistakenly allocated 18% of the data to the positive class, and a true positive rate (TPR) of 0.69, indicating that the model properly assigned 69 percent of the observations to the positive class. The AUC for the KNN Model is 0.75. Figures 12a,b, 13a,b and 14a,b exhibit the ROC curve values for the bagged trees, LR, and the Gaussian naive Bayes. The AUC curve of the bagged tree is 0.88, but the AUC curves of the LR and the Gaussian naive Bayes are both 0.89.

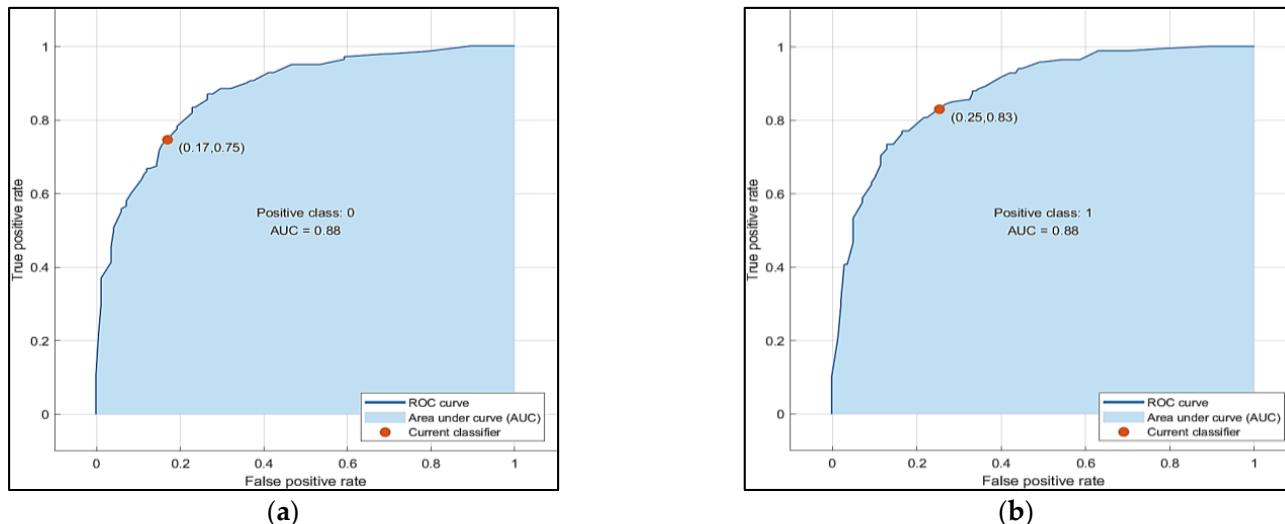


Figure 12. (a) Bagged tree model ROC curve corresponding to class 0. (b) Bagged tree model ROC curve corresponding to class 1.

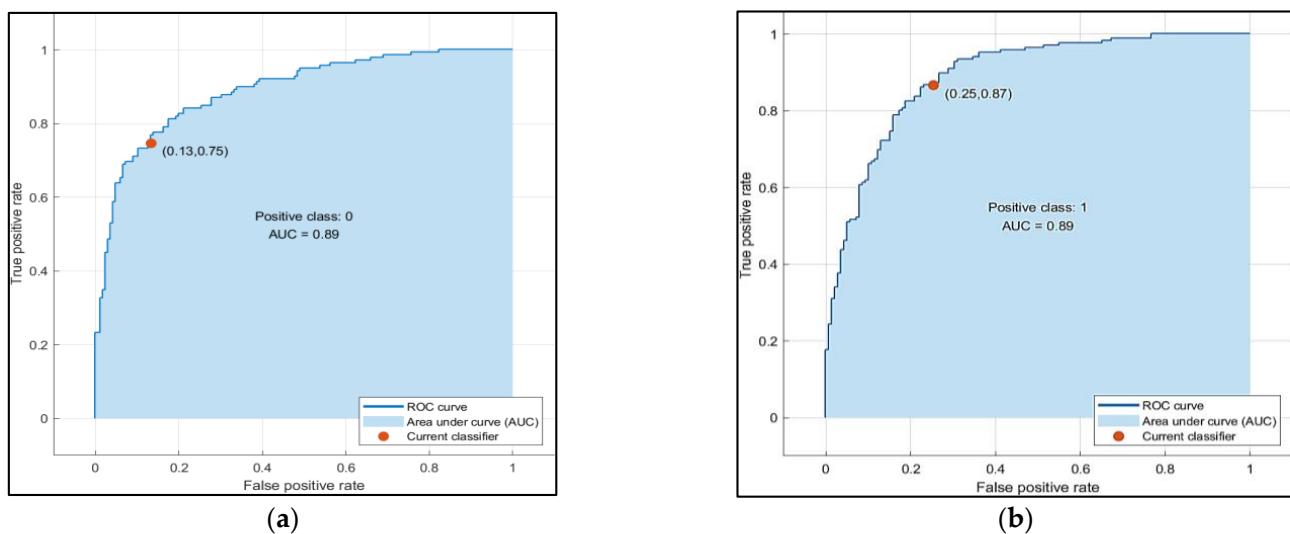


Figure 13. (a) LR model ROC curve corresponding to class 0 (b) LR model ROC curve corresponding to class 1 using the LR.

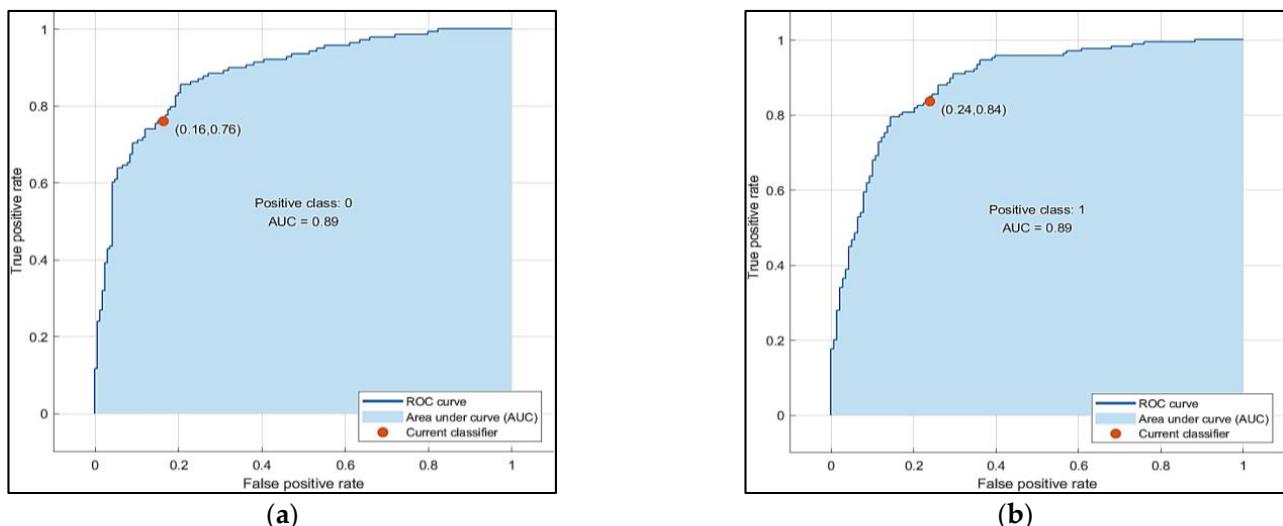


Figure 14. (a) Gaussian naive Bayes model ROC curve corresponds to class 0. (b) Gaussian naive Bayes model ROC curve corresponds to class 1.

The values obtained from the confusion matrix for the performance evaluations of all three ML models are shown in Table 5. It is derived from the table, that the F-measure and recall values achieved by the SVM are 84.7% and 91.5%, which is the highest among the other algorithms, AdaBoost, KNN, bagged trees, LR, and the Gaussian naive Bayes.

The performances of various existing classification models for heart disease prediction are compared with the XAI-driven models, in terms of the AUC classifiers. The obtained experimental values are presented in Table 6. From Table 5, it is inferred that the ensemble classifier results are low, compared to the other models. Certain reasons after it involves the pre-processing techniques on the sample dataset shown in Table 1, are carried out to detect the outliers, to derive the maximum margin solution, and to avoid the issues of underfitting, and overfitting of the results. Overfitting can be defined as the disparity between the biased training and testing work. Overfitted models may lead to lacking conclusions that may wrongly or even harmfully shape clinical decision-making [81]. The feature selection, calculating a score, based on the weight coefficient times feature value and optimization, as shown in Figure 2 in particular, supports the models to give finer results than the ensembles. The SVM, naive Bayes, and the LR produce prediction results

alongside the influencing variables, which makes the prognosis fully explainable and drives the final decision. Another important factor is that the SVM and logistic regression models perform well with less data and more features.

Table 6. Comparison of the area under the curve (AUC) value of the various ensemble classifiers.

Reference	Disease	Dataset	Classifiers	AUC (Best Classifier)
[82]	Heart Disease	University Federico II	ADA, AdaBoost, LR, Naive Bayes, Random Forest (RF), Rpart, SVM, XGBoost	LR with 75%
[83]	Heart Disease	UK Biobank dataset for heart disease	Artificial Neural Network (ANN), Naive Bayes, RF, Lasso, Ridge Regression (RR), SVM, LR	RF with 79.9%
[84]	Heart Disease	Kazakh population in Xinjiang	Naive Bayes, RF, SVM, XGB DT, kNN, LR	LR with 87.2%
[85]	Heart Disease	Cleveland dataset-For heart disease by UCI, Statlog heart disease dataset.	LR, Multiple Linear Regression, RR, Decision Tree, SVM, kNN, Naive Bayes	Naive Bayes with 83%
[86]	Coronary Artery Disease	NEU Hospital dataset for heart disease	KNN, SVM, RF, ANN, Naive Bayes, LR	LR with 81.3%
Proposed XAI-Ensemble Classifiers	Heart Disease	Cleveland dataset-For heart disease by UCI repository	AdaBoost, SVM, KNN, Bagging, LR, Naive Bayes	Naive Bayes, LR, SVM with 89%

The efficiency of the proposed XAI-driven ensemble classifiers is better than the conventional classification models. Moreover, the XAI framework is more interpretable, and the decisions are trustworthy. The performance of the proposed model on evaluating the real-time scenarios is far apart from the dataset used in the evaluation. Moreover, this can be considered a potential challenge of the model. Many underlying factors are associated with analyzing the real-time case study, as the training samples must be acquired from the same demographic population. The people living in the same demographic location would have a similar lifestyle and health standards. Acquiring adequate training and testing is a challenging task, linked with technical and economic feasibilities. The future study may address the aforementioned issue in acquiring the real-time data and evaluating the ensembled model, over the XAI framework.

6. Conclusions

AI, along with ML and IoT, is playing a significant role in the healthcare industry and is emerging as a potential tool for developing and applying intelligent systems in the healthcare industry. AI-based applications save patients time and money and provide them with clinical assistance on a preliminary basis. AI applications integrated with ML models can perform diagnostics, prognosis, procedures, interpretations, and the investigation of patients with more precision and assist physicians and radiologists in achieving conclusive results. In a certain disease, it is not possible to interpret correctly. The AI, ML, and IoT emerging technologies are becoming a boon for patients living in remote areas and receiving preliminary assistance, saving them money, and preventing discomfort. In the current study, the contribution of AI, and ML in healthcare is discussed, followed by the advantages and challenges of ML and AI in the healthcare industry. The experimental approach for predicting heart disease in patients using machine learning techniques is also discussed. The SVM algorithm has exhibited a better performance with an accuracy of 82.5%, among all of the classifiers used in heart disease classification.

SVM, KNN, AdaBoost, bagged trees, LR, and Gaussian naive Bayes machine learning algorithms are discussed in the proposed work. In the in-depth analyses of the XAI techniques, including feature selection, explainable feature weight initialization, normalization, and optimization, are explored for a proper explainability and interpretation of the results on a heart disease-based dataset. The limitations of the existing models are that the dataset taken has a limited scope of attributes and sample size. However, the classification models are explained in the current study, and the XAI techniques are discussed for the results obtained. There is also a demand for self-learning models to perform the classification, based on minimal data. As a future scope of work, other ML classifications and clustering techniques can be implemented on different healthcare datasets, to explore the applications of AI, ML, and IoT-rich intelligent systems in the healthcare industry. Deep learning and neural network fields with larger datasets, related to cardiovascular disease, can be explored as future work, along with the integration of the trained machine learning models into proper explainable AI interface systems, such that the predictive results obtained from the machine learning models have a proper explainability, transparency, and are trusted by the doctors and clinical support staff. In the current study, the XAI-driven model for the classifiers for heart disease prediction is evaluated over a single dataset, the Cleveland dataset. The future dimension of the research could include the statistical analysis of the model over the divergent datasets, such as the UK Biobank dataset, University Federico II, Statlog heart disease dataset, and the NEU Hospital dataset for heart disease.

Author Contributions: The work was conceptualized by P.G., S.A. and P.N.S.; The work is the formal analysis by N.A., S.A. and F.K.A.; Methodology and Implementation by P.N.S., S.A. and P.G.; The investigation of the study is carried out by S.A., N.A. and P.N.S.; Under the administration of F.K.A., S.A. and N.A.; draft of the manuscript, and the draft post review is supported by P.N.S., P.G. and S.A. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia. [Grant No. 1996].

Acknowledgments: The authors acknowledge the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia, under project Grant No. 1996.

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

References

1. Dhillon, A.; Singh, A. Machine learning in healthcare data analysis: A survey. *J. Biol. Today's World* **2019**, *8*, 1–10.
2. Srinivasu, P.N.; Bhoi, A.K.; Nayak, S.R.; Bhutta, M.R.; Woźniak, M. Blockchain Technology for Secured Healthcare Data Communication among the Non-Terminal Nodes in IoT Architecture in 5G Network. *Electronics* **2021**, *10*, 1437. [\[CrossRef\]](#)
3. Ma, L.; Wang, X.; Wang, X.; Wang, L.; Shi, Y.; Huang, M. TCDA: Truthful combinatorial double auctions for mobile edge computing in industrial Internet of Things. *IEEE Trans. Mob. Comput.* **2021**, *21*, 4125–4138. [\[CrossRef\]](#)
4. Felkey, B.G.; Fox, B.I. Is This the First Adherence-Focused Multidisciplinary Care Team App? *Hosp. Pharm.* **2016**, *51*, 94–95. [\[CrossRef\]](#)
5. Pan, Y.; Fu, M.; Cheng, B.; Tao, X.; Guo, J. Enhanced deep learning assisted convolutional neural network for heart disease prediction on the internet of medical things platform. *IEEE Access* **2020**, *8*, 189503–189512. [\[CrossRef\]](#)
6. Amato, F.; López, A.; Peña-Méndez, E.M.; Vaňhara, P.; Hampl, A.; Havel, J. Artificial neural networks in medical diagnosis. *J. Appl. Biomed.* **2013**, *11*, 47–58. [\[CrossRef\]](#)
7. Srinivasu, N.P.; Rao, S.T.; Srinivas, G.; Reddy, P.P.V.G.D. A Computationally Efficient Skull Scraping Approach for Brain MR Image. *Recent Adv. Comput. Sci. Commun.* **2020**, *13*, 833–844. [\[CrossRef\]](#)
8. Naga, S.P.; Rao, T.; Balas, V. A systematic approach for identifying tumor regions in the human brain through HARIS algorithm. In *Deep Learning Techniques for Biomedical and Health Informatics*; Academic Press: Cambridge, MA, USA, 2020; pp. 97–118.
9. Keleş, A.; Keleş, A.; Yavuz, U. Expert system based on neuro-fuzzy rules for diagnosis breast cancer. *Expert Syst. Appl.* **2011**, *38*, 5719–5726. [\[CrossRef\]](#)
10. Paydar, S.; Pourahmad, S.; Azad, M.; Bolandparvaz, S.; Taheri, R.; Ghahramani, Z.; Abbasi, H.R. The evolution of a malignancy risk prediction model for thyroid nodules using the artificial neural network. *Middle East J. Cancer* **2016**, *7*, 47–52.
11. Javed, F.; Venkatachalam, P.A.; Hani, A.F.M. Knowledge based system with embedded intelligent heart sound analyzer for diagnosing cardiovascular disorders. *J. Med. Eng. Technol.* **2007**, *31*, 341–350. [\[CrossRef\]](#)

12. Daecher, A.; Cotteleer, M.; Holdowsky, J. The Internet of Things: A Technical Primer. Deloitte, 2018. Available online: <https://www2.deloitte.com/us/en/insights/focus/internet-of-things/technical-primer.html> (accessed on 10 October 2022).
13. King, W. The ‘healthcare internet of things’. *Pharm. Exec.* **2017**, *37*, 34–35.
14. Rao, A.S.; Verweij, G. *Sizing the Prize: What’s the Real Value of AI for Your Business, and How Can You Capitalize*; PwC Publication: London, UK, 2017; Available online: <https://www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html> (accessed on 10 October 2022).
15. Sun, J.; Guo, Y.; Wang, X.; Zeng, Q. mHealth for aging China: Opportunities and challenges. *Aging Dis.* **2016**, *7*, 53. [CrossRef]
16. Rubí, J.N.S.; Gondim, P.R.L. IoMT Platform for Pervasive Healthcare Data Aggregation, Processing, and Sharing Based on OneM2M and OpenEHR. *Sensors* **2019**, *19*, 4283. [CrossRef]
17. Chen, M.; Hao, Y.; Hwang, K.; Wang, L.; Wang, L. Disease prediction by machine learning over big data from healthcare communities. *IEEE Access* **2017**, *5*, 8869–8879. [CrossRef]
18. Guleria, P.; Ahmed, S.; Alhumam, A.; Srinivasu, P.N. Empirical Study on Classifiers for Earlier Prediction of COVID-19 Infection Cure and Death Rate in the Indian States. *Healthcare* **2022**, *10*, 85. [CrossRef]
19. Marakhimov, A.; Joo, J. Consumer adaptation and infusion of wearable devices for Healthcare. *Comput. Hum. Behav.* **2017**, *76*, 135–148. [CrossRef]
20. Shahmiri, S. Wearing your data on your sleeve: Wearables, the FTC, and the privacy implications of this new technology. *Tex. Rev. Ent. Sport. L.* **2016**, *18*, 25.
21. Callahan, A.; Shah, N.H. Machine learning in Healthcare. In *Key Advances in Clinical Informatics*; Academic Press: Cambridge, MA, USA, 2017; pp. 279–291.
22. Chen, P.H.C.; Liu, Y.; Peng, L. How to develop machine learning models for Healthcare. *Nature Mater.* **2019**, *18*, 410–414. [CrossRef]
23. Qayyum, A.; Qadir, J.; Bilal, M.; Al-Fuqaha, A. Secure and robust machine learning for Healthcare: A survey. *arXiv* **2020**, arXiv:2001.08103. [CrossRef]
24. Vongsingthong, S.; Smanchat, S. Internet of things: A review of applications and technologies. *Suranaree J. Sci. Technol.* **2014**, *21*, 359–374.
25. Konstantinidis, E.I.; Antoniou, P.E.; Bamparopoulos, G.; Bamidis, P.D. A lightweight framework for transparent cross platform communication of controller data in ambient assisted living environments. *Inf. Sci.* **2015**, *300*, 124–139. [CrossRef]
26. Tjoa, E.; Guan, C. A survey on explainable artificial intelligence (xai): Toward medical xai. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *32*, 4793–4813. [CrossRef] [PubMed]
27. Pawar, U.; O’Shea, D.; Rea, S.; O’Reilly, R. Explainable ai in Healthcare. In Proceedings of the 2020 International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA), Dublin, Ireland, 15–19 June 2020; pp. 1–2.
28. Paul, S.M.V.; Balasubramaniam, S.; Panchatcharam, P.; Kumar, P.M.; Mubarakali, A. Intelligent Framework for Prediction of Heart Disease using Deep Learning. *Arab. J. Sci. Eng.* **2022**, *47*, 2159–2169. [CrossRef]
29. Absar, N.; Das, E.K.; Shoma, S.N.; Khandaker, M.U.; Miraz, M.H.; Faruque, M.R.I.; Tamam, N.; Sulieman, A.; Pathan, R.K. The Efficacy of Machine-Learning-Supported Smart System for Heart Disease Prediction. *Healthcare* **2022**, *10*, 1137. [CrossRef] [PubMed]
30. Muhammad, L.J.; Al-Shourbaji, I.; Haruna, A.A.; Mohammed, I.A.; Ahmad, A.; Jibrin, M.B. Machine learning predictive models for coronary artery disease. *SN Comput. Sci.* **2021**, *2*, 350. [CrossRef]
31. Hsu, Y.C.; Tsai, I.J.; Hsu, H.; Hsu, P.W.; Cheng, M.H.; Huang, Y.L.; Lin, C.Y. Using Anti-Malondialdehyde Modified Peptide Autoantibodies to Import Machine Learning for Predicting Coronary Artery Stenosis in Taiwanese Patients with Coronary Artery Disease. *Diagnostics* **2021**, *11*, 961. [CrossRef]
32. Arrieta, A.B.; Díaz-Rodríguez, N.; Del Ser, J.; Bennetot, A.; Tabik, S.; Barbado, A.; Herrera, F. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf. Fusion* **2020**, *58*, 82–115. [CrossRef]
33. Gilpin, L.H.; Bau, D.; Yuan, B.Z.; Bajwa, A.; Specter, M.; Kagal, L. Explaining explanations: An overview of interpretability of machine learning. In Proceedings of the 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA), Turin, Italy, 1–3 October 2018; pp. 80–89.
34. Lipton, Z.C. The mythos of model interpretability. *Queue* **2018**, *16*, 31–57. [CrossRef]
35. Bhatt, U.; Xiang, A.; Sharma, S.; Weller, A.; Taly, A.; Jia, Y.; Ghosh, J.; Puri, R.; Moura, J.M.; Eckersley, P. Explainable machine learning in deployment. In Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, Barcelona, Spain, 27–30 January 2020; pp. 648–657.
36. Fellous, J.M.; Sapiro, G.; Rossi, A.; Mayberg, H.; Ferrante, M. Explainable artificial intelligence for neuroscience: Behavioral neurostimulation. *Front. Neurosci.* **2019**, *13*, 1346. [CrossRef]
37. Dave, D.; Naik, H.; Singhal, S.; Patel, P. Explainable AI meets Healthcare: A Study on Heart Disease Dataset. *arXiv* **2020**, arXiv:2011.03195.
38. Calegari, R.; Ciatto, G.; Dellaluce, J.; Omicini, A. Interpretable Narrative Explanation for ML Predictors with LP: A Case Study for XAI. In WOA; 2019; pp. 105–112. Available online: <https://www.semanticscholar.org/paper/Interpretable-Narrative-Explanation-for-ML-with-LP%3A-Calegari-Ciatto/1e345972e7625c771554c15d362b98fd2e86d8f4> (accessed on 10 October 2022).
39. Porto, R.; Molina, J.M.; Berlanga, A.; Patricio, M.A. Minimum Relevant Features to Obtain Explainable Systems for Predicting Cardiovascular Disease Using the Statlog Data Set. *Appl. Sci.* **2021**, *11*, 1285. [CrossRef]

40. Aggarwal, R.; Podder, P.; Khamparia, A. ECG Classification and Analysis for Heart Disease Prediction Using XAI-Driven Machine Learning Algorithms. In *Biomedical Data Analysis and Processing Using Explainable (XAI) and Responsive Artificial Intelligence (RAI)*; Springer: Singapore, 2022; pp. 91–103.
41. Westerlund, A.M.; Hawe, J.S.; Heinig, M.; Schunkert, H. Risk prediction of cardiovascular events by exploration of molecular data with explainable artificial intelligence. *Int. J. Mol. Sci.* **2021**, *22*, 10291. [CrossRef]
42. Adadi, A.; Berrada, M. Peeking inside the black-box: A survey on Explainable Artificial Intelligence (XAI). *IEEE Access* **2018**, *6*, 52138–52160. [CrossRef]
43. Temenos, A.; Tzortzis, I.N.; Kaselimi, M.; Rallis, I.; Doulamis, A.; Doulamis, N. Novel Insights in Spatial Epidemiology Utilizing Explainable AI (XAI) and Remote Sensing. *Remote Sens.* **2022**, *14*, 3074. [CrossRef]
44. Guleria, P.; Sood, M. Explainable AI and machine learning: Performance evaluation and explainability of classifiers on educational data mining inspired career counseling. *Educ. Inf. Technol.* **2022**, *1*–36. [CrossRef]
45. Moreno-Sánchez, P.A. Development of an Explainable Prediction Model of Heart Failure Survival by Using Ensemble Trees. In Proceedings of the 2020 IEEE International Conference on Big Data (Big Data), Atlanta, GA, USA, 10–13 December 2020; pp. 4902–4910.
46. Graham, S.A.; Lee, E.E.; Jeste, D.V.; Van Patten, R.; Twamley, E.W.; Nebeker, C.; Depp, C.A. Artificial intelligence approaches to predicting and detecting cognitive decline in older adults: A conceptual review. *Psychiatry Res.* **2020**, *284*, 112732. [CrossRef]
47. Peng, J.; Zou, K.; Zhou, M.; Teng, Y.; Zhu, X.; Zhang, F.; Xu, J. An Explainable Artificial Intelligence Framework for the Deterioration Risk Prediction of Hepatitis Patients. *J. Med. Syst.* **2021**, *45*, 61. [CrossRef]
48. Moradi, M.; Samwald, M. Explaining black-box text classifiers for disease-treatment information extraction. *arXiv* **2020**, arXiv:2010.10873.
49. Muddamsetty, S.M.; Jahromi, M.N.S.; Moeslund, T.B. Expert level evaluations for explainable AI (XAI) methods in the medical domain. In *ICPR-2020 Workshop Explainable Deep Learning-AI*; Springer: Berlin/Heidelberg, Germany, 2020.
50. Srinivasu, P.N.; Sandhya, N.; Jhaveri, R.H.; Raut, R. From blackbox to explainable ai in healthcare: Existing tools and case studies. *Mob. Inf. Syst.* **2022**, *2022*, 8167821. [CrossRef]
51. Guidotti, R.; Monreale, A.; Ruggieri, S.; Turini, F.; Giannotti, F.; Pedreschi, D. A survey of methods for explaining black box models. *ACM Comput. Surv.* **2018**, *51*, 1–42. [CrossRef]
52. Yang, G.; Ye, Q.; Xia, J. Unbox the Black-box for the Medical Explainable AI via Multi-modal and Multi-centre Data Fusion: A Mini-Review, Two Showcases and Beyond. *arXiv* **2021**, arXiv:2102.01998. [CrossRef]
53. West, D.M. *The Future of Work: Robots, AI, and Automation*; Brookings Institution Press: Washington, DC, USA, 2018.
54. Ravi, V.; Narasimhan, H.; Pham, T.D. A cost-sensitive deep learning-based meta-classifier for pediatric pneumonia classification using chest X-rays. *Expert Syst.* **2022**, *39*, e12966. [CrossRef]
55. Ravi, V.; Narasimhan, H.; Chakraborty, C.; Pham, T.D. Deep learning-based meta-classifier approach for COVID-19 classification using CT scan and chest X-ray images. *Multimed. Syst.* **2022**, *28*, 1401–1415. [CrossRef]
56. Sharma, S.; Parmar, M. Heart diseases prediction using deep learning neural network model. *Int. J. Innov. Technol. Explor. Eng.* **2020**, *9*, 124–137. [CrossRef]
57. Yadav, P.; Menon, N.; Ravi, V.; Vishvanathan, S. Lung-gans. Unsupervised representation learning for lung disease classification using chest ct and x-ray images. *IEEE Trans. Eng. Manag.* **2021**, *1*–13. [CrossRef]
58. Pham, T.; Ravi, V.; Liu, N.; Luo, B.; Fan, C.; Sun, X. Tensor Decomposition of Largest Convolutional Eigenvalues Reveals Pathological Predictive Power of Rhob in Rectal Cancer Biopsy. 2021. Available online: <https://www.researchsquare.com/article/rs-957359/v1> (accessed on 10 October 2022). [CrossRef]
59. Gunning, D. Explainable artificial intelligence (xai). *Def. Adv. Res. Proj. Agency Web* **2017**, *2*, 1. [CrossRef]
60. Laborda, J.; Ryoo, S. Feature Selection in a Credit Scoring Model. *Mathematics* **2021**, *9*, 746. [CrossRef]
61. Vinh, L.T.; Lee, S.; Park, Y.T.; d'Auriol, B.J. A novel feature selection method based on normalized mutual information. *Appl. Intell.* **2012**, *37*, 100–120. [CrossRef]
62. Alhumam, A. Software Fault Localization through Aggregation-Based Neural Ranking for Static and Dynamic Features Selection. *Sensors* **2021**, *21*, 7401. [CrossRef]
63. Saha, S.; Priyoti, A.T.; Sharma, A.; Haque, A. Towards an Optimized Ensemble Feature Selection for DDoS Detection Using Both Supervised and Unsupervised Method. *Sensors* **2022**, *22*, 9144. [CrossRef]
64. Ertekin, S. *Learning in Extreme Conditions: Online and Active Learning with Massive, Imbalanced and Noisy Data*; The Pennsylvania State University: State College, PA, USA, 2009.
65. Amarappa, S.; Sathyaranayana, S.V. Data classification using Support vector Machine (SVM), a simplified approach. *Int. J. Electron. Comput. Sci. Eng.* **2014**, *3*, 435–445.
66. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]
67. Xia, T. Support vector machine based educational resources classification. *Int. J. Inf. Educ. Technol.* **2016**, *6*, 880. [CrossRef]
68. Son, Y.J.; Kim, H.G.; Kim, E.H.; Choi, S.; Lee, S.K. Application of support vector machine for prediction of medication adherence in heart failure patients. *Healthc. Inform. Res.* **2010**, *16*, 253–259. [CrossRef]
69. Ashish Kumar, The Ultimate Guide to AdaBoost Algorithm | What Is AdaBoost Algorithm? Available online: <https://www.mygreatlearning.com/blog/adaboost-algorithm> (accessed on 10 October 2022).
70. Ying, C.; Qi-Guang, M.; Jia-Chen, L.; Lin, G. Advance and prospects of AdaBoost algorithm. *Acta Autom. Sin.* **2013**, *39*, 745–758.

71. AdaBoost. Wikipedia. Available online: <https://en.m.wikipedia.org/wiki/AdaBoost> (accessed on 10 October 2022).
72. Smith, K. *Precalculus: A Functional Approach to Graphing and Problem Solving*; Jones & Bartlett Publishers: Burlington, MA, USA, 2013; p. 8. ISBN 978-0-7637-5177-7.
73. Cohen, D.; Lee, T.; Sklar, D. *Precalculus: A Problems-Oriented Approach*, 6th ed.; Thomson-Brooks/Cole: Belmont, CA, USA, 2005; p. 698. ISBN 978-0-534-40212-9.
74. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140. [CrossRef]
75. Yaman, E.; Subasi, A. Comparison of bagging and boosting ensemble machine learning methods for automated EMG signal classification. *BioMed Res. Int.* **2019**, *2019*, 9152506. [CrossRef]
76. Truett, J.; Cornfield, J.; Kannel, W. A multivariate analysis of the risk of coronary heart disease in Framingham. *J. Chronic Dis.* **1967**, *20*, 511–524. [CrossRef]
77. Hosmer, D.W., Jr.; Lemeshow, S.; Sturdivant, R.X. *Applied Logistic Regression*; John Wiley & Sons: Hoboken, NJ, USA, 2013; p. 398.
78. Murty, M.N.; Devi, V.S. *Pattern Recognition: An Algorithmic Approach*; Springer Science & Business Media: Berlin, Germany, 2011.
79. Srinivasu, P.N.; JayaLakshmi, G.; Jhaveri, R.H.; Praveen, S.P. Ambient Assistive Living for Monitoring the Physical Activity of Diabetic Adults through Body Area Networks. *Mob. Inf. Syst.* **2022**, *2022*, 3169927. [CrossRef]
80. Narkhede, S. Understanding auc-roc curve. *Towards Data Sci.* **2018**, *26*, 220–227.
81. Kernbach, J.M.; Staartjes, V.E. Foundations of Machine Learning-Based Clinical Prediction Modeling: Part II—Generalization and Overfitting. In *Machine Learning in Clinical Neuroscience*; Springer: Cham, Switzerland, 2022; pp. 15–21.
82. Megna, R.; Petretta, M.; Assante, R.; Zampella, E.; Nappi, C.; Gaudieri, V.; Mannarino, T.; D'Antonio, A.; Green, R.; Cantoni, V.; et al. A Comparison among Different Machine Learning Pretest Approaches to Predict Stress-Induced Ischemia at PET/CT Myocardial Perfusion Imaging. *Comput. Math. Methods Med.* **2021**, *2021*, 3551756. [CrossRef] [PubMed]
83. Sharma, D.; Gotlieb, N.; Farkouh, M.; Patel, K.; Xu, W.; Bhat, M. Machine Learning Approach to Classify Cardiovascular Disease in Patients with Nonalcoholic Fatty Liver Disease in the UK Biobank Cohort. *J. Am. Heart Assoc.* **2022**, *11*, e022576. [CrossRef] [PubMed]
84. Jiang, Y.; Zhang, X.; Ma, R.; Wang, X.; Liu, J.; Keerman, M.; Yan, Y.; Ma, J.; Song, Y.; Zhang, J.; et al. Cardiovascular Disease Prediction by Machine Learning Algorithms Based on Cytokines in Kazakhs of China. *Clin. Epidemiol.* **2021**, *13*, 417–428. [CrossRef] [PubMed]
85. Patro, S.; Padhy, N.; Chiranjivi, D. Ambient assisted living predictive model for cardiovascular disease prediction using supervised learning. *Evol. Intell.* **2020**, *14*, 941–969. [CrossRef]
86. Yuval, M.; Yaman, B.; Tosun, Ö. Classification Comparison of Machine Learning Algorithms Using Two Independent CAD Datasets. *Mathematics* **2022**, *10*, 311. [CrossRef]