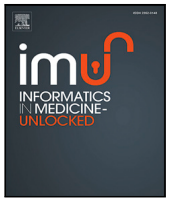
Informatics in Medicine Unlocked 49 (2024) 101535

Contents lists available at ScienceDirect

Informatics in Medicine Unlocked

journal homepage: www.elsevier.com/locate/imu



A one-dimensional convolutional neural network-based deep learning approach for predicting cardiovascular diseases

Dhafer G. Honi *, Laszlo Szathmary

University of Debrecen, Department of IT, H-4002 Debrecen, Hungary

ARTICLE INFO

Dataset link: https://archive.ics.uci.edu/dataset/45/heart+disease

Keywords:
Deep learning
Machine learning
Neural networks

ABSTRACT

Early detection of cardiovascular diseases (CVDs) is crucial for managing cardiovascular diseases and improving patient outcomes. Deep neural networks have the potential to reduce the reliance on costly and time-consuming clinical tests, leading to cost savings for patients and healthcare systems. This study proposes the development of specialized convolutional neural networks for the automated selection of essential variables, employing various preprocessing procedures. It evaluates the approach using the UCI repository heart disease dataset, focusing on early-stage heart disease identification to enhance early prediction and intervention for CVD. To address the challenge of achieving higher accuracy, we introduce an approach using one-dimensional convolutional neural networks, incorporating extensive testing to optimize the network architecture and enhance predictive performance. Additionally, recognizing the impact of features on accuracy, a comprehensive data analysis was performed. Through a meticulous selection process, we identified and utilized key features that significantly influenced the accuracy of our model, contributing to more reliable predictions. Finally, cross-validation techniques were implemented to precisely evaluate the efficacy of our work. Numerous experiments were conducted to demonstrate the relevance of our research. The prediction accuracy was found to be 99.95% when employing a train–test approach, while it was approximately 98.53% when employing K-Fold cross-validation. In comparison to existing literature, our approach outperforms a recent best study that proposed a Catboost model, achieving an F1-score of about 92.3% and an average accuracy of 90.94%. This signifies a substantial improvement in predictive performance, with a percentage improvement of approximately 9.90% compared to the Catboost model.

1. Introduction

Cardiovascular diseases (CVDs) present a considerable health challenge worldwide, requiring attention from the global community. According to the World Health Organization (WHO), CVDs are responsible for approximately 18 million deaths worldwide every year, causing them to have the highest mortality rate. The failure to promptly diagnose these conditions and the lack of effective monitoring technologies are responsible for approximately 32% of deaths worldwide. WHO, as mentioned in reports from 2015, highlights that CVDs are accounted for the deaths of roughly 18 million people [1,2]. Additionally, heart attacks and stroke-related complications account for a staggering 85% of these fatalities. In the United States, there is an occurrence of cardiovascular-associated death every 34 s [3,4]. Consequently, mitigating the risk of complications related to CVDs and improving patient outcomes is crucial [5]. Predicting the onset of CVDs. can assist healthcare administrations in preventing or mitigating its impact. Organizations can leverage their human resources (HR) data to develop predictive models aimed at anticipating cardiovascular disease occurrences. In recent years, the benefits of artificial intelligence applications have emerged as valuable tools across various sectors, including healthcare, education, economics and administration [6,7]. Consequently, the use of artificial intelligence methods in predicting cardiovascular disease has attracted significant research attention. Additionally, the increasing accessibility of data related to this field has spurred further research endeavors in this area [8,9]. Cardiovascular disease, also known as heart disease, is the leading cause of death worldwide. The heart muscle is responsible for maintaining blood flow throughout the body [10]. Applying machine learning techniques has shown fascinating results in predicting certain medical conditions. However, these techniques have not yet been employed to predict the individual survival of hypertensive patients with CVD using regularly collected and imposing digital administrative health data [11]. It may be possible to optimize the use of accumulated datasets to support in predicting patient outcomes, planning individualized patient care,

* Corresponding author.
E-mail addresses: dhafer.honi@inf.unideb.hu (D.G. Honi), szathmary.laszlo@inf.unideb.hu (L. Szathmary).

https://doi.org/10.1016/j.imu.2024.101535
Received 3 April 2024; Received in revised form 19 June 2024; Accepted 19 June 2024
Available online 22 June 2024

monitoring resource utilization, and improving institutional performance if a machine learning algorithm can be used to exploit the large administrative dataset. It could be possible to maximize the use of collected datasets to help with patient outcome prediction, personal care planning, resource management, and institutional performance enhancement [12].

This research addresses several key challenges in the management of CVD and applying machine learning techniques for early detection and prediction. Firstly, there is a pressing need for improved methods of early detection and prediction of CVD, as current approaches may not adequately identify the disease in its early stages when symptoms are not apparent [13,14]. Additionally, existing literature suggests a gap in the utilization of machine learning techniques for predicting individual CVD survival, particularly in hypertensive patients using routinely obtained big digital electronic administrative health data [15,16]. This highlights a critical research problem in leveraging advanced computational methods to analyze large datasets for better prognostic evaluation and treatment decisions. Moreover, the research aims to address challenges related to feature selection and algorithm optimization, recognizing the need for appropriate techniques to enhance prediction accuracy and support early intervention [17]. Finally, there is a broader need to provide reliable advice to health and medical specialists regarding significant changes in the healthcare sector, particularly in the context of adopting machine learning techniques for CVD management [18,19]. These research problems underscore the importance of the study's objectives in advancing the understanding and application of machine learning in the field of cardiovascular health.

In this research study, we propose an optimized Convolutional Neural Network (CNN) model for predicting CVD and identifying the most influential factors based on their correlation set values. Subsequently, various machine learning and deep learning techniques have been employed to analyze CVD. Our main contributions can be summarized as follows.

1. In the initial phase of our study, we applied a deep learning technique complemented by preprocessing steps to enhance the accuracy of predicting CVD, showcasing the model's details and workings for accurate predictions. This will assist doctors in determining predictive outcomes and diagnosing heart disease effectively.
2. Dataset features are analyzed to reveal their correlation with each other and to identify the most important features.
3. To ensure reliable outcomes, we tested our model using both cross-validation and split validation. We employed a deep neural networks algorithm with a focus on accuracy performance for predicting CVD.
4. To the best of our knowledge, we are the first to introduce optimized prediction accuracy achieved through the careful selection of relevant features. Early detection of cardiovascular diseases (CVD) is facilitated by a robust deep neural network algorithm. This approach enables timely intervention and enhances the identification of critical features to support recovery algorithms.
5. The proposed solution for CVD is worth considering because of limited dataset and this solution is to work well in this type of situations.

The remaining sections of this paper are organized in the following manner: Section 2 introduces the techniques and methods discussed in the literature. Section 3 expounds upon the methodology that was employed. Section 4 details the experimental evaluation. Section 5 comprises a discussion and comparison of the results. Finally, Section 6 concludes the paper.

2. Related work

Many academics use diverse data mining approaches to analyze a range of heart disease expectation models. Besides contributing for evaluating results and possible future research on the scheme, they used datasets and other computations to provide more beneficial outcomes. In order to develop effective methods and high accuracy for detecting cardiac problems, researchers conducted comprehensive and precise studies.

Pattekari et al. [20] propose a heart disease prediction system that utilizes the Naïve Bayes data mining approach, while Tran [21] introduces a web-based intelligent system for heart disease diagnosis using the same technique. Both systems collect information from users through predefined questions and extract hidden patterns from datasets. By comparing user responses against these datasets, they provide solutions to complex inquiries about heart disease and enable more informed clinical decisions. Although these systems offer significant advancements in predicting and diagnosing heart disease, there are limitations to consider. One limitation is their reliance on predefined questions that may not capture all relevant information from users, potentially leading to misdiagnosis or overlooking important symptoms. Furthermore, although the Naïve Bayes algorithm is efficient, it assumes independence between features, oversimplifying complex interactions within heart disease data. Therefore, while promising, further research is necessary to enhance accuracy and reliability for clinical practice. Additionally, Gonsalves et al. [22] leveraged historical medical data to predict Coronary Heart Disease (CHD) using Naïve Bayes along with other machine learning techniques. While Naïve Bayes is known for its ease of implementation and computational efficiency, its assumption of feature independence may not hold true in complex medical datasets, potentially limiting its predictive power compared to more sophisticated algorithms. Shouman et al. [23] proposed a hybrid approach that integrates K-means clustering with Naïve Bayes to enhance the precision of heart disease diagnosis. This hybrid model seeks to bolster Naïve Bayes' performance by preprocessing the data with K-means clustering, identifying clusters of similar instances to reduce noise and improve predictive accuracy. However, several considerations limit its efficacy. Firstly, employing K-means clustering as a preprocessing step may risk information loss or data structure distortion, especially in intricate medical datasets with high dimensions. Secondly, the effectiveness of K-means clustering heavily relies on selecting the appropriate number of clusters (K), which can be subjective and may not always yield optimal clustering. Additionally, while noise reduction through clustering can enhance predictive accuracy, it may inadvertently discard valuable information or introduce bias, potentially compromising the model's generalization capability. Elma et al. [24] applied a new classifier that combines the distance-based algorithm K-Nearest Neighbors and statistical-based Naïve Bayes Classifier. By integrating these two techniques, the model benefits from both local and global information, potentially improving its ability to capture complex patterns in the data and enhance predictive performance. However, there are some potential drawbacks to this approach. Firstly, integrating K-Nearest Neighbors (KNN) with Naïve Bayes introduces additional computational complexity, especially in large datasets, which may limit scalability and real-time processing capabilities. Secondly, combining KNN with Naïve Bayes assumes that both algorithms complement each other, but this may not always hold true, especially if their underlying assumptions or principles contradict each other. Finally, the hybrid approach may increase model complexity and require careful parameter tuning, making it challenging to interpret and optimize effectively. Dulhare [25] enhanced heart disease prediction techniques by incorporating of particle swarm optimization and Naïve Bayes. While optimizing Naïve Bayes parameters using particle swarm optimization (PSO) can improve predictive accuracy, it may also lead to overfitting if not regularized or validated appropriately. PSO is sensitive to initialization and parameter settings, and convergence to the global

optimum is not guaranteed, which may result in suboptimal model performance. Additionally, the computational cost of PSO may be prohibitive, particularly in large-scale datasets, and the improvement in predictive accuracy may not justify the additional computational burden.

Gnaneswar and Jebarani [26] introduced a forward neural network architecture to develop a predictive model for heart rate, taking into account the cycling cadence. In this context, heart rate and cadence serve as the primary data sources. The outcome of this model is the forecasted heart rate for the subsequent second. By utilizing a feed-forward neural network structure, the statistical relationship between heart rate and cycling cadence is effectively captured and represented. The performance of neural networks excels at capturing nonlinear relationships and significantly depends on data quality, preprocessing techniques that used on the dataset, and suitable hyperparameter optimization. Many hyperparameters in a neural network can contribute to achieving the best accuracy for the proposed model.

Mutijarsa et al. [27] proposed enhanced healthcare services. In light of these decisions, it is essential to take into account the substantial developments in remote communication technology that have been made for the prevention of heart illness. Data mining techniques have shown to be incredibly useful in this situation for the localization and diagnosis of cardiac disorders. A thorough comparison examination of numerous single and hybrid data mining algorithms is conducted in their assessment with the aim of determining the algorithm that exhibits the best degree of accuracy in predicting coronary disease.

Yeshvendra et al. [28] and Tripoliti et al. [29] conducted research on the use of ensemble methods, specifically random forests, for predicting heart disease. The study in [28] found that utilizing machine learning algorithms could significantly improve prediction accuracy by integrating feature selection with random forest methodology to achieve a successful predictive rate of 85.81%. These findings potentially suggest early identification and intervention in high-risk individuals can help save lives through this effective technology application. In [29] authors presented a novel ensemble learning technique that breaks away from traditional methods by adapting the number of classifiers in real-time during forest development. This dynamic approach guarantees both diversity and accuracy within the group, making it an advantageous classifier option. The effectiveness of this method was evaluated across eight biomedical datasets utilizing different random forests algorithm variations; remarkable achieving 90% classification accuracy for identifying tree count. Nevertheless, while ensembles often deliver high prediction precision rates, they come with potential concerns such as excessive computational workload or overfitting when imbalanced data is present.

Patil et al. [30] proposed a novel decision support system (DSS) that incorporates three data mining techniques: Classical Random Forest (CRF), Modified Random Forest (MRF), and Weighted Random Forest (WRF). CRF builds a collection of trees, whereas MRF constructs the tree dynamically using an online fitting procedure. WRF introduces a weight factor to improve prediction accuracy. Random forests represent a significant modification of bagging, involving a three-step forest development process. Oikonomou et al.. [31] introduced a novel method for predicting cardiac risk through the analysis of the radiomic profile of coronary perivascular adipose tissue (PVAT). Employing a machine learning algorithm, specifically a random forest, the study trained and validated a fat radiomic profile (FRP) signature. Subsequent testing in two patient cohorts demonstrated that the FRP signature outperformed traditional risk stratification methods in predicting major adverse cardiac events (MACE). Notably, the FRP signature exhibited stability, remaining unchanged six months after an acute myocardial infarction, in contrast to the variability observed in the FAI signature. Ibrahim et al. [32] proposed a multi-label active learning (MLLAL) model that enhances heart disease prediction by selecting informative unlabeled instances for expert labeling, reducing labeling costs. Their study compares five MLLAL strategies, showing improved model performance compared to passive learning methods. The optimal strategy varies based on evaluation metrics, highlighting MLLAL's potential in enhancing predictive accuracy and efficiency for preventive care. Pratiyush et al. [33] a novel ensemble classification model was proposed within an Explainable AI (XAI) framework to predict heart disease from the cardiovascular dataset. The dataset comprised 303 instances and 14 attributes, and the classification techniques employed were Support Vector Machine (SVM), AdaBoost, K-nearest neighbor (KNN), bagging, logistic regression (LR), and Naïve Bayes. The outcomes of each algorithm were compared, and the performance of the XAI-based classification models was found to be superior to conventional classification models. Nadiahet et al. [34] proposed a Catboost model built on effective machine learning techniques that is intended for the automatic identification of essential characteristics and the early diagnosis of heart disease.

Peter and Somasundaram [35] and Javeed et al. [36] proposed various feature selection techniques that aim to identify the most relevant attributes for heart disease prediction. In [35] authors introduced a method that combines the Bayes Theorem with CFS for feature selection, offering a comparative analysis against traditional methods like the conventional algorithm and the hybrid Algorithm CFS+FilterSubsetEval. On the other hand, in [36] authors presented a feature selection technique named Floating Window Adaptive Size Elimination (FWAFE), employing both artificial neural networks (ANN) and deep neural networks (DNN) as classification frameworks after the feature reduction. While effective feature selection can enhance model interpretability and generalization, it may inadvertently overlook crucial but less obvious predictors, potentially resulting in suboptimal performance.

Rupali [37] proposed the Heart Disease Prediction System (HDPS) that employs Laplacian smoothing to approximate significant patterns in the data while mitigating noise. Ali et al. [38] proposed a two-stage system for diagnosing heart disease. The first stage uses a linear and L1 regularized SVM to eliminate irrelevant features from the dataset. The second stage uses an L2 regularized SVM with a linear or RBF kernel to classify patients as having or not having heart disease. The authors manually tune the hyperparameters of the SVM models to optimize performance. Samuel et al. [39] proposed an enhanced heart failure (HF) risk prediction system that integrates the fuzzy analytic hierarchy process (Fuzzy_AHP) technique with artificial neural networks (ANNs). The system first utilizes Fuzzy_AHP to determine the relative contributions of 13 commonly used HF attributes. These weights are then incorporated into the ANN model, which is trained on a dataset of 297 HF patients. Ali and Bukhari [40] proposed a two-stage decision support system to improve the generalization performance of HF risk prediction models. The first stage utilizes a mutual information-based statistical model to identify the most relevant features from the dataset. The second stage employs a neural network to train on the selected features and make predictions. The proposed system was evaluated on the HF subset of the Cleveland heart disease database and demonstrated superior performance compared to other HF risk prediction models, achieving a mean percent error (MPE) of 8.8% and an accuracy rate of 93.33%.

Verma et al. [41] and Alaa et al. [42] proposed hybrid machine learning approaches that integrate feature selection, clustering, and classification algorithms for diagnosing heart disease. In [41] authors introduced a hybrid approach specifically tailored for diagnosing Coronary Artery Disease (CAD), which combines risk factor identification, feature selection, and classification. They used CFS selection with PSO search method and K-means clustering to identify relevant features, followed by training four classification algorithms (MLP, MLR, FURIA, and C4.5) using these selected features. Their method achieved high prediction accuracy, with MLR reaching 88.4% on one dataset and 89.6% on another. In [42] authors proposed a hybrid method aimed at improving coronary heart disease (CHD) prediction. Their approach

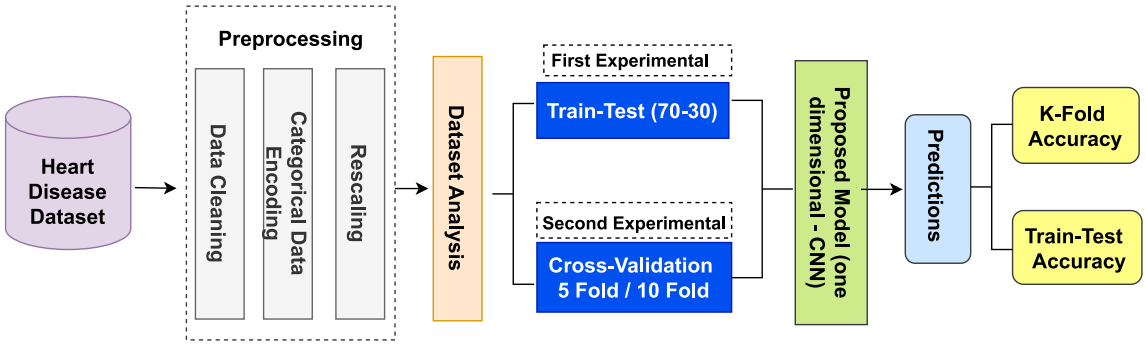


Fig. 1. The proposed methodology.

employs recursive feature elimination (RFE) to identify relevant features for CHD prediction, followed by three classification algorithms (logistic regression, decision tree, and Support Vector Machine) applied to these selected features. A voting mechanism then combines the predictions of these algorithms to generate the final CHD prediction. The method proposed in [42] outperformed individual classification algorithms on two datasets, showcasing the effectiveness of hybrid approaches in enhancing heart disease diagnosis. Despite the advantages of hybrid approaches in leveraging complementary strengths of different techniques, they may introduce additional complexity in model interpretation and implementation.

The literature has witnessed a growing effort to develop methods for predicting CVD. However, traditional machine learning algorithms often struggle to incorporate textual data into their predictive processes due to its high dimensionality [43–47]. To address this shortcoming and enable early CVD detection and timely interventions, this paper explores and develops a suite of reliable machine learning techniques.

While existing solutions show potential, there is opportunity for improvement in their predictive accuracy to further strengthen confidence in these models. To achieve this goal, the proposed research utilizes convolutional neural networks in conjunction with data preprocessing techniques and cross-validation (5-Fold to 10-Fold) and train–test (70–30) partitions.

3. Our methodology

The methodology of our study consists of several key steps aimed to comprehensively analyzing heart disease prediction. Firstly, the dataset utilized for this study was obtained from the UCI Heart Disease dataset [48]. Next, we preprocessed the data to ensure its quality and suitability for analysis. This involved missing values, outliers, and encoding categorical data to facilitate accurate and meaningful analysis. We performed feature selection and scaling to further refine the dataset and prepare it for in-depth analysis. Following preprocessing, we conducted a dataset analysis of the dataset to identify the main features associated with Cardiovascular Disease (CVD) using statistical and graphical techniques. Subsequently, we developed and trained a Convolutional Neural Network (CNN) model with the preprocessed data. In our proposed model, we utilized the Rectified Linear Unit (ReLU) function, defined as $f(x) = \max(0, x)$, as the hidden layer transfer function. The selection of ReLU function was based on its efficient and effective performance in deep neural network models, particularly in avoiding the vanishing gradient problem. Additionally, we designed the architecture of our CNN model with a specific number of hidden layers to balance complexity and computational efficiency. We fine-tuned hyperparameters to optimize performance and used the model to predict CVD. We optimized prediction accuracy by selecting the best model configuration based on validation results, as discussed in Section 3.5 to include a more detailed discussion of the model architecture, including the rationale behind the number of layers and the choice of transfer functions.

Table 1
UCI Heart disease dataset features.

No	Type	Feature name
1	Age	Age in years
2	Sex	Male = 1; Female = 0
3	Cp	Chest pain type (4 values)
4	Trestbps	Resting blood pressure (in mm Hg on admission to the hospital)
5	Chol	Serum cholesterol in mg/dl
6	Fbs	Fasting blood sugar >120 mg/dl (1 = true; 0 = false)
7	Restecg	Resting electrocardiographic results
8	Thalach	Maximum heart rate achieved
9	Exang	Exercise induced angina (1 = yes; 0 = no)
10	Oldpeak	ST depression induced by exercise relative to rest
11	Slope	The slope of the peak exercise ST segment
12	Ca	Number of major vessels (0–3) colored by fluoroscopy
13	Thal	1 = normal; 2 = fixed defect; 3 = reversible defect
14	Target	0 = no disease and 1 = disease

After making predictions, we evaluated the model's performance and interpreted the results. We identified factors influencing CVD risk and prediction accuracy, comparing the model's predictions with actual outcomes. We calculated the error rate and generated a confusion matrix to assess model performance. Finally, we measured the accuracy of the model using two metrics: K-Fold Cross Validation Accuracy and Train–Test Accuracy. We compared the accuracy scores of our model with those of other existing models and methods for predicting CVD, as discussed in Section 4.

The suggested methodology is divided into three phases: (1) preprocessing, (2) training, and (3) classification as shown in Fig. 1.

3.1. Dataset description

This study utilized the UCI Heart Disease dataset [48]. The dataset comprises 14 attributes of 1025 patients. Table 1 outlines the properties of the dataset and their corresponding categories. Fig. 2 depicts the target variable, indicating if a patient has heart disease ("Yes" or "No"). The "Target" feature reflects this distinction.

3.2. Implementation platform

During the experiments of the study, we used DeepNote[1] as the implementation platform for deep learning models. DeepNote is a cloud-based Jupyter notebook environment that offers unrestricted entry to computing resources at no cost. It includes a virtual machine with 16 GB of RAM and up to 100 GB of hard disk space. The virtual

[1] https://deepnote.com.

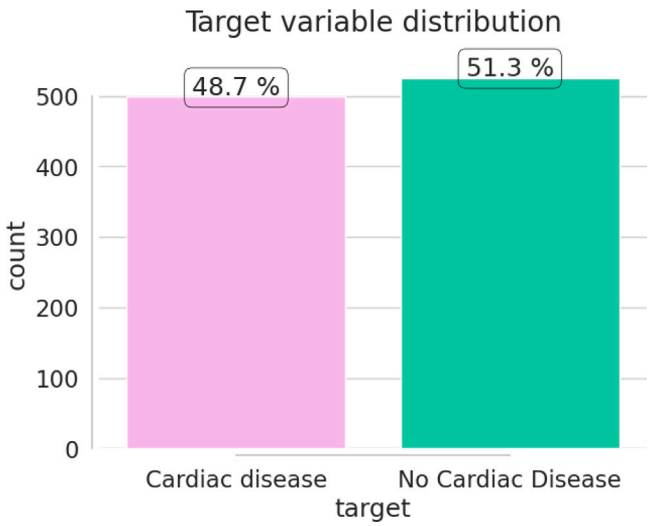


Fig. 2. Target variable distribution.

machine runs on DeepNote's servers and enables researchers to train deep learning models efficiently. Additionally, DeepNote provides a wide range of preinstalled libraries and tools, making it easy to install and use the necessary dependencies. These libraries include Numpy and Pandas for efficient data manipulation, Tensorflow and Keras for building and training deep learning models. Finally, visualization capabilities are enhanced through Seaborn and Matplotlib for plotting graphs and charts.

3.3. Preprocessing

Data preprocessing is a crucial step in machine learning and deep learning models, significantly impacting model performance. It encompasses three primary tasks: data cleansing, categorical data encoding, and rescaling. These processes are briefly addressed in the following subsections.

3.3.1. Data cleaning

Upon analyzing the dataset, it was discovered that certain variables, such as ID, social security number (ccf), chest pain location (painloc) and patient names, were identical for all patients. These redundant features were consequently removed at this stage. Furthermore, since the values of the additional 58 attributes had no bearing on our classification task, they were also excluded from further analysis.

3.3.2. Categorical data encoding

Some of the dataset's features do not have numerical values but rather categorical (nominal) values. Most machine learning and deep learning algorithms cannot directly handle categorical features. To address this issue, we need to convert these features into numerical representations. There are several categorical features in the original dataset, including sex, chest pain type, fasting blood sugar >120 mg/dL, resting electrocardiographic results, exercise-induced angina, the slope of the peak exercise ST segment, Thalassemia, and target. One-hot encoding, a technique that first identifies the unique values and their corresponding counts, is employed to transform these categorical features into numerical representations. Each unique value is then assigned a one-hot binary vector. For instance, the sex feature, which has two values (Male, Female), is converted into the binary vectors (1, 0) and (0, 1), respectively.

3.3.3. Rescaling

During our study, some factors have wide ranges, which can make them seem more important than others. This can lead to inconsistent results when categorizing data. To fix this, we need to make sure all factors are on the same scale. We do this by rescaling their values to fit within the range [0,1]. This makes it easier to compare and analyze the data accurately. In our dataset, we apply this rescaling to all attributes to ensure uniformity and consistency in our analysis. Eq. (1) displays the normalizing formula :

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}}, \tag{1}$$

where $X_{\min}$ and $X_{\max}$ are the minimum and maximum values of the given feature, respectively.

3.4. Dataset analysis

The correlation matrix is usually used to understand the relationship among the dataset features. Fig. 3 depicts the correlation matrix of our dataset. The cell colors vary from blue to red color. Grey cells represent no correlation, while red variations represent high correlation. Blue variations represent a negative correlation among dataset features.

The correlation analysis revealed significant relationships between various cardiovascular features. "Thalach" exhibited a strong positive correlation with "slope" (Fig. 4a) and with "cp" (Fig. 4b). Likewise, "exang" showed a positive correlation with "oldpeak" (Fig. 4c). Furthermore, "age" demonstrated significant positive correlations with "trestbps" (Fig. 4d) and "ca" (Fig. 5a). "Chol" displayed a moderate positive correlation with "age" (Fig. 5b). Moreover, "ca" exhibited a positive association with texttt"oldpeak" (Fig. 5c). Finally, "age" showed a moderate correlation with "oldpeak" (Fig. 5d).

In our prediction process, we aimed to evaluate the significance of different features, employing the Chi-square χ ranking method for this purpose. This method is particularly suitable for our task due to its simplicity, interpretability, effectiveness with nominal and ordinal data, sensitivity to feature interactions, scalability to large datasets, and consistency with other feature selection methods [49,50]. The results of our analysis provided valuable insights into the importance of specific features. According to the Chi-square ranking, **cp**, **ca**, **oldpeak**, and **thalach** emerged as the most influential features in shaping our predictive model, as visually depicted in Fig. 6. This indicates that these particular features have a significant impact on the accuracy and effectiveness of our predictive analysis. Interpreting these results, a higher importance assigned to certain features suggests that they carry more predictive power in determining the outcome of our model. In other words, variations in these features have a stronger influence on the predicted outcomes compared to other features. This underscores the critical role of importance features in shaping the predictive performance of our model. This result is particularly interesting because it provides beneficial insights into which features contribute the most to the predictive accuracy of our model. By focusing on the most relative importance of different features, we can prioritize them in our analysis and potentially refine our predictive model to achieve even better performance. Additionally, identifying these key features can aid in the interpretation of the model's predictions and provide valuable guidance for decision-making in real-life applications.

Firstly, the model utilizes Convolutional layers,

3.5. Proposed architecture

The prediction model is one of the most important part of any successful prediction model, particularly in the field of CVDs. A variety of machine learning techniques have been employed for classifying CVDs, including decision trees, random forests, Naïve Bayes, logistic regression, and SVM. In this study, a convolutional neural networks prediction model is used to predict CVDs. To avoid overfitting and

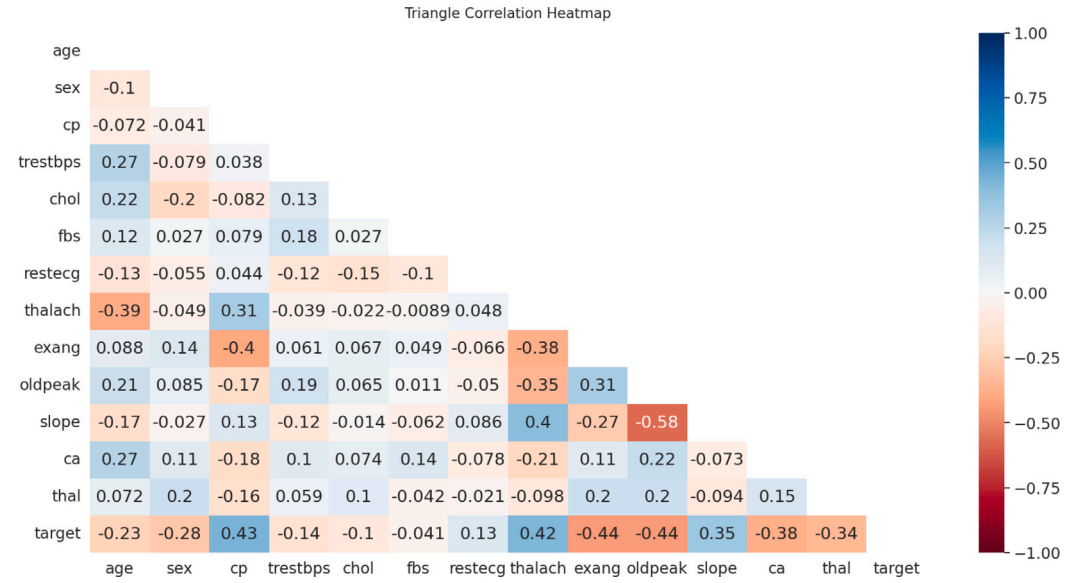


Fig. 3. Triangle correlation heat-map. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

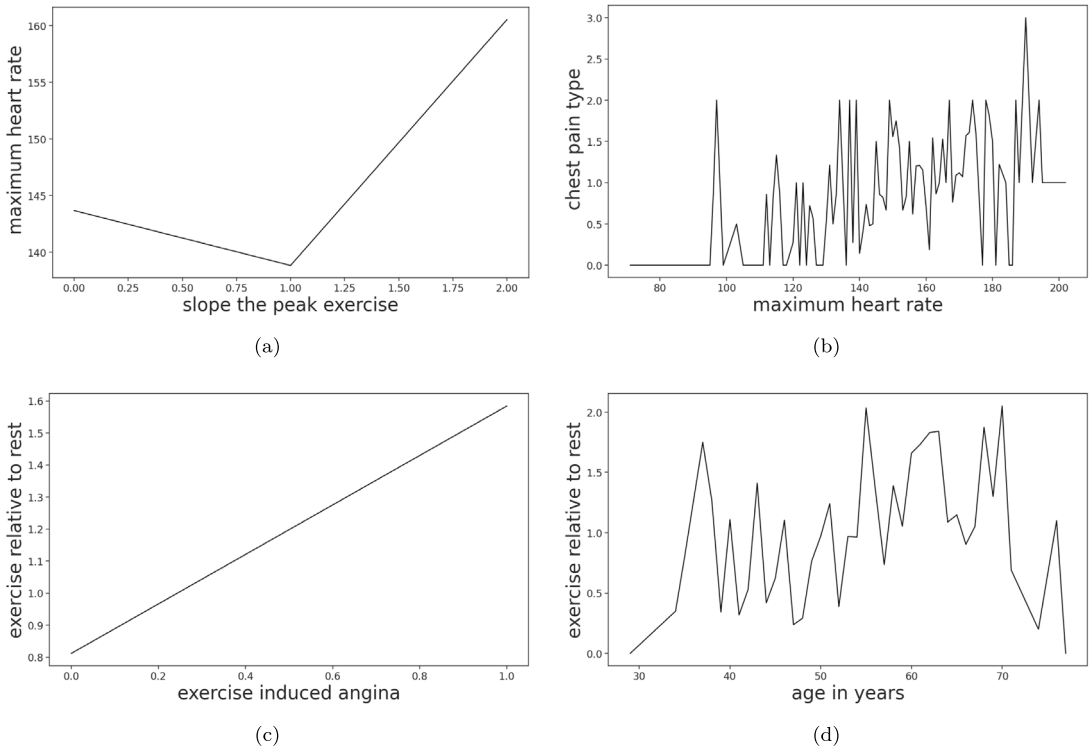


Fig. 4. Relation among several features. (a) correlation between thalach and slope. (b) correlation between thalach and cp. (c) correlation between exang and oldpeak. (d) correlation between age and trestbps.

underfitting, the hyperparameters of the model must be carefully chosen. Firstly, the model utilizes Convolutional layers, which extract important features from the input data through convolution operations with a specified filter. The output is calculated by performing element-wise multiplication between the filter weights and the input values within a sliding window [51]. In addition, batch normalization layers are essential for minimizing the dependency of the activation function on scaling parameters or initial values. By regulating excessively large values into a narrower range, these layers ensure a more stable parameter distribution in each layer, leading to faster learning rates [52]. Dropout layers are a technique used to prevent overfitting in neural networks. During training, the idea is that each neuron is temporarily disabled with a probability of p. This means that all inputs and outputs to the neuron will also be disabled during that iteration. The dropped-out neurons are then resampled with probability p at every training step, so a neuron that was disabled at one step can be active at the next one. The hyperparameter p is commonly referred to as the *dropout-rate* and is usually set to around 0.5, resulting in 50% of the neurons being dropped out [53]. Moreover, flatten layers play a crucial role in neural network architectures by reducing the dimensionality of the data. This reduction leads to fewer parameters in the subsequent fully connected layers, which is important for managing model complexity and avoiding overfitting [54]. Hidden layers work as intermediary phases that connect input and output in a neural network. When too

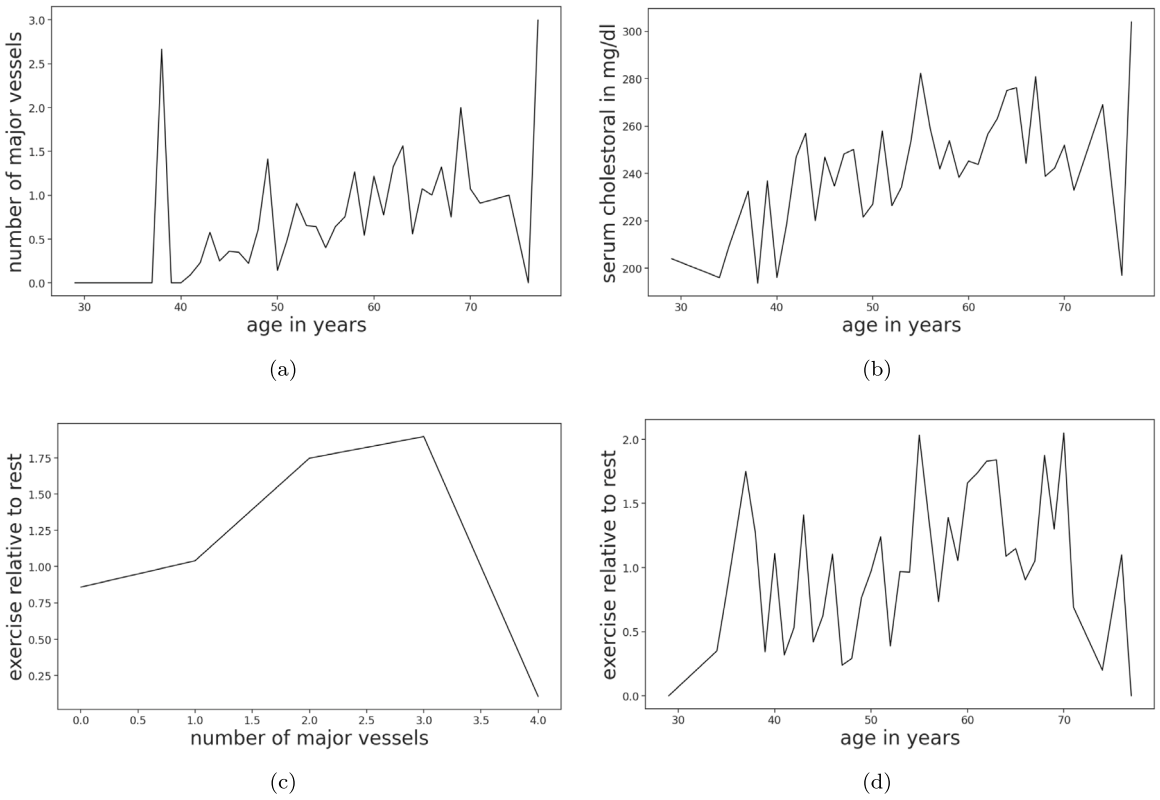


Fig. 5. Relation among several features. (a) correlation between age and ca. (b) correlation between chol and age. (c) correlation between ca and oldpeak. (d) correlation between age and oldpeak.

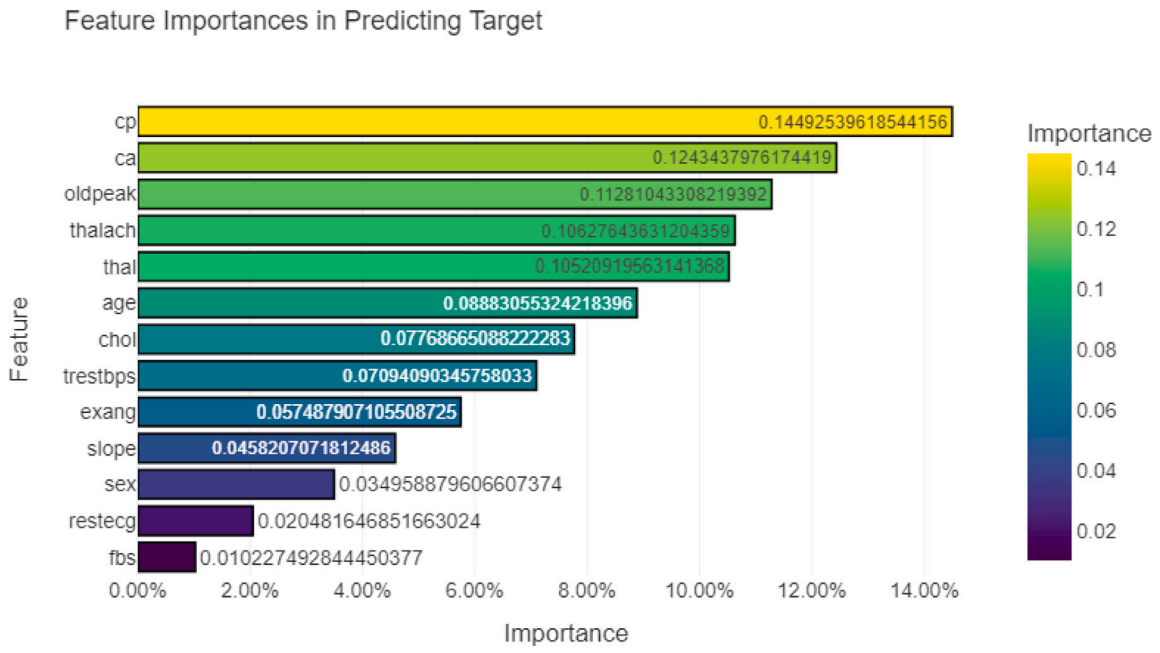


Fig. 6. Features importance.

few neurons are employed in these hidden layers, underfitting may occur, as the network struggles to efficiently detect signals within a complex dataset. Conversely, overfitting may occur when there are too many neurons in the hidden layers [55]. Number of neurons in a neural network architecture plays a crucial role in determining the complexity and performance of the architecture. Increasing the number of neurons generally improves the model's ability to learn complex patterns in the data. However, excessive neurons can lead to overfitting [56]. Therefore, activation functions and other relevant parameters should also be selected meticulously.

The resulting model comprises an input layer that receives the preprocessed cardiovascular data. This data is then fed into the convolutional layers, which extract features using filters of varying sizes (32, 64, 128). The extracted features are then passed through batch

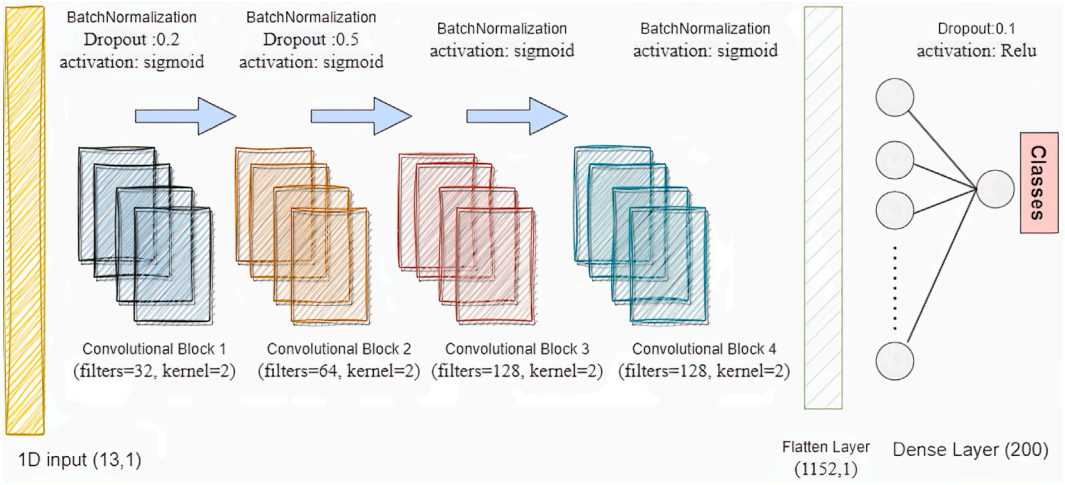
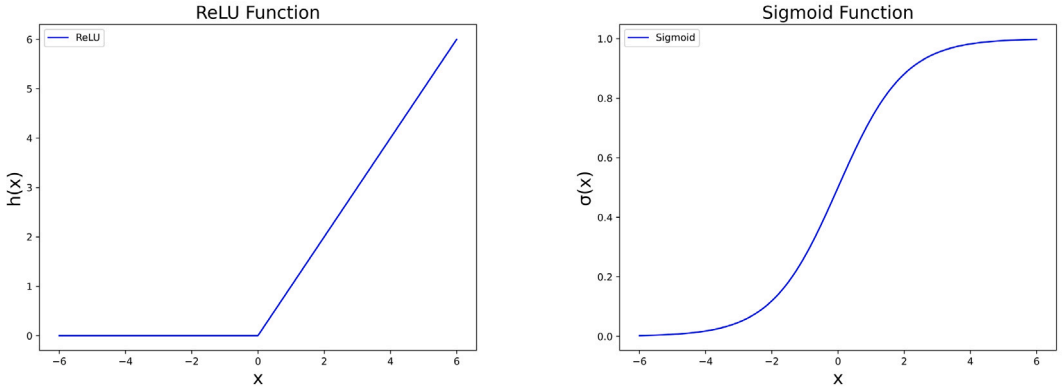


Fig. 7. Proposed one-dimensional-convolutions-neural-network architecture.



(a) Relu function (b) sigmoid function

Fig. 8. Activation functions: (a) ReLU function. (b) sigmoid function.

normalization layers to normalize the activations and ensure smoother and more stable training. To prevent overfitting, dropout layers are applied after each convolutional layer, discarding 20%, 50%, and 10% of the units, respectively.

The extracted features are then flattened into a vector, which is passed through two dense layers. The first dense layer has 200 neurons. For each neuron, the rectified linear unit function (ReLU), as shown in Fig. 8a, is applied to compute the output. This function returns the weighted sum of the input data if it is positive, effectively passing through positive values unchanged, and outputs zero if the input is negative. This activation function introduces non-linearity to the model, enabling it to learn complex patterns and relationships in the data, while the second dense layer has a single neuron with a sigmoid activation as shown in Fig. 8b. The sigmoid activation function outputs a probability between 0 and 1, indicating the likelihood of whether a patient has CVDs. The proposed architecture is depicted in Fig. 7. Table 2 summarizes the model's specifications.

Table 2
Proposed network architecture.

Layer type	Output shape	Number of parameters	Activation function
Input layer	(13, 1)	96	Sigmoid
Batch normalization	(13, 32)	128	–
Dropout	(13, 32)	0	–
Convolutional layer	(12, 32)	2080	Sigmoid
Batch normalization	(12, 64)	256	–
Dropout	(12, 64)	0	–
Convolutional layer	(11, 64)	8256	Sigmoid
Batch normalization	(11, 128)	512	–
Convolutional layer	(10, 128)	32 896	Sigmoid
Batch normalization	(10, 128)	512	–
Flatten layer	1280	0	–
Dense layer	200	256 200	ReLU
Dropout	200	0	–
Dense layer (Output layer)	1	201	Sigmoid

3.6. Justification of the proposed architecture

The reason for this proposed architecture lies in several factors that were considered. Firstly, domain expertise played a crucial role, as the proposal stemmed from a thorough understanding of the domain, particularly in CVDs prediction. Through extensive research and consultation with domain experts, it was identified that CNNs have shown promising results in predictions, medical image analysis and classification tasks, making them a suitable choice for analyzing complex cardiovascular data. Additionally, the proposal was informed by a comprehensive literature review in the field of deep learning and CVDs. By analyzing previous studies and methodologies, insights were gained into the strengths and limitations of different approaches, leading to the identification of CNNs as a potential solution. Furthermore, experimental exploration played a significant role in the development of the proposed architecture. It involved iterative experimentation and

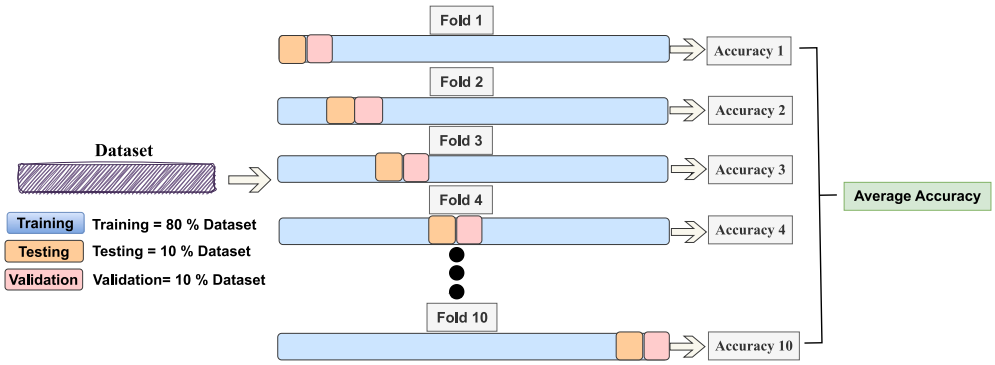


Fig. 9. Schematic overview of K-Fold cross-validation method.

exploration of various model configurations, where different combinations of layers, activation functions, and hyperparameters were tested and evaluated using validation techniques to identify the most effective architecture for CVD predictions.

3.7. Development of the proposed architecture

The proposed architecture was developed through a careful process that began with a comprehensive justification rooted in domain expertise and extensive literature review. Based on insights from the field of CVD prediction, CNNs appeared as a suitable solution. Experimentation played a crucial part in improving the architecture, with various configurations tested to optimize predictive performance.

3.8. Validation

Prediction models are evaluated by partitioning the dataset into two distinct sets: a training set and a testing set. This study employs train–test sets and K-Fold cross-validation for validation.

1. Train–test validation sets: In this method, 70% of the dataset is allocated for training the model, while the remaining 30% serves as the validation set.
   The 70/30 train–test split is a common practice in machine learning and statistics. It involves dividing the dataset into 70% for training the model and 30% for evaluating its performance. This split allows the model to learn from a substantial portion of the data while also having a separate set to assess its generalization [57]. And the other reason to use these splits is to achieve a balance between ample training data 70% for effective model learning and a substantial test set 30% for reliable performance evaluation. This approach is versatile and widely applicable to various machine learning tasks. However, the optimal ratio depends on factors such as dataset size, model complexity and the specific task. In limited datasets, a larger training proportion may be suitable, while larger test sets might be necessary for complex models to ensure accurate generalization assessment [58,59].
2. K-Fold cross-validation
   Train–test sets present a potential issue in that they may not always provide a fair assessment of the model's performance. This is because including the testing samples in the training set can lead to an overly optimistic evaluation. To obtain a more realistic assessment and avoid overfitting, cross-validation is required. This technique involves dividing the dataset into k partitions, with each partition serving as the testing set in one iteration while the remaining $k-1$ partitions are used for training. This process is repeated k times [49]. An average of the accuracy values obtained over these k iterations is used to determine the final accuracy as shown in Fig. 9.

Table 3
Model performance metrics.

	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)	Support
Training	99.94	99.96	99.98	99.96	150
Testing	99.93	99.96	99.97	99.95	158

4. Experimental results

To evaluate the performance of our model, we performed two experiments: a train–test split and a K-Fold cross-validation.

4.1. First experiment

The dataset was split into two sets, 70% for training and the remaining 30% for testing. Various performance metrics such as accuracy, precision, recall and F1-score were used to evaluate the classification effectiveness of 1D convolutional neural networks (1D-CNN). The equations in [(2), (3), (4), (5)] provide the definitions of these metrics. The results of the CNN classifier are summarized in Table 3.

$$\text{Accuracy} = \frac{TP + TN}{\text{Total Sample}} \tag{2}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{3}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{4}$$

$$\text{F1-Score} = \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{5}$$

TP: the test correctly identifies a patient with cardiovascular disease.
FP: the test incorrectly identifies a patient without cardiovascular disease as having the disease.
TN: the test correctly identifies a patient without cardiovascular disease.
FN: the test incorrectly identifies a patient with cardiovascular disease as not having the disease.

The results demonstrated that our model outperformed all other methods listed in Table 4 in terms of precision, recall, F1-score, and accuracy. This result can be credited to the strong predictive abilities of deep learning, especially its skill in finding complex patterns in the data, along with our careful preprocessing steps like normalization, scaling features, and reducing noise. During the 150 epochs of training, the model showed impressive results achieving 99.96% accuracy in training, 99.94% precision, 99.96% recall, and 99.98% F1-score. This indicates that the model effectively identified patterns within the training dataset and successfully generalized its learnings to new data instances. Similarly, on the testing dataset, the model exhibited strong performance, achieving an accuracy of 99.95%, precision of 99.93%,

Table 4
Comparative results on the Dataset using ML and DL.

Author	Methods	Accuracy	Precision	Recall	F1-Score
Peter and Somas [35]	Naïve Bayes	80.37	–	–	–
	Multilayer perceptron	85.18	–	–	–
	KNN	85.55	–	–	–
Shouman et al [23]	Naïve Bayes	84.5	–	–	–
Rupali [37]	Laplacian smoothing	86	–	–	–
Elma et al [24]	Naïve Bayes and KNN	85.92	–	–	–
	BoostFSNB	82.42	–	–	–
	AdaboostC4.5	78.79	–	–	–
	BaggingC4.5	79.86	–	–	–
	Bagging NB	83.24	–	–	–
	Adaboost NB	83.14	–	–	–
Dulhare [25]	Particle swarm optimization with NB	87.91	–	–	–
Nadiah [34]	Catboost_tuned	90.94	93.17	91.46	92.31
Javeed et al [36]	Feature selection based on FWAFE + ANN	91.11	–	–	–
	Feature selection based on FWAFE + DNN	93.33	–	–	–
Ali et al [38]	Stacked SVMs	92.22	–	–	–
Samuel et al [39]	ANN-fuzzy-AHP	91.10	–	–	–
Ali and Bukhari [40]	Mutual information + DNN	90	–	–	–
Verma et al [41]	CFS-PSO	88.4	–	–	–
Alaa et al [42]	Naïve Bayes	84	–	–	–
	SVM	82	–	–	–
Proposed method	1D CNN (train–test)	99.95	99.93	99.96	99.97
	1D CNN (K-Fold = 5)	98.53	99	99.92	99.91
	1D CNN (K-Fold = 10)	99.12	99.94	99.93	99.98

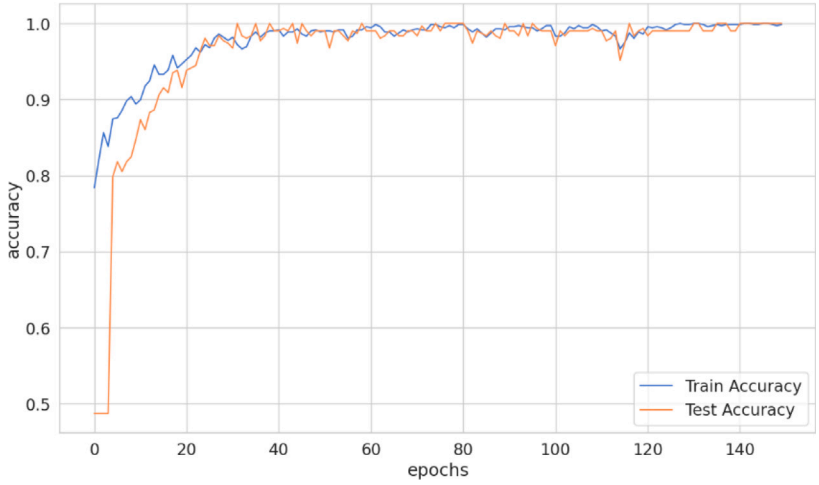


Fig. 10. Training and validation accuracy per epoch.

recall of 99.96%, and F1-score of 99.97%. While most researchers still use traditional machine learning techniques, some are starting to use more advanced methods such as ANN, deep learning, and CNN as their primary models. The accuracy of our proposed approach surpasses these methods employing the same dataset. Fig. 10 illustrates the model's overall performance throughout the 150 epochs, demonstrating consistent high accuracy in both training and testing phases. Additionally, discernible trends in the accuracy curve, such as an initial rapid increase followed by a stabilization, are observable. Fig. 11 illustrates the reduction in the model's loss throughout the training and validation phases. As the model underwent training, its loss gradually diminished, indicative of its improved ability to predict values that closely align with the actual ones. The extent of the loss reduction can be highlighted to emphasize the model's learning progress and enhanced predictive performance. Fig. 12, depicted as a confusion matrix, provides a visual representation of the model's prediction performance for two classes. It offers insights into the frequency of correctly classified instances, including true positives, true negatives, false positives, and false negatives. Specific values within the confusion matrix can be analyzed to interpret the model's accuracy in identifying positive cases accurately.

The accuracy and loss curves are crucial for assessing the learning progress and performance of a machine learning model over time. They provide valuable insights into the model's effectiveness and evolution throughout the training process.

4.2. Second experiment

To evaluate the effectiveness of the proposed model, we employed both 10-Fold and 5-Fold cross-validation procedures. By applying these cross-validation techniques to the original dataset, we obtained performance outcomes that were consistent and reliable. When compared to other approaches that utilized the same dataset [48], our model outperformed them in terms of accuracy. Our model achieved an accuracy of 98.53% under the 5-Fold cross-validation scenario and an impressive

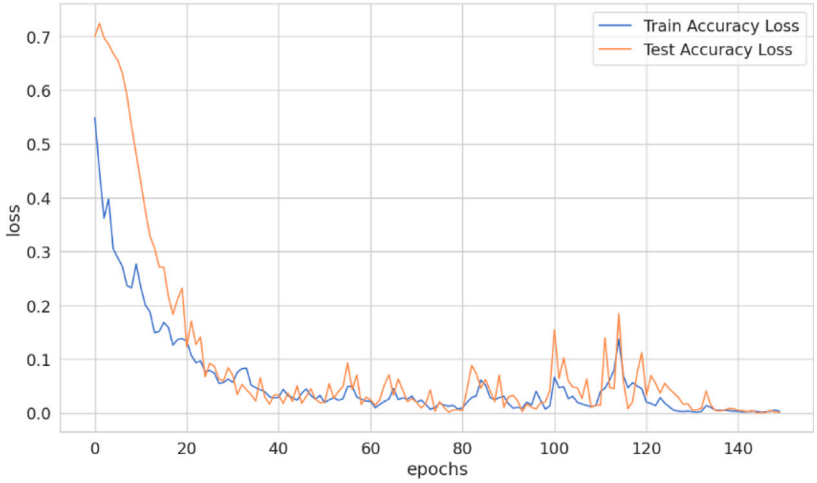


Fig. 11. Training and validation losses per epoch.

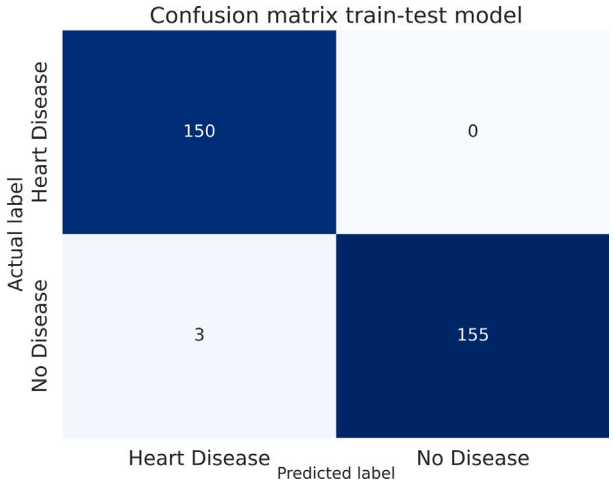


Fig. 12. Confusion matrix train–test model.

99.12% under the 10-Fold cross-validation scenario. The results of our proposed technique are presented in Table 4.

For a visual representation of the classification model's performance, Fig. 13 depicts the confusion matrices for both the 5-Fold and 10-Fold models. These visual representations provide valuable insights into the model's performance in different cross-validation scenarios.

In Fig. 14, we present the accuracy of our model across multiple iterations of cross-validation. This visualization depicts the performance of the model not only on the training data but also on the validation data, providing a holistic view of its predictive capabilities throughout the cross-validation process. By examining the accuracy trends over different folds of the cross-validation, we gain insights into the model's consistency and generalization ability across diverse subsets of the dataset. Additionally, analyzing the accuracy on both training and validation sets allows us to assess potential overfitting or underfitting issues and fine-tune the model accordingly to improve its overall performance.

Additionally, Fig. 15 visually shows how our model's loss changes over time during both training and validation. This graph helps us see how well our model learns from the training data and how it performs on new, unseen data. By looking at these curves, we can understand if our model is getting better at making predictions without getting too good at just memorizing the training data. It is like watching a learning process unfold and knowing when our model is doing well and when it might need some adjustments to improve further.

5. Discussion

Our proposed 1D-CNN model is systematically compared to various methodologies such as traditional machine learning approaches, advanced machine learning and ensemble techniques, and hybrid methods in terms of effectiveness. Firstly, various studies in the traditional ML domain, including those carried out by Peter and Somas [35], Shouman et al. [23], Rupali [37], Elma et al. [24] and Dulhare [25] have achieved remarkable accuracy ranging from 80.37% to 87.l1% utilizing methods such as Naïve Bayes, Multilayer Perceptron, and KNN. However, despite their noteworthy performance capabilities, these methods often lack precision when compared with our proposed lD-CNN model, which capitalize on deep learning's ability to capture intricate patterns in data. Secondly, advanced ML and ensemble techniques, researchers such as Nadiah. [34], Javeed et al. [36], and Ali et al. [38] have demonstrated impressive accuracies of up to 93.33% with methods such as Catboost_tuned, FWAFE + ANN, FWAFE + DNN, and Stacked SVMs. Despite their successes, our 1D CNN models outshine them across all metrics, including accuracy, precision, recall, and F1-Score, showcasing the efficacy of deep learning in cardiovascular disease prediction. Furthermore, Hybrid approaches, as explored by Samuel et al. [39], Ali and Bukhari [40], Verma et al. [41], and Alaa et al. [42], have also shown promise, with accuracies ranging from 84% to 91.10%. Nevertheless, our 1D-CNN models exhibit enhanced predictive capabilities, surpassing these hybrid methods and underscoring the effectiveness of deep learning in capturing nuanced data relationships. Finally, our proposed 1D-CNN model consistently outperforms all referenced methods across various evaluation metrics, achieving outstanding accuracies of up to 99.95%. Our models demonstrate potential in their ability to predict and generalize, especially evident in K-Fold cross-validation scenarios.

6. Conclusions

Cardiovascular disease prediction is crucial for assisting doctors in diagnosing patients early. Deep learning will not replace clinicians but rather complement the clinical portfolio by enhancing clinical decisions made by humans. In addition, a number of expensive clinical and laboratory tests will no longer be necessary, saving money for both patients and the healthcare system. This is achieved using deep learning techniques. The overall objective of this research is to employ 1D-CNN to improve the prediction accuracy of heart failure. The research employs two distinct validation methods. The first technique utilizes a 1D-CNN with a train–test split and achieves an exceptional classification accuracy of 99.95%. In the second approach, a 1D-CNN is divided into either 5-Fold or 10-Fold using cross-validation. Additionally, an

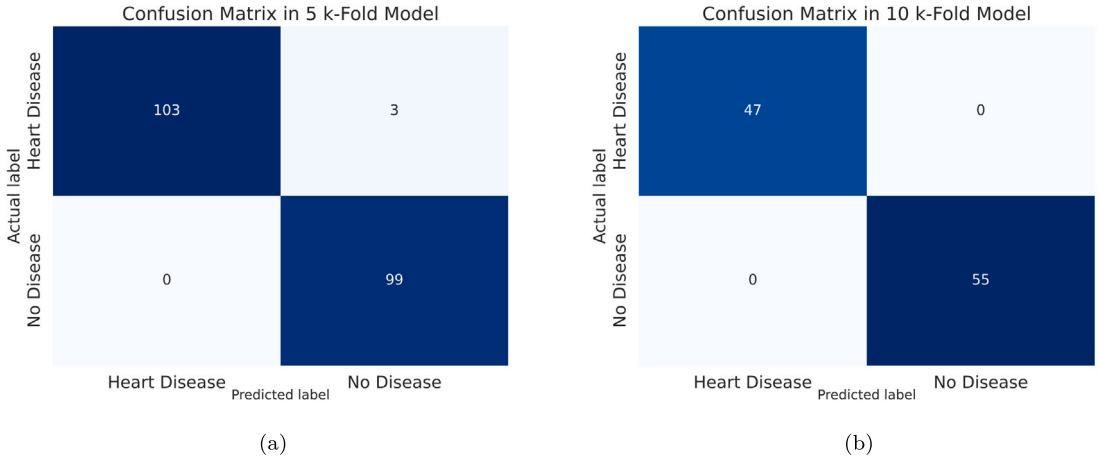


Fig. 13. Confusion matrices for (a) 5-Fold, (b) 10-Fold cross-validation models.

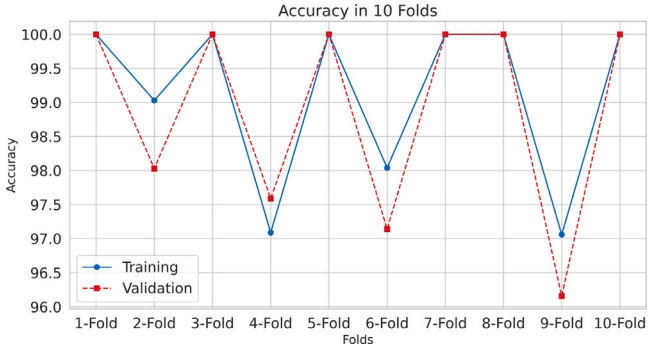


Fig. 14. Accuracy for training and validation.

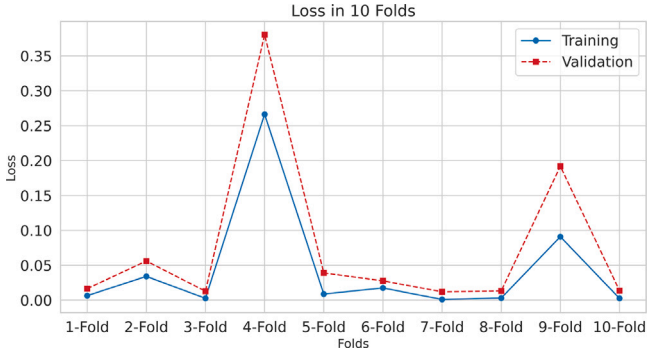


Fig. 15. Loss performance for training and validation.

accuracy of 99.12% is achieved through 5-Fold cross-validation, while an accuracy of 98.53% is obtained from 10-Fold cross-validation. Our method has undergone comprehensive evaluation based on accuracy, precision, recall, and F1-score. The results of the experiment unequivocally demonstrate that our proposed approach performs better than any previously reported approaches using the same dataset. While our study demonstrates promising results with the 1D-CNN approach, it is essential to acknowledge certain limitations. Firstly, inappropriate or biased datasets can result in unworthy model performance or even exacerbate existing biases in predictions. Additionally, the reliance on electronic health record (EHR) data may raise concerns about completeness and accuracy. Further exploration of the long-term effectiveness and impact on clinical decision-making and patient outcomes is also warranted. Secondly, the critical aspect of the neural network model is the choice of hyperparameters. In our model, important parameters including convolutional layers, hidden layers, filter sizes, and dropout rates introduce a degree of subjectivity and require careful tuning to optimize performance. Addressing these limitations is crucial for the responsible and effective utilization of machine learning in cardiovascular disease prediction.

CRediT authorship contribution statement

Dhafer G. Honi: Writing – original draft, Validation, Software, Methodology, Investigation, Data curation, Conceptualization. **Laszlo Szathmary:** Writing – review & editing, Validation, Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Availability of data and material

All the data used in this study are available at https://archive.ics.uci.edu/dataset/45/heart+disease.

References

[1] Kutrani H, Eltalhi S. Cardiac catheterization procedure prediction using machine learning and data mining techniques. IOSR J Comput Sci 2019;21(1):86–92.
[2] Saba T, Mohamed A, Rehman A, Almazyad AS, Fati SM. Image enhancement and segmentation techniques for detection of knee joint diseases: A survey. Curr Med Imaging 2018;14(5):704–15. http://dx.doi.org/10.2174/1573405613666170912164546.
[3] Shah K, Sharma K, Saxena D. Health technology assessment in cardiovascular diseases. Front Cardiovasc Med 2023;10:1108503. http://dx.doi.org/10.3389/fcvm.2023.1108503.
[4] Rahim MSM, Ali MA, Rahman MM, et al. 3D bones segmentation based on CT images visualization. Biomed Res 2017;28(8):3641–4.
[5] Ali OT, Nassif AB, Capretz LF. Business intelligence solutions in healthcare: transforming OLTP system to BI solution. In: 2013 third international conference on communications and information technology. ICCIT, IEEE; 2013, p. 209–14, Available at: http://dx.doi.org/10.1109/ICCITechnology.2013.6579551.
[6] Varian H. Artificial intelligence, economics, and industrial organization. In: Agrawal A, Gans J, Goldfarb A, editors. The economics of artificial intelligence: an agenda. Chicago: University of Chicago Press; 2019, p. 399–422.
[7] Vardarlier P, Zafer C. Use of artificial intelligence as business strategy in recruitment process and social perspective. In: Hacioglu U, editor. Digital business strategies in blockchain ecosystems. Contributions to management science, Cham: Springer; 2020, p. 265–78. http://dx.doi.org/10.1007/978-3-030-29739-8_17.

[8] Alduayj SS, Rajpoot K. Predicting employee attrition using machine learning. In: 2018 international conference on innovations in information technology. IIT, Al Ain, United Arab Emirates; 2018, p. 93–8. http://dx.doi.org/10.1109/INNOVATIONS.2018.8605976.
[9] Usha PM, Balaji NV. Analyzing employee attrition using machine learning. Karpagam J Comput Sci 2019;13:277–82.
[10] World Health Organization. WHO consultation to adapt influenza sentinel surveillance systems to include COVID-19 virological surveillance: virtual meeting, 6–8 October 2020. Tech. rep. WHO/WHE/GIH/GIP/2021.1, 2022.
[11] Ambale-Venkatesh B, Yang X, Wu CO, Liu K, Hundley WG, McClelland R, Gomes AS, Folsom AR, Shea S, Guallar E, Bluemke DA, Lima JAC. Cardiovascular event prediction by machine learning: the multi-ethnic study of atherosclerosis. Circ Res 2017;121(9):1092–101. http://dx.doi.org/10.1161/CIRCRESAHA.117.311312.
[12] Feng Y, Leung AA, Lu X, et al. Personalized prediction of incident hospitalization for cardiovascular disease in patients with hypertension using machine learning. BMC Med Res Methodol 2022;22:325. http://dx.doi.org/10.1186/s12874-022-01814-3.
[13] Haïssaguerre M, Derval N, Sacher F, Jesel L, Deisenhofer I, et al. Sudden cardiac arrest associated with early repolarization. N Engl J Med 2008;358(19):2016–23. http://dx.doi.org/10.1056/NEJMoa071968, PMID: 18463377.
[14] Kumar PM, Lokesh S, Varatharajan R, Babu GC, Parthasarathy P, et al. Cloud and IoT based disease prediction and diagnosis system for healthcare using fuzzy neural classifier. Future Gener Comput Syst 2018;86:527–34. http://dx.doi.org/10.1016/j.future.2018.04.036.
[15] Ambale-Venkatesh B, Yang X, Wu CO, Liu K, Hundley WG, et al. Cardiovascular event prediction by machine learning: the multi-ethnic study of atherosclerosis. Circ Res 2017;121(9):1092–101. http://dx.doi.org/10.1161/CIRCRESAHA.117.311312, Epub 2017 Aug 9. PMID: 28794054; PMCID: PMC5640485.
[16] Ali AH, Jasim HM, Abduljabbar ZA, Nyangaresi VO, Umran SM, Ma J, Honi DG. Provably efficient and fast technique for determining the size of a brain tumor in T1 MRI images. In: 2024 international conference on artificial intelligence in information and communication. ICAIIC, IEEE; 2024, p. 608–13. http://dx.doi.org/10.1109/ICAIIC60209.2024.10463420.
[17] Esfahani HA, Ghazanfari M. Cardiovascular disease detection using a new ensemble classifier. In: 2017 IEEE 4th international conference on knowledge-based engineering and innovation. KBEI, Tehran, Iran; 2017, p. 1011–4. http://dx.doi.org/10.1109/KBEI.2017.8324946.
[18] Malki Z, Atlam E, Dagnew G, Alzighaibi AR, Ghada E, Gad I. Bidirectional residual LSTM-based human activity recognition. Comput Inf Sci 2020;13(3):40. http://dx.doi.org/10.5539/cis.v13n3p40.
[19] Malki Z, Atlam E-S, Hassanien AE, Dagnew G, Elhosseini MA, Gad I. Association between weather data and COVID-19 pandemic predicting mortality rate: Machine learning approaches. Chaos Solitons Fractals 2020;138:110137. http://dx.doi.org/10.1016/j.chaos.2020.110137.
[20] Pattekari SA, Parveen A. Prediction system for heart disease using Naïve Bayes. Int J Adv Comput Math Sci 2012;3(3):290–4.
[21] Tran H. A survey of machine learning and data mining techniques used in multimedia system. Tech. rep., Richardson, TX, USA: Dept. Comput. Sci. Univ. Texas Dallas Richardson; 2019.
[22] Gonsalves AH, Thabtah F, Mohammad RMA, Singh G. Prediction of coronary heart disease using machine learning: an experimental analysis. In: Proceedings of the 2019 3rd international conference on deep learning technologies. ICDLT'19, Xiamen, China. New York: Association for Computing Machinery; 2019, p. 51–6. http://dx.doi.org/10.1145/3342999.3343015, 2019.
[23] Shouman M, Turner T, Stocker R. Integrating decision tree and k-means clustering with different initial centroid selection methods in the diagnosis of heart disease patients. In: Proceedings of the international conference on data science. ICDATA, WorldComp; 2012.
[24] Ferdousy EZ, Islam MM, Matin MA. Combination of Naive Bayes classifier and K-nearest neighbor (cNK) in the classification based predictive models. Comput Inf Sci 2013;6(3):48. http://dx.doi.org/10.5539/cis.v6n3p48.
[25] Dulhare UN. Prediction system for heart disease using Naive Bayes and particle swarm optimization. Biomed Res 2018;29(12):2646–9.
[26] Gnaneswar B, Jebarani MRE. A review on prediction and diagnosis of heart failure. In: 2017 international conference on innovations in information, embedded and communication systems. ICIIECS, Coimbatore, India; 2017, p. 1–3. http://dx.doi.org/10.1109/ICIIECS.2017.8276033.
[27] Mutijarsa K, Ichwan M, Utami DB. Heart rate prediction based on cycling cadence using feedforward neural network. In: 2016 international conference on computer, control, informatics and its applications. IC3INA, Tangerang, Indonesia; 2016, p. 72–6. http://dx.doi.org/10.1109/IC3INA.2016.7863026.
[28] Singh YK, Sinha N, Singh SK. Heart disease prediction system using random forest. In: Singh M, Gupta P, Tyagi V, Sharma A, Ören T, Grosky W, editors. Advances in computing and data sciences. ICACDS 2016. Commun comput inf sci., Singapore: Springer; 2017, p. 721. http://dx.doi.org/10.1007/978-981-10-5427-3_63.
[29] Tripoliti EE, Fotiadis DI, Manis G. Automated diagnosis of diseases based on classification: dynamic determination of the number of trees in random forests algorithm. IEEE Trans Inf Technol Biomed 2012;16(4):615–22. http://dx.doi.org/10.1109/TITB.2011.2175938, PMID: 28895575.
[30] Patil PR, Kinariwala SA. Automated diagnosis of heart disease using random forest algorithm. Int J Adv Res Ideas Innov Technol 2017;4(1):64–7.
[31] Oikonomou EK, Williams MC, Kotanidis CP, Desai MY, Marwan M, Antonopoulos AS, et al. A novel machine learning-derived radiotranscriptomic signature of perivascular fat improves cardiac risk prediction using coronary CT angiography. Eur Heart J 2019;40(43):3529–43. http://dx.doi.org/10.1093/eurheartj/ehz592.
[32] El-Hasnony IM, Elzeki OM, Alshehri A, Salem H. Multi-label active learning-based machine learning model for heart disease prediction. Sensors 2022;22(3):1184. http://dx.doi.org/10.3390/s22031184.
[33] Guleria P, Naga Srinivasu P, Ahmed S, Almusallam N, Alarfaj FK. XAI framework for cardiovascular disease prediction using classification techniques. Electronics 2022;11(24):4086. http://dx.doi.org/10.3390/electronics11244086.
[34] Baghdadi NA, Farghaly Abdelaliem SM, Malki A, et al. Advanced machine learning techniques for cardiovascular disease early detection and diagnosis. J Big Data 2023;10:144. http://dx.doi.org/10.1186/s40537-023-00817-1.
[35] Peter TJ, Somasundaram K. Study and development of novel feature selection framework for heart disease prediction. Int J Sci Res Publ 2012;2(10):1–7.
[36] Javaid A, Zghyer F, Kim C, Spaulding EM, Isakadze N, Ding J, et al. Medicine 2032: the future of cardiovascular disease prevention with machine learning and digital health technology. Med Am J Prev Cardiol 2022;12:100379. http://dx.doi.org/10.1016/j.ajpc.2022.100379, PMID: 36090536; PMCID: PMC9460561.
[37] Patil RR. Heart disease prediction system using Naive Bayes and Jelinek-Mercer smoothing. Int J Adv Res Comput Commun Eng 2014;3(5):6787–9.
[38] Ali L, Niamat A, Khan JA, Golilarz NA, Xingzhong X, Noor A, et al. An optimized stacked support vector machines based expert system for the effective prediction of heart failure. IEEE Access 2019;7:54007–14. http://dx.doi.org/10.1109/ACCESS.2019.2909969.
[39] Samuel OW, Asogbon GM, Sangaiah AK, Fang P, Li G. An integrated decision support system based on ANN and fuzzy_AHP for heart failure risk prediction. Expert Syst Appl 2017;68:163–72. http://dx.doi.org/10.1016/j.eswa.2016.10.020.
[40] Ali L, Bukhari SAC. An approach based on mutually informed neural networks to optimize the generalization capabilities of decision support systems developed for heart failure prediction. IRBM 2021;42(5):345–52. http://dx.doi.org/10.1016/j.irbm.2020.04.003.
[41] Verma L, Srivastava S, Negi PC. A hybrid data mining model to predict coronary artery disease cases using non-invasive clinical data. J Med Syst 2016;40(7):178. http://dx.doi.org/10.1007/s10916-016-0536-z.
[42] Alaa AM, Bolton T, Angelantonio EDi, Rudd JHF, van der Schaar M. Cardiovascular disease risk prediction using automated machine learning: A prospective study of 423;604UK biobank participants. PLoS One 2019;14(5):e0213653. http://dx.doi.org/10.1371/journal.pone.0213653, PMID: 31091238; PMCID: PMC6519796.
[43] Ward A, Sarraju A, Chung S, Li J, Harrington R, Heidenreich P, et al. Machine learning and atherosclerotic cardiovascular disease risk prediction in a multi-ethnic population. NPJ Digit Med 2020;3:125. http://dx.doi.org/10.1038/s41746-020-00331-1, PMID: 33043149; PMCID: PMC7511400.
[44] Jamthikar A, Gupta D, Khanna NN, Araki T, Saba L, Nicolaides A, Sharma A, et al. A special report on changing trends in preventive stroke/cardiovascular risk assessment via B-mode ultrasonography. Curr Atheroscler Rep 2019;21(7):25. http://dx.doi.org/10.1007/s11883-019-0788-4, PMID: 31041615.
[45] Suri JS, Bhagawati M, Paul S, Protogeron A, Sfikakis PP, Kitas GD, et al. Understanding the bias in machine learning systems for cardiovascular disease risk assessment: The first of its kind review. Comput Biol Med 2022;142:105204. http://dx.doi.org/10.1016/j.compbiomed.2021.105204, Epub 2022 Jan 4. PMID: 35033879.
[46] Vulli A, Srinivasu PN, Sashank MSK, Shafi J, Choi J, Ijaz MF. Fine-tuned DenseNet-169 for breast cancer metastasis prediction using FastAI and 1-cycle policy. Sensors 2022;22(8):2988. http://dx.doi.org/10.3390/s22082988.
[47] Chatzimichail T, Hatjimihail AT. A software tool for calculating the uncertainty of diagnostic accuracy measures. Diagnostics 2021;11(3):406. http://dx.doi.org/10.3390/diagnostics11030406.
[48] UCI: Heart failure prediction dataset. UCI Machine Learning Repository, https://archive.ics.uci.edu/ml/machine-learning-databases/heart-disease/. [Accessed 2023].
[49] Stone M. Cross-validatory choice and assessment of statistical predictions. J R Stat Soc Ser B (Methodol) 1974;36(2):111–33. http://dx.doi.org/10.1111/j.2517-6161.1974.tb00994.x.
[50] Hamed SH, Alhafidh OK, Younis YS, Yahia WB, Meften S, Ali AH. A modified wavelet threshold approach for reducing various noise with statistical results. In: 2022 international conference on artificial intelligence of things. ICAIoT, IEEE; 2022, p. 1–6. http://dx.doi.org/10.1109/ICAIoT57170.2022.10121838.
[51] Chollet F. Deep learning with python. Simon & Schuster; 2017, Available at: https://www.simonandschuster.com/books/Deep-Learning-with-Python/Francois-Chollet/9781617294433.
[52] Vulli A, Srinivasu PN, Sashank MS, Shafi J, Choi J, Ijaz MF. Fine-tuned DenseNet-169 for breast cancer metastasis prediction using fastai and 1-cycle policy. Sensors 2022;22(8):2988, https://ieeexplore.ieee.org/document/9301880.
[53] Zhang H, Cricri F, Tavakoli HR, Santamaria M, Lam YH, Hannuksela MM. Learn to overfit better: finding the important parameters for learned image compression. In: 2021 international conference on visual communications and image processing. VCIP, IEEE; 2021, p. 1–5. http://dx.doi.org/10.1109/VCIP53242.2021.9675360.

[54] Al-Darraji S, Honi DG, Fallucchi F, Abdulsada AI, Giuliano R, Abdulmalik HA. Employee attrition prediction using deep neural networks. Computers 2021;10(11):141. http://dx.doi.org/10.3390/computers10110141.
[55] Hinton GE, Osindero S, Teh YW. A fast learning algorithm for deep belief nets. Neural Comput 2006;18(7):1527–54. http://dx.doi.org/10.1162/neco.2006.18.7.1527, PMID: 16764513.
[56] LeCun Y, Bengio Y, G. Hinton. Deep learning. Nature 2015;521:436–44. http://dx.doi.org/10.1038/nature14539.
[57] Kohavi R, John GH. Wrappers for feature subset selection. Artif Intell 1997;97(1–2):273–324. http://dx.doi.org/10.1016/S0004-3702(97)00043-X.
[58] Ting KM, Witten IH. Issues in data mining with text: Text data mining vs. Traditional data mining. In: Data mining and knowledge discovery handbook. 2nd ed.. Boston, MA: Springer; 2011, p. 405–35.
[59] Gholamy A, Kreinovich V, Kosheleva O. Why 70/30 or 80/20 relation between training and testing sets: A pedagogical explanation. Int J Intell Technol Appl Stat 2018;11(2):105–11.