International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
IJISAE
ISSN:2147-6799 www.ijisae.org Original Research Paper

NARXnet Machine Learning Prediction Cardiovascular Disease Event ECG and PCG

Islam D. S. Aabdalla, Dr. D. Vasumathi*[2]

Submitted:11/03/2024 **Revised**: 26/04/2024 **Accepted**: 03/05/2024

Abstract: In the modern era, cardiovascular diseases (CVD) have become a major health concern. Trained neural network classifiers are widely used to predict abnormalities related to cardiovascular disease (CVD) by analyzing ECG signals and PCG sounds. It effectively conveys the process of deriving signals from the R interval of the ECG signal and processing and segmenting them simultaneously. Different estimation techniques are used after the chaotic, time, and Frequency domain characteristics are derived. This work focuses on using the Bayesian regularization (BR), Levenberg-Marquardt (LM), and Scaled Conjugate Gradient (SCG) optimization algorithms to train the NARX network. Nonetheless, utilizing a variety of models, including LR, SLR, DT, SVM, Ensemble, GPR, NN, and others, predicts many features extracted from labeled PCG sound and ECG signals. This study evaluates the trained NARX model's prediction performance concerning the three optimization algorithms utilized during the training phase. It compares various machine learning methods for estimating CVD and evaluates the estimation results based on performance criteria. The NARX-BR artificial neural network detects CVD with an R-squared value of 0.968 and MSE of 0.0738 and the highest accuracy, achieved at 98.2%, is observed for Decision Tree for predication cardiovascular.

Keywords*: PCG sound, machine learning, ECG signals, NARX, LM, SCG, BR.*

1. Introduction

In today's world, with a global population of 7.7 billion, heart diseases stand as one of the leading causes of mortality[1]. ECG and PCG are commonly used diagnostic tools in assessing cardiovascular disease. The ECG records the heart's electrical activity, providing insights into its rhythm, rate, and potential irregularities. By analyzing the ECG waveform, medical can detect various CVD conditions such as arrhythmias, myocardial infarction, atrial fibrillation, ventricular hypertrophy, and others. Widely used in clinical practice, emergency care, and continuous monitoring, the ECG is pivotal for assessing cardiac function and diagnosing abnormalities[2] PCG offers detailed heart sound data that assists in medical assessments, records heart sounds using a microphone or transducer, capturing mechanical events like S1 and S2 sounds, murmurs, clicks, and other anomalies. PCG is valuable for identifying heart valve issues like stenosis or regurgitation, congenital defects, and structural abnormalities. Healthcare providers analyze heart sound timing, intensity, and characteristics to evaluate valve and chamber function[3][4]. Machine learning techniques are increasingly used to predict cardiovascular disease (CVD) from ECG and PCG data. These methods analyze vast amounts of signal data to develop models that identify patterns linked to CVD. NARX, a neural network architecture, models dynamic systems by considering past inputs and outputs, making it superior for nonlinear time series prediction influenced by external factors. NARX networks are particularly effective in predicting cardiovascular abnormalities from ECG and PCG signals[5]. The NARX network comprises two distinct configurations: closed-loop and open-loop. During training, the open-loop setup is preferred, leveraging real historical values within the time series. Subsequently, the trained open-loop NARX network is transformed into a closed-loop configuration, offering benefits for multi-step predictions during testing. The primary aim of neural network training is to minimize a substantial cost function[6][7]. The NARX model, utilized for function determination, offers versatility in signal processing applications, with the simplest approach involving a feedforward neural network incorporating embedded memory. It effectively conveys the idea that the network's design relies on previous sequence elements, providing a limited view of the series through neighboring sequence elements, referred to as a time window. A state space depiction illustrates the recurrent nature of NARX neural networks, tailored to model time series data by considering past values of both input and output variables in predicting the current output. Essentially, NARX networks predict the next value in a sequence based on historical input and output data. The NARX neural network model has proven effective in cardiovascular health by predicting missing data points in impedance cardiography signals, thereby improving the identification of vital hemodynamic parameters such as stroke volume and cardiac output[8][9]. Furthermore, CNN architectures have been employed to diagnose heart conditions using electrocardiogram (ECG) images with remarkable accuracy,

PhD Scholar in Computer Science and Engineering at JNTUH University, Hyderabad, India[1].
Computer Science and Engineering at JNTUH University, Hyderabad, India[2].*

enabling diagnosis even with ECG images captured by smartphones, which is particularly valuable in settings lacking expert diagnosis resources[10]. Additionally, CNNs have been utilized to automatically diagnose heart diseases based on ECG signals, achieving high classification accuracy, especially when incorporating post-processing filters[11]. Hence, the amalgamation of machine learning techniques like NARX and CNNs with ECG and PCG data shows great potential in enhancing the diagnosis and monitoring of heart diseases.

1.1. Advantages of NARX Networks:

Memory of Past Predictions: The feedback connections enable NARX networks to retain past predictions, aiding in capturing long-term dependencies in time series data. Nonlinear Relationships: NARX networks excel at capturing nonlinear relationships between input and output variables, common in real-world datasets. Prediction Accuracy: Properly trained NARX networks can yield highly accurate predictions for time series data. This study aims to use machine learning to predict cardiovascular diseases (CVD) from Electrocardiogram (ECG) and Phonocardiogram (PCG) data. ECG and PCG record heart activity, aiding in CVD assessment. The focus is on using NARX neural networks to identify CVD patterns[12].

1.2. The Main Contribution Of This Study:

1- Extract features such as Scalogram Power, Spectrogram, and Persistence Spectrum for both datasets.

2- Determining "mu" (learning rate) and "gradient" involves calculating the parameter update step size, highlighting the importance of precise parameter tuning for reliable predictions.

3-Employ NARX with three learning algorithms using Bayesian-Regularization, Levenberg-Marquardt, and SCG to predict CVD using ECG and PCG datasets to enhance diagnostic accuracy.

4- Apply different machine learning models such as LR, SLR, DT, SVM, Ensemble, GPR, and NN.

5- Calculate RMSE, MSE, R-squared, and MAE for different algorithms.

6- Compare results with previous work.

This study is organized into sections, with Section 2 offering a summary of relevant literature, Section 3 outlines the methodology, Section 4 presents the results and discussion, and Section 5 provides the conclusion.

2. Literature Review

Cardiovascular diseases, a leading cause of death, prompt research on predictive algorithms using health datasets to identify associated risk factors. Modern advances in ML tools drive the development of techniques for cardiac disease diagnosis, with approaches including clustering and classification. Numerous machine learning models proposed by researchers aim to detect cardiovascular disease[13].Umit et al[2]. Aaddressed cardiovascular diseases (CVD), concentrating on cuffless blood pressure estimation through PPG and ECG signals for uninterrupted vascular access. The study preprocessed signals, segmented data, and eliminated noise using moving averages. Blood pressure measurement involved extracting features from chaotic, time, and frequency domains, utilizing Support Vector Regression (SVR), Nonlinear Autoregressive Neural Networks NARX-NN, Coarse Tree, and Linear Regression techniques. Specifically, LSTM-NN, and NARX-NN. Demonstrated accuracy in estimation without vascular occlusion, suggesting applications in wearable technology for continuous blood pressure monitoring. Furthermore, new artificial neural networks (ANNs) such as MLP, ENN, and RBF were developed for diagnosing valve-related physiological heart diseases.

Pathak, et al[14]. Introduced a novel method for predicting coronary artery disease (CAD) using PCG signals. The proposed approach analyzed the time-varying frequency characteristics of PCG signals, particularly focusing on both systolic and diastolic phases. Experiments were conducted on 960 PCG recordings from heart disease and normal subjects, demonstrating promising results. The method, which utilized the synchrosqueezing transform of the cardiac cycle, achieved improved accuracy in CAD detection compared to existing techniques. The study emphasized the potential of developing a non-invasive, cost-effective CAD prediction system using PCG signals, which could enhance healthcare accessibility, especially for marginalized populations.

Monish et al[15]. Developed the importance of real-time ECG monitoring for the early classification of life-threatening conditions. They presented a NARX-wavelet model that addressed motion artifacts in ECG signals by merging NARX neural networks with wavelet-based filtering. Using an Arduino Uno, an AD8232 ECG sensor module, and a laptop, they focused on motion artifact removal. The results favored the Scaled Conjugate Gradient algorithm, which enhanced the signal-to-noise ratio through wavelet denoising. The model provided efficient artifact removal, sampling rate independence, and cross-platform portability for ECG monitoring systems.

Esti et al[16] introduced NN-FCA, which ranked features and examined correlations to enhance CHD risk prediction. Tested on a Korean dataset, NN-FCA outperformed the Framingham risk score (FRS) with superior accuracy and a larger ROC curve. Crucial features such as BMI, cholesterol levels, and blood pressure were emphasized, with NN-FCA highlighting correlations such as BMI with blood pressure and cholesterol. This model showed significant potential in

CHD prediction, potentially aiding in personalized treatment and prevention.

Alromema[17] Focused on analysing heart sounds, recognized for their complex behaviour with murmurs, using nonlinear dynamic models. A method was presented for extracting features from these models to enhance the classification of phonocardiograms (PCGs). Heart sound auscultation, depicted graphically by PCGs, was crucial for the early prediction of CVD. The methodology involved signal pre-processing, cardiac cycle segmentation, and feature extraction from nonlinear dynamical modelling. The results highlighted the effectiveness of the RUSBoosted Tree ensemble classifier, indicating the potential for improved heart sound analysis using nonlinear features.

Benouar,[8]introduced a predictive model that employed a non-linear autoregressive neural network (NARX) to forecast missing points in Impedance Cardiography (ICG) complexes, notably the X point. The NARX model significantly improved the detection rates of ICG characteristic points (X, Y, O, Z), ranging from 75% to 88%, a notable increase from previous rates of 21% to 30%. This model enhanced the extraction of hemodynamic parameters crucial for assessing left ventricular pre-ejection time (LVET). It addressed the variability of ICG complex subtypes and showed promise for personalized X-point selection. Further investigations explored its impact on LVET calculation, requiring comparison with thermodilution measurements for validation.

Fayaz, et al[18]. Developed the GWLM-NARX model to improve CVD risk prediction, combining the Grey Wolf Levenberg algorithms and NN for enhanced accuracy in early disease detection and prevention. The methodology included developing and validating the model with clinical and demographic data, utilizing Python, Pandas, and Sklearn tools for analysis, and applying L2 regularization to combat overfitting. Results showed over 90% accuracy in disease prediction, leveraging Autoregressive (AR) and(NARX) models to identify key risk factors. Despite acknowledging limitations like dataset quality and potential overfitting, the model demonstrated effectiveness in risk identification and personalized treatment planning, paving the way for more accurate prediction systems in cardiovascular healthcare.

Khaled, Sara, et al[19]. Utilized NARX to classify PCG signals, which are crucial for diagnosing heart issues, comparing BR, LM, and SCG optimization algorithms. NARX with BR outperformed LM and SCG, improving results. PCG signals, important for heart health assessment, were analysed alongside ECG and PPG. NARX identified normal/abnormal signals, focusing on s1 and s2 heartbeats and key regions. The network's open/closed-loop structures minimized costs via BR, LM, and SCG—BR excelled for PCG. The paper reviewed related works and presented methodology, results, and conclusions. It emphasized BR's effectiveness, utilizing NARX's MLP architecture for nonlinear mapping. Efficient in time series modelling, NARX suggested empirical optimizations with SCG, LM, and BR—BR proved superior. PhysioNet 2016 data-informed optimal neuron/delay setups for classification.

Amin, et al[20]. A personalized real-time hybrid model was introduced for predicting the severity of patients' conditions during their Emergency Department (ED) stay. This model utilized a combination of Nonlinear Autoregressive Exogenous (NARX) and Ensemble Learning (EL), leveraging vital signs such as Pulse Rate (PR), Respiratory Rate (RR), Arterial Blood Oxygen Saturation, and Systolic Blood Pressure (SBP) automatically collected during treatment. It forecasted the severity of illness in the upcoming hour based on vital signs recorded in the preceding two hours. Two EL techniques, namely Random Forest (RF) and Adaptive Boosting (AdaBoost), were utilized. Comparative evaluations with alternative models, including Auto-Regressive Integrated Moving Average (ARIMA), NARX coupled with Linear Regression (LR), Support Vector Regression (SVR), and K-Nearest Neighbors Regression (KNN), demonstrated significantly improved accuracy with the proposed NARX-EL models. Particularly, NARX-RF excelled in predicting sudden fluctuations and unexpected adverse events in patients' vital signs, achieving an R^2 score of 0.978 and an NRMSE of 6.16%

Amos, et al[21]. explored the effectiveness of Neural Network architectures in screening residential rental applications within the Nigerian property market. Comparing Recursive Neural Network (RNN) and Feedforward Neural Network (FFNN), ten training algorithms were assessed using data from 53 property managers in Lagos. Performance metrics included sensitivity, specificity, precision, and various scores. Results indicated satisfactory performance for both FFNN and RNN, with Bayesian regularization (BR) outperforming other algorithms. Conversely, Gradient Descent, Resilient Backpropagation (RP), and Scaled Conjugate Gradient Backpropagation (SCG). Showed lower performance. The study suggested BR-trained FFNN and RNN as optimal for rental application screening. Limitations included scope and algorithm selection.

Young Lee, et al[22]. Validated a deep learning-based artificial intelligence algorithm for detecting myocardial infarction (MI) using 6-lead electrocardiography (ECG), addressing conventional interpretation limitations and emphasizing rapid MI diagnosis. Trained on 400,000 ECGs, the algorithm achieved promising results in validation. With a variational autoencoder (VAE), the DLA reconstructed precordial 6-lead ECGs from limb 6-lead ECGs, consistently detecting MI across diverse ECG data subsets. Cardiologists reviewed patient records to ensure dataset accuracy, focusing on type 1 and 2 MIs. The DLA architecture, employing multiple hidden layers, achieved MI probability indication, with TensorFlow and Python implementation adapting the algorithm for real-time MI detection in wearable devices, promising significant advancements in healthcare technology.

Landry, et al[23]. Aimed to develop a cuffless method for accurately estimating blood pressure (BP) waveform and extracting key BP features like systolic BP, diastolic BP, and mean BP. Access to the entire waveform offered advantages over previous cuffless BP estimation methods, enhancing accuracy and providing additional cardiovascular health indicators. NARX was employed via artificial neural networks to predict BP waveforms using ECG or PPG signals as inputs. The model's effectiveness was compared to a PAT model using data from 15 subjects in the MIMIC II database. Two training methods were tested: predictive training on the initial eight minutes per subject, testing on the rest (up to 5.2 hours), and interval training using the first and last eight-minute segments, with intermittent testing. Initially, both methods showed similar results, but interval training proved more accurate over longer durations. Treating BP as a dynamic system improved precision in estimating SBP, DBP, and MAP compared to the PAT model. Additionally, the NARX model provided deeper insights into patient health by furnishing the BP waveform.

Jyothi, et al[24]. Introduced the Gaussian Kaiming Variance-based Deep Learning Neural Network (GKVDLNN) classifier for detecting heart disease (HD). Explored heart disease detection utilizing ECG and PCG signals, ensuring accuracy with extensive datasets. ECG and PCG signals were obtained from publicly available datasets, and pre-processing was performed using Improved Empirical Mode Decomposition (IEMD). Signal features were extracted from decomposed bands, selected, concatenated, and classified by GKVDLNN. Experimental findings showcased 96.103% accuracy with reduced costs, emphasizing the significance of early heart disease detection and proposing an innovative approach integrating ECG and PCG signals with advanced machine learning techniques.

Turker, et al[25]. Introduced a novel approach to generating graph-based features by employing the Petersen graph pattern (PGP) and a new decomposition technique known as tent pooling (TEP) decomposition. Through the integration of TEP and PGP, they developed a multilevel feature generation network. Feature selection was carried out using iterative neighbourhood component analysis (INCA). These chosen features were then inputted into decision tree (DT), linear discriminant (LD), bagged tree (BT), and support vector machine (SVM) classifiers to classify them into five categories automatically. The method achieved a remarkable 100.0% classification accuracy using KNN with ten-fold cross-validation. The DT, LD, BT, and SVM classifiers attained accuracies of 95.10%, 98.30%, 98.60%, and 99.90%, respectively. This high classification accuracy underscores the potential of utilizing PCG signals for heart sound classification with the proposed PGP and TEP-based model.

Parasto, et al[26] proposed method's efficacy was assessed using the Physionet Challenge 2016 database, employing a 10-fold cross-validation approach. To address dataset imbalance, the Synthetic Minority Over-sampling Technique (SMOTE) was applied to create balanced datasets. This methodology outperformed existing approaches in the literature, demonstrating higher accuracy, sensitivity, and specificity metrics. Notably, it achieved an accuracy of 98.03%, sensitivity of 97.64%, and specificity of 98.43% in distinguishing normal from abnormal heart sounds within the Physionet database, surpassing results obtained by previously established methods evaluated in the Physionet 2016 challenge database. Researchers have developed predictive algorithms and systems to assist medical practitioners and cardiologists in analysing data, using diverse models for accuracy, and comparing them with existing approaches. In our study, we analysed the ECG and PCG dataset and, NARX with different algorithms for improved results.

3. **Methods**

The proposed method, as explained in Fig. 1, outlines a CVD prediction framework Initial preprocessing involves recording ECG and PCG, followed by noise reduction and feature extraction for NARX neural network training with three algorithms firstly Bayesian Regularization (BR) algorithm is utilized to regulate the neural network (NN) using Bayesian methods and determine the optimal parameters. The training procedure will halt automatically if there is no enhancement in generalization, as indicated by the rise in the mean square error of validation trials. The NARX model undergoes retraining with each change in initial configurations and adjustments in data sets, resulting in diverse results. Scaled Conjugate Gradient (SCG) is preferred for its simplicity, despite its slower execution, being more suitable for uncomplicated objective functions. Levenberg-Marquardt (LM) is utilized to tackle the challenge of non-linear least square curve fitting, combining Gauss-Newton and gradient descent methods for minimization[27]. LM is known for its speed and efficient memory usage. Forevermore extracting a feature vector from PCG and ECG signals and predicting them using the NARX model for CVD. This NARX utilizes recurrent connections and activation functions, iteratively adjusting weights and biases for accurate predictions. The proposed models utilize LM, BR, and SCG training, sharing a common activation function, performance metric

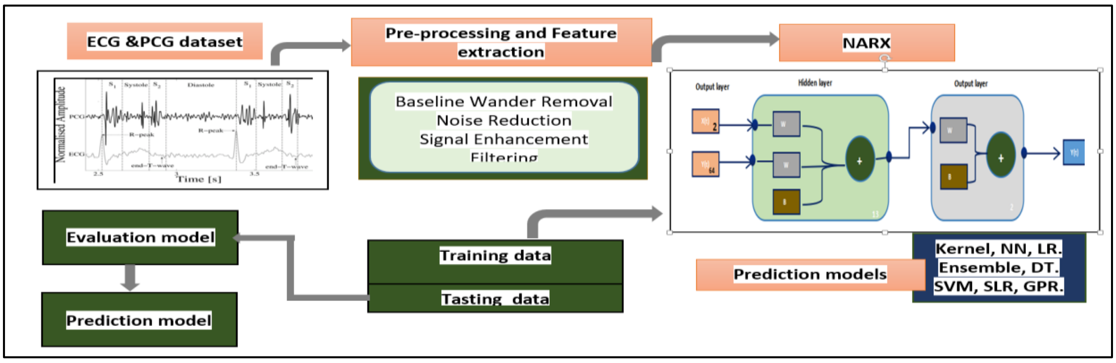


Fig. 1. Flowchart of the proposed NARX network to predict CVD.

3.1. Dataset description

The datasets analyzed in this paper from the PhysioNet 2016 challenge, are publicly accessible on the website [30]. A total of 6,316 feature vectors, each of length 27, were extracted from the recordings and signals. These comprised 3,158 vectors from healthy hearts and an equal number from unhealthy hearts. Its 80% of the feature vectors are allocated for training, with the remaining set aside for testing. The training database comprises 1580 samples, each with 64 predictor variables and 2 response variables. Similarly, the testing database consists of 395 samples with the same predictor and response variable structure. These databases are employed to train and assess models designed to analyze PCG and ECG signals for diverse objectives, including disease diagnosis and prognosis prediction.

3.2. Dataset Preprocessing

The preprocessing of ECG and PCG signals is crucial for CVD prediction, involving preprocessing and division into the initial heart sound S1 and the second heart sound S2. We conducted this data processing using the methodology outlined in a prior publication [8], which consists of four main stages: preprocessing, peak detection, rejection of extra peaks, and identification of S1 and S2. Consequently, we aim to validate whether incorporating ECG data alongside PCG data enhances diagnostic performance compared to relying solely on PCG signals**.**

3.2.1. QRS complex detection

ECG preprocessing, including PCG and R peak detection, employs three filters, signal squaring, and adaptive thresholding[28]. Enhancements included: Rejecting the first R peak if it's extreme, verifying QRS complexes' prominence, and rejecting peaks with intervals under 0.5s; these instructions were also integrated into the PCG classification to enhance S1 and S2 identification.

Table1. PCG dataset preprocessing

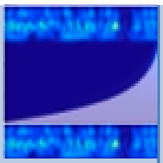
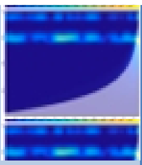
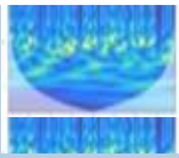
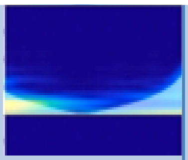
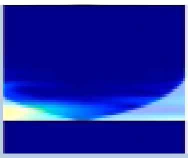
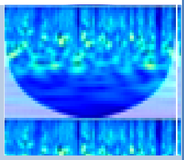
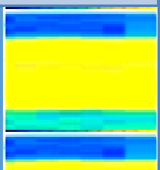
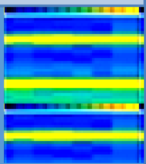
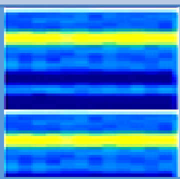
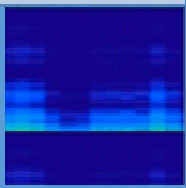
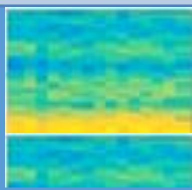
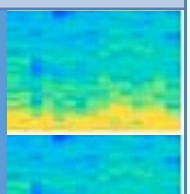
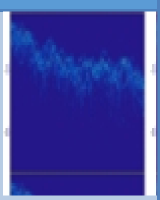
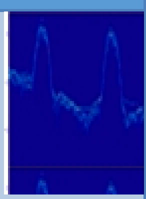
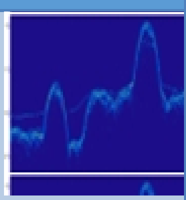
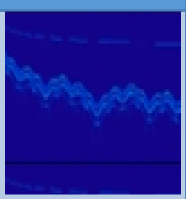
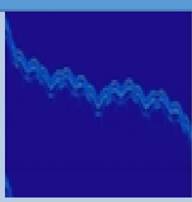
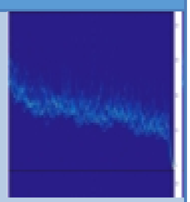
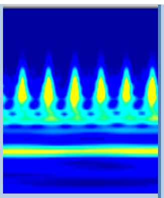
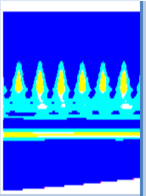
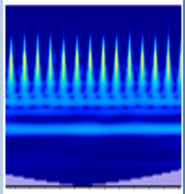
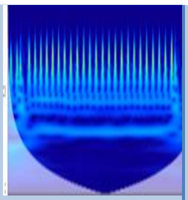
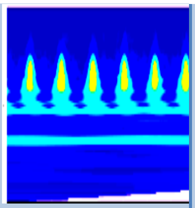
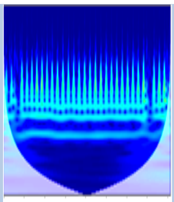
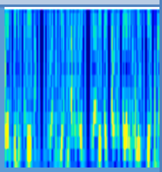
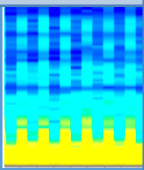
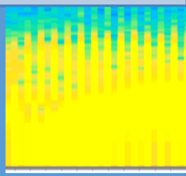
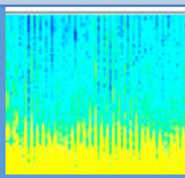
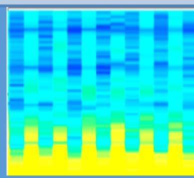
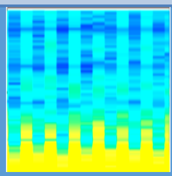
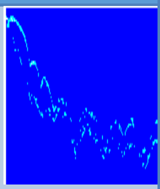
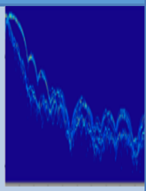
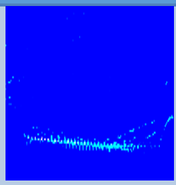
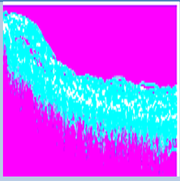
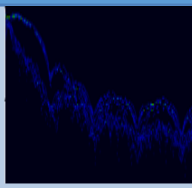
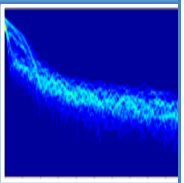
	BandPass	BandStop	HighPass	Denoise	Smooth	Envelope
Scalogram						
Power Spectrogram						
Persistence Spectrum						

Table 2. ECG dataset preprocessing

	BandPass	BandStop	HighPass	Denoise	Smooth	Envelope
Scalogram						
Power Spectrogram						
Persistence Spectrum						

These filters and methods seem to serve different purposes in signal processing, from noise reduction to frequency band manipulation and feature extraction. Each would be chosen based on the specific requirements of the application and the characteristics of the signal being processed.

Table 3. Parameters of pre-process Filtering

Filter Name	Filter Specification	
Bandpass	Frequency unit	Sample/xπ radians
	Lower Frequency	0.25
	Upper Frequency	0.75
	Lower Band Steepness	0.85
	Lower Band Steepness	0.85
	Stopband Attenuation	60
Band Stop	Frequency unit	Sample/xπ radians
	Lower Frequency	0.25
	Upper Frequency	0.75
	Lower Band Steepness	0.85
	Lower Band Steepness	0.85
HighPass	Frequency unit	Sample/xπ radians
	Passband	0.5
	Steepness	0.85
	Stopband Attenuation	60
Denoise	Wavelet	Sym No 4
	Method	Bayse
	Level	8
	Rule	Median
	Noise estimate	Level independent
Smooth	Smooth methods	Moving mean
	Window type	Smoothing factor 0.5
Envelope	Envelope type	Lower Envelope
	Methods	hilbert

3.2.2. Bandpass Filter:

A bandpass filter is specialized to allow only a precise range of frequencies to pass through, effectively filtering out any signals outside of this designated range. Its specifications typically include defining the lower and upper limits of the frequency range it permits, indicating the steepness of the transition between passbands and stopbands, and specifying the level of attenuation in the stopband to effectively block unwanted frequencies.

$$y(t) = f^{-1}\{f\{x(t)\}.H_{BP\,(F)}\} \quad (1)$$

Where x(t) is the input signal, y(t) is the filtered output signal, denotes the Fourier transform, and $H_{BP\,(F)}$ is the frequency response of the bandpass filter.

3.2.3. Band Stop Filter: is commonly designed to inhibit a particular range of frequencies while permitting all others to pass unaffected. Much like its counterpart, the bandpass filter, it is characterized by specifications detailing the targeted frequencies, the degree of frequency alteration, and the extent of attenuation in the stopband.

$$y(t) = f^{-1}\{f\{x(t)\}.H_{Bs\,(F)}\} \quad (2)$$

Similar to the bandpass filter, but $H_{Bs\,(F)}$ represents the frequency response of the bandstop filter.

3.2.4. HighPass Filter: abbreviated as HPF, is an electronic component engineered to enable frequencies exceeding a predetermined cutoff threshold to transit, while concurrently diminishing frequencies below this threshold. This filter delineates its operation through specifications detailing the designated passband frequency, the rate of signal attenuation within this passband, and the degree of attenuation in the stopband.

$$y(t) = f^{-1}\{f\{x(t)\}.H_{HP\ (F)}\} \quad (3)$$

Similar to the other filters, but $H_{HP\ (F)}$ represents the frequency response of the high-pass filter.

3.2.5. Signal Denoising: Signal Denoising is a technique aimed at eradicating noise while preserving essential information within a signal. This process is executed through a denoising function, denoted as y(t) = Denoising Function(x(t), parameters), where the specific denoising algorithm and its corresponding parameters dictate the transformation applied to the signal.

3.2.6. Envelope Extraction: This technique focuses on delineating amplitude fluctuations across time, delineating the specific envelope type, typically the lower envelope, and employing the Hilbert transform for extraction.

Equation: The mathematical expression governing this process is represented as

y(t) = Envelope Function(x(t)), where the envelope function encompasses methodologies like the Hilbert transform or peak detection, facilitating the extraction of the signal's envelope from x(t).

3.3. Features extraction

PCA reduces dimensionality by transforming high-dimensional data into a lower-dimensional space, retaining maximum original data variance via orthogonal principal components. It involves computing eigenvectors and eigenvalues from the covariance matrix, sorting them by eigenvalues, and projecting the data onto these vectors for transformation[6]. Feature selection for classification can be achieved through methods assessing frequency distribution disparities. Additionally, frequency suppression within a defined range can be performed while allowing others to pass[23].

3.4. NARX

The NARX (Nonlinear AutoRegressive with eXogenous inputs) neural network is adept at modeling and predicting time series data. Unlike traditional autoregressive models, NARX networks incorporate exogenous inputs, enhancing their ability to capture complex data relationships. They feature input, hidden, and output layers, and are trained using algorithms like backpropagation[1]. NARX networks excel in tasks such as time series prediction, system identification, and control. The fundamental expression representing the NARX model is:

$$y(t) = f(y(t-1), \ldots, y(t-dy), x(t-1), \ldots, x(t-dx)) \quad (4)$$

The equation comprises the mapping function f(...) utilized by the neural network. In this context, y(t) denotes the NARX output at time t, representing the predicted value of y for that moment. The terms y(t -1), ..., y(t-dy) represent previous outputs of the NARX, while x(t), ..., x(t -dx) denote the inputs. Here, dx signifies the number of input delays, and dy represents the number of output delays. The NARX model integrates a two-layer feed-forward neural network to approximate the function[30]. Contrarily, Recurrent Neural Networks (RNNs) are recognized for their intricacy, often necessitating prolonged training and learning periods.

The NARX model employs SCG, LM, and BR optimization algorithms. SCG enhances convergence speed, LM is effective in networks with numerous weights, and BR mitigates overfitting by balancing error and complexity. The study evaluates their impact on classifying PCG signals, exploring optimal activation and loss function combinations

Three specific algorithms used are Scaled Conjugate Gradient (SCG), and Scaled Conjugate Gradient is an optimization algorithm commonly used for training neural networks.

The equation for the parameter vector θ at iteration k is as follows:

$$\theta_{k+1} = \theta_k + \alpha_k p_k \quad (5)$$

Where $\alpha_k$is the step size determined by line search, is the search direction, and θ_k is the parameter vector at iteration k.

Levenberg-Marquardt (LM) is a widely employed optimization technique specifically tailored for nonlinear least squares curve fitting tasks. The formula governing the update of the parameter vector θ at iteration k is as follows:

$$\theta_{k+1} = \theta_k - (J_k^t + \lambda_k I)^{-1} J_k^t R_k \quad (6)$$

Where the J_k is the Jacobian matrix of the residuals R_k concerning θ, $\lambda_k$is the damping parameter,

I am the identity matrix.

and Bayesian regularization (BR). Bayesian regularization is a technique used for regularization in machine learning models. The regularization term is added to the cost function, which is minimized during training. The cost function for Bayesian regularization can vary depending on the specific model being trained, but it typically involves a regularization term that penalizes large parameter values[29]. The optimization problem is then formulated as:

$$\min\theta = L(\theta) + \lambda R(\theta) \quad (7)$$

Where L(θ) is the loss function, R(θ) is the regularization term, and λ is the regularization parameter, controlling the trade-off between fitting the data and penalizing large parameter values.

3.5. Metrics for Time Series Modelling

3.5.1. The Mean Square Error (MSE): This measures the average squared deviation between actual and predicted values by summing their squared differences and dividing by the total number of observations, offering a comprehensive evaluation of alignment with true values.

$$MSE = 1/\mathrm{n}\sum_{\mathrm{i=1}}^{\mathrm{n}}(\mathrm{Yi} - \hat{\mathrm{Y}}\mathrm{i})^2 \quad (8)$$

3.5.2. Mean absolute error(MAE)

measure the average absolute difference between predicted and observed values. By disregarding outliers, it avoids harsh penalties for large errors, with lower values signaling superior model performance**[2]**.

$$\mathrm{MAE} = \sum_{\mathrm{i=1}}^{\mathrm{n}} |\mathrm{Yi} - \hat{\mathrm{Y}}\mathrm{i}| \quad (9)$$

3.5.3. Root mean squared error(RMSE)

It is clear from the name RMSE that it is a square root of MSE.

$$\mathrm{RMSE}=\sqrt{MSE} \quad (10)$$

Coefficient of determination

3.5.4. The Coefficient of Determination (R2)

The statistical criterion *R*2 is calculated as follows:

$$R^2 = 1 - \frac{\sum_{\mathrm{i=1}}^{\mathrm{n}}(\mathrm{Yi}-\hat{\mathrm{Y}}\mathrm{i})^2}{\sum_{\mathrm{i=1}}^{\mathrm{n}}(\mathrm{Yi}-\bar{\mathrm{Y}}\mathrm{i})^2} \quad (11)$$

$$\mathrm{Accuracy} = \frac{\mathrm{TP+TN}}{\mathrm{TP+TN+FP+FN}} \quad (12)$$

4. Result and Discussion

Accuracy is crucial for assessing NARXnet and machine learning algorithms, primarily evaluated through Root Mean Squared Error (RMSE) and R-squared. A subsequent experiment utilizes Mean Absolute Error (MAE) to highlight differences between measured and estimated blood glucose levels.

From Table 4NARXnet training on ECG and PCG datasets for cardiovascular disease prediction employed LM, SCG, and BR algorithms. Results varied in performance (ranging from 0.88 to 1.00) and MSE (from 3.47e-08 to 0.0669), with differing epoch counts (8 to 10) and additional parameters for LM and SCG.

Table 4. NARX Neural Network Prediction CVD Performance

Dataset	Algorithms	No. epoch	Performance	Gradient	Mu
PCG	Levenberg-Marquardt (LM)	8	0.85	0.04529 1e-07	0.0001 1e+10
	Scaled Conjugate Gradient (SCG)	32	0.95	0.0167 1e-06	-
	Bayesian Regularization (BR)	35	0.77	3.47e-08	50 1e+10
ECG	Levenberg-Marquardt (LM)	10	0.89	0.0669 1e-07	0.0001 1e+10
	Scaled Conjugate Gradient (SCG)	38	0.97	0.0137 1e-06	-
	Bayesian Regularization (BR)	10	0.91	8.99e-07	0.005 1e+10

From table 5 The provided table displays numeric evaluation metrics for various machine learning models. Both Linear Regression (LR) and Simple Linear Regression (SLR) show identical results with an RMSE and MSE of approximately 0.0906 and 0.00821 respectively, and an MAE of about 0.9672. Decision Tree stands out with the lowest RMSE of around 0.0665, but a slightly higher MAE of 0.9823. Support Vector Machine (SVM) demonstrates moderate performance with an RMSE of 0.0785, MSE of 0.00617, and MAE of 0.9754. The

Ensemble model combines models to achieve an RMSE of 0.0697, MSE of 0.00486, and MAE of 0.9806. Gaussian Process Regression (GPR) and Neural Network (NN) exhibit comparable performance with an RMSE of around 0.081 and 0.0717 respectively, along with MSE and MAE values. Lastly, the Kernel model showcases an RMSE of 0.0743, MSE of 0.00552, and MAE of 0.978. In summary, the decision tree and ensemble models tend to offer the lowest RMSE, implying better predictive accuracy, while the choice of model depends on factors such as interpretability and computational efficiency.

Table5. Overall Performance Analysis of Proposed Machine Learning Algorithms for Predication CVD.

	Model	RMSE	MSE	R^2	MAE
ECG dataset	LR	0.090601	0.0082086	0.96721	0.04834
	SLR	0.090601	0.0082086	0.96721	0.04834
	Tree	0.066535	0.004427	0.98231	0.0074279
	SVM	0.078541	0.0061687	0.97536	0.041193
	Ensemble	0.069685	0.0048559	0.9806	0.0076959
	GPR	0.081047	0.0065687	0.97376	0.0098408
	NN	0.071717	0.0051434	0.97945	0.012024
	Kernel	0.07427	0.005516	0.97796	0.02538
PCG dataset	LR	0.22747	0.051741	0.79334	0.15228
	SLR	0.22906	0.052469	0.79043	0.15354
	Tree	0.25551	0.065285	0.73924	0.11509
	SVM	0.22022	0.048496	0.8063	0.13457
	Ensemble	0.2342	0.054852	0.78091	0.10865
	GPR	0.19125	0.036576	0.85391	0.10591
	NN	0.262	0.068646	0.72581	0.084435
	Kernel	0.21872	0.047839	0.80892	0.12002

Fethermore PCG result The evaluation metrics for various machine learning models shows that Linear Regression (LR) and Simple Linear Regression (SLR) yield similar outcomes, Decision Tree (Tree) boasts the lowest RMSE at approximately 0.0665, while the Ensemble model achieves an RMSE of about 0.0697. Gaussian Process Regression (GPR) and Support Vector Machine (SVM) perform moderately with an RMSE of around 0.081 and 0.0785, respectively. Neural Network (NN) and Kernel model also exhibit competitive performance with an RMSE of approximately 0.0717 and 0.0743. Decision tree and ensemble models shine with their lower RMSE, but model selection should also consider interpretability, efficiency, and specific problem requirements.

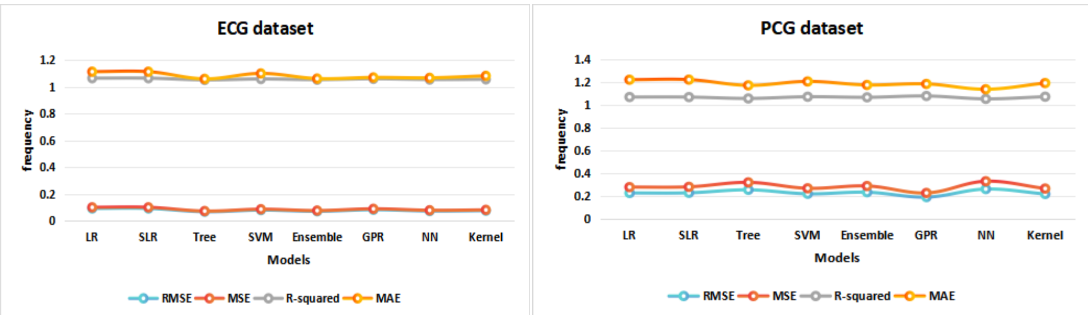


Fig. CVD Prediction Model Performance Comparison Across ECG and PCG Datasets.

From table 6 the PCG dataset, the SCG model performed the best with an accuracy of 94.67%. For the ECG dataset, the BR model had the highest accuracy of 96.80%. In terms of Mean Squared Error (MSE), SCG achieved the lowest MSE of 0.0288 on the PCG dataset, and similarly, SCG also achieved the lowest MSE of 0.0436 on the ECG dataset.

Table 6. Performance Scaled Conjugate Gradient, Levenberg-Marquardt, and Bayesian regularization.

Dataset	Performance	Levenberg-Marquardt (LM)	Bayesian regularization (BR)	Scaled Conjugate Gradient (SCG)
PCG	MSE	0.0723	0.3305	0.0288
	R^2	0.8438	0.7688	0.9467
ECG	MSE	0.0534	0.0738	0.0436
	R^2	0.8869	0.968	0.9052

In the comparison of algorithms on PCG and ECG datasets, Scaled Conjugate Gradient (SCG) consistently outperforms others with the lowest MSE values and highest R-squared R^2values. Levenberg-Marquardt (LM) shows competitive performance with moderate MSE and R^2 scores. Bayesian Regularization (BR) consistently performs the worst with high MSE and low R^2 values.

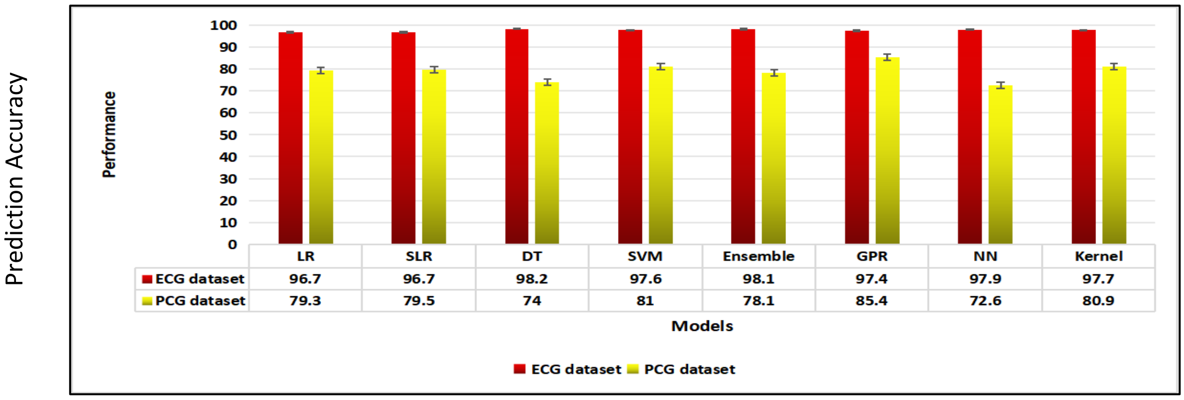


Fig3. Performance CVD Prediction Model.

Accuracy of the proposed algorithm and competition results of the competition from figure3. In the ECG dataset, the DT model performs the best with an accuracy of 98.2%, while in the PCG dataset, the GPR model achieves the highest accuracy of 85.4%.

Table7. Comparison of Model Performance in CVD Time Series Prediction

References	Models	R^2 Prefprmances	MSE
Bhattacharjee(2020)	NARX	0.948	-
Shiva et al (2022)	Lasso Regression	0.227	0.418
Diovu , et al(2022)	Liner Regression	0.36	0.141
Umit,et al(2020)	NARX,KNN,RF	0.78	6.16
Sheikh (2023)	GWLM-NARX	0.86	0.137
Proposed	NARX-BR	0.968	0.0738

Table 7 presents various models along with their Mean Squared Error (MSE) values as reported in different studies. The models include NARX, Lasso regression, Linear regression, NARX-NN, NARX combined with other algorithms like KNN, SVR, and RF, GWLM-NARX, and a proposed model named NARX-BR. The MSE values range from 0.2278 to 2.09, with NARX-BR having the highest MSE at 0.968 and the lowest being for the Lasso regression model at 0.2278.

Table 8: Summary of Prediction Accuracies Reported by Different Studies

References	Prediction Accuracy (%)
Senturk , et al[4]	88
Khaled, Sara, et al[19]	93
Talha, et al[8]	88
Brery,et al[32]	89.10
Sheikh Amir, et al.[18]	90
Cinzia, et al.[33]	70.77
Mudsir, et al.[34]	80.00
Abduh, et al[3]	90
Sara, et al.[35]	91
Proposed	**98.2**

4. Conclusion

Machine learning, notably the NARX model, exhibits considerable potential in cardiovascular disease (CVD) prediction through the ECG and PCG dataset This investigation extensively delved into the training of the NARX model, employing a range of optimization algorithms including Levenberg-Marquardt (LM), Scaled Conjugate Gradient (SCG), and Bayesian Regularization (BR), SCG stood out as the most effective among them. Comparative assessment against alternative models underscored the superior accuracy of NARX. Particularly noteworthy was the performance of the NARX-BR variant, which demonstrated remarkable accuracy metrics, boasting a MSE of 0.0738 and an R-squared value of 0.968, while the Decision Tree model achieved an outstanding accuracy rate of 98.2%. future work should concentrate on augmenting the NARX model with advanced feature engineering techniques and leveraging larger datasets to further enhance predictive capabilities. Additionally, exploring ensemble learning methodologies and validating the models in real-world clinical settings are imperative steps toward ensuring their practical applicability.

References

[1] A. Yadav, A. Singh, M. K. Dutta, and C. M. Travieso, "Machine learning-based classification of cardiac diseases from PCG recorded heart sounds," *Neural Computing and Applications,* vol. 32, no. 24, pp. 17843-17856, 2020.

[2] U. Senturk, K. Polat, and I. Yucedag, "A non-invasive continuous cuffless blood pressure estimation using dynamic recurrent neural networks," *Applied Acoustics,* vol. 170, p. 107534, 2020.

[3] H. Dong *et al.*, "Non-destructive detection of CAD stenosis severity using ECG-PCG coupling analysis," *Biomedical Signal Processing and Control,* vol. 86, p. 105328, 2023.

[4] M. M. Suhail and T. A. Razak, "Cardiac disease classification from ECG signals using hybrid recurrent neural network method," *Advances in Engineering Software,* vol. 174, p. 103298, 2022.

[5] M. Alkhodari and L. Fraiwan, "Convolutional and recurrent neural networks for the detection of valvular heart diseases in phonocardiogram recordings," *Computer Methods and Programs in Biomedicine,* vol. 200, p. 105940, 2021.

[6] F. Chakir, A. Jilbab, C. Nacir, and A. Hammouch, "Recognition of cardiac abnormalities from synchronized ECG and PCG signals," *Physical and Engineering Sciences in Medicine,* vol. 43, pp. 673-677, 2020.

[7] T. Kurian, "Deep Convolution Neural Network-Based Classification and Diagnosis of Heart Disease using ElectroCardioGram (ECG) Images," in *2023 IEEE 8th International Conference for Convergence in Technology (I2CT)*, 2023: IEEE, pp. 1-6.

[8] S. Benouar, M. Kedir-Talha, and F. Seoane, "Time-series NARX feedback neural network for forecasting impedance cardiography ICG missing points: a predictive model," *Frontiers in Physiology,* vol. 14, p. 1181745, 2023.

[9] V. M. Shervegar, "Heart sound classification using wavelet scattering transform and support vector machine," *Intelligent Data Analysis,* no. Preprint, pp. 1-17, 2023.

[10] T. Kurian, "Deep Convolution Neural Network-Based Classification and Diagnosis of Heart Disease using ElectroCardioGram (ECG) Images," 2023: IEEE, pp. 1-6.

[11] R. Avanzato and F. Beritelli, "Heart disease recognition based on extended ECG sequence database and deep learning techniques," in *2022 IEEE International Conference on Internet of Things and Intelligence Systems (IoTaIS)*, 2022: IEEE, pp. 117-121.

[12] A. Jain, R. Dubey, and V. V. Thakare, "A Novel Method for Diagnosis of Cardiac Disease Using ECG on Proposed CNN," in *Proceedings of International Conference on Communication and Computational Technologies: ICCCT 2022*, 2022: Springer, pp. 47-54.

[13] I. D. Suliman and D. Vasumathi, "Prediction of Heart Disease Using Machine Learning Algorithms," in *2023 14th International Conference on Computing*

Communication and Networking Technologies (ICCCNT), 2023: IEEE, pp. 1-12.
[14] A. Pathak, P. Samanta, K. Mandana, and G. Saha, "Detection of coronary artery atherosclerotic disease using novel features from synchrosqueezing transform of phonocardiogram," *Biomedical Signal Processing and Control,* vol. 62, p. 102055, 2020.
[15] [U. Bhattacharjee and M. Chakraborty, "NARX-wavelet based active model for removing motion artifacts from ECG," in *2020 International Conference on Computer, Electrical & Communication Engineering (ICCECE)*, 2020: IEEE, pp. 1-6.
[16] E. Suryani, S. Setyawan, and B. P. Putra, "The Cost-Based Feature Selection Model for Coronary Heart Disease Diagnosis System Using Deep Neural Network," *IEEE Access,* vol. 10, pp. 29687-29697, 2022.
[17] W. Alromema, E. Alduweib, and Z. Abduh, "Heart Sound Classification using the Nonlinear Dynamic Feature Approach along with Conventional Classifiers," *Engineering, Technology & Applied Science Research,* vol. 13, no. 3, pp. 10808-10813, 2023.
[18] S. A. Fayaz, M. Zaman, S. Kaul, and W. J. Bakshi, "Optimizing Cardiovascular Disease Prediction: A Synergistic Approach of Grey Wolf Levenberg Model and Neural Networks," *Journal of Information Systems Engineering & Business Intelligence,* vol. 9, no. 2, 2023.
[19] S. Khaled, M. Fakhry, H. Esmail, A. Ezzat, and E. Hamad, "Analysis of training optimization algorithms in the NARX neural network for classification of heart sound signals," *International Journal of Scientific and Engineering Research,* vol. 13, no. 2, pp. 382-390, 2022.
[20] A. Naemi, M. Mansourvar, T. Schmidt, and U. K. Wiil, "Prediction of patients severity at emergency department using NARX and ensemble learning," in *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2020: IEEE, pp. 2793-2799.
[21] J. A. Oguntokun and A. O. Adewusi, "Residential Rental Applications Screening: A Comparative Performance of Feedforward And Recursive Neural Networks Architectures," *Moroccan Journal of Quantitative and Qualitative Research,* vol. 5, no. 3, 2023.
[22] Y. Cho *et al.*, "Artificial intelligence algorithm for detecting myocardial infarction using six-lead electrocardiography," *Scientific Reports,* vol. 10, no. 1, p. 20495, 2020.
[23] C. Landry, S. D. Peterson, and A. Arami, "Nonlinear dynamic modeling of blood pressure waveform: Towards an accurate cuffless monitoring system," *IEEE Sensors Journal,* vol. 20, no. 10, pp. 5368-5378, 2020.
[24] P. Jyothi and G. Pradeepini, "Heart disease detection system based on ECG and PCG signals with the aid of GKVDLNN classifier," *Multimedia Tools and Applications,* pp. 1-26, 2023.
[25] T. Tuncer, S. Dogan, R.-S. Tan, and U. R. Acharya, "Application of Petersen graph pattern technique for automated detection of heart valve diseases with PCG signals," *Information Sciences,* vol. 565, pp. 91-104, 2021.
[26] P. S. Nia and H. D. Hesar, "Abnormal Heart Sound Detection using Time-Frequency Analysis and Machine Learning Techniques," *Biomedical Signal Processing and Control,* vol. 90, p. 105899, 2024.
[27] A. F. Gündüz and A. KARCİ, "Heart sound classification for murmur abnormality detection using an ensemble approach based on traditional classifiers and feature sets," *Computer Science,* vol. 5, no. 1, pp. 1-13, 2022.
[28] F. Tueche, Y. Mohamadou, A. Djeukam, L. C. N. Kouekeu, R. Seujip, and M. Tonka, "Embedded algorithm for QRS detection based on signal shape," *IEEE Transactions on Instrumentation and Measurement,* vol. 70, pp. 1-12, 2021.
[29] G. M. Rao, D. Ramesh, V. Sharma, A. Sinha, M. M. Hassan, and A. H. Gandomi, "AttGRU-HMSI: enhancing heart disease diagnosis using hybrid deep learning approach," *Scientific Reports,* vol. 14, no. 1, p. 7833, 2024.
[30] S. S. Reddy, N. Sethi, and R. Rajender, "Risk Assessment of myocardial infarction for diabetics through multi-aspects computing," *EAI Endorsed Transactions on Pervasive Health and Technology,* vol. 6, no. 24, pp. e3-e3, 2020.
[31] R. Diovu, "Performance Evaluation of Regression-Based Machine Learning Algorithms for Myocardial Infarction Prediction," *Performance Evaluation,* vol. 12, no. 43, pp. 5106-5111, 2022.
[32] M. Fakhry and A. F. Brery, "Comparison of window shapes and lengths in short-time feature extraction for classification of heart sound signals," *International Journal of Electrical and Computer Engineering (IJECE),* vol. 12, no. 6, pp. 6090-6102, 2022.
[33] C. Perrino *et al.*, "Improving translational research in sex-specific effects of comorbidities and risk factors in ischaemic heart disease and cardioprotection: position paper and recommendations of the ESC Working Group on Cellular Biology of the Heart," *Cardiovascular Research,* vol. 117, no. 2, pp. 367-385, 2021.
[34] M. Ashraf, Y. K. Salal, S. Abdullaev, M. Zaman, and M. A. Bhut, "Introduction of feature selection and leading-edge technologies viz. tensorflow, pytorch, and keras: An empirical study to improve prediction

accuracy of cardiovascular disease," in *International Conference on Innovative Computing and Communications: Proceedings of ICICC 2021, Volume 3*, 2022: Springer, pp. 19-31.

[35] S. Ghorashi *et al.*, "Leveraging regression analysis to predict overlapping symptoms of cardiovascular diseases," *IEEE Access*, 2023.