



# Automated health detection of congestive heart failure subject using rank multiresolution wavelet packet attributes and 1-norm linear programming ELM

Demissie J. Gelmecha<sup>1</sup> · Ram S. Singh<sup>1</sup> · Devendra K. Sinha<sup>2</sup> · Dereje Tekilu<sup>1</sup>

Received: 29 October 2020 / Revised: 22 August 2021 / Accepted: 20 September 2021 /  
Published online: 8 January 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

## Abstract

As far as the mortality of the global population is concerned, it is cardiovascular diseases which cause the highest death rate worldwide, mostly due to the Congestive Heart Failure (CHF). Therefore, an initial detection and diagnosis of CHF becomes essential. This manuscript presents a novel approach to detect health of CHF subject which is based on Multiresolution Wavelet Packet (MRWP) decomposition method, attributes ranking approach, kernel principle component analysis (KPCA) and 1 – Norm Linear Programming Extreme Learning Machine (1 – NLPELM). For this investigation, the heart rate variability (HRV) signal has been decomposed up to 5-level using MRWP decomposition method. The sixty three log root mean square (LRMS) attributes were extracted from the decomposed HRV signal. The top ten attributes are selected by ranking approaches such as Fisher, Wilcoxon, Entropy, Bhattacharya, and receiver operating characteristic (ROC). The ten ranked attributes were then mapped to one new feature by KPCA and fed to 1 – NLPELM. The HRV database of normal subjects (normal sinus rhythm (NSR), age 22–45 years old and elderly (ELY), age 60–82 years old) and CHF subjects (age 32–71 years old) were obtained from PhysioNet ATM. The simulation results demonstrated that Bhattacharya + KPCA with 1 – NLPELM approach achieved an accuracy of  $98.44 \pm 1.4\%$ ,  $99.13 \pm 1.85\%$  for NSR – CHF and ELY – CHF respectively. Out of all ranking methods, Bhattacharya combined with KPCA + 1 – NLPELM provided the highest degree of accuracy for all datasets. In addition, the proposed method has also achieved very good generalization performance and less execution time as compared to 1 – NLPELM, KPCA + PNN, KPCA + SVM, probabilistic neural network (PNN) and support vector machine (SVM).

**Keywords** Bhattacharya ranking method · Regularization parameters · Multiresolution Wavelet Packet (MRWP) decomposition · Quadrature mirror filter

---

✉ Ram S. Singh  
ramsewaknitj@gmail.com

Extended author information available on the last page of the article

## Abbreviations

AI	Artificial intelligence
ANS	Autonomic nervous system
CRD	Cardiac resynchronization-defibrillator
CVDS	Cardio vascular diseases
CHF	Congestive heart failure
GDA	Generalized discriminant analysis
HRV	Heart rate variability
1-NLPELM	1-Norm linear programming extreme learning machine
ELY	Elderly
ELM	Extreme learning machine
ECG	Electrocardiogram
KPCA	Kernel principle component analysis
LRMSF	Log root mean square features
ML	Machine learning
$\mu_1$ and $\mu_2$	Mean of 1st and 2nd group of Jth attributes (features)
MRWP	Multiresolution wavelet packet
NSR	Normal sinus rhythm
N1	Number of first group attributes
N2	Number of second group attributes
PNN	Probabilistic neural network
RBF	Radial basis function
ROC	Receiver operating characteristic
SCA	Sudden cardiac arrest
SVM	Support vector machine
$\beta$	The weight between hidden node and output
V1 and V2	Variance of 1st and 2nd group of Jth attributes
W	Wavelet functions
WP	Wavelet packet
G(W,B,X)	Activation function(wight, biase, input attribute)
L	Class label for given attributes

## 1 Introduction

Congestive Heart Failure (CHF) is a cardiac disorder that affects four million people globally per year without any externally instantly recognizable signs [38]. CHF is caused by spectacular forfeiture of cardiac function and is therefore more linked to an electrical pulse heart condition than an arterial obstruction. Sudden cardiac arrest (SCA) is always lead by CHF, this is the heart's failure in which heart does not effectively transfer blood to numerous tissues, resulting in lack of oxygen and the subject becomes eventual unconsciousness after one minute [14]. A cardiac resynchronization-defibrillator (CRD) has been used to resurrect CHF subjects by providing an electrical current around one to forty Joules energy of heart to activate its electrical activity [9]. The cardiopulmonary resuscitation can also help CHF subjects to avoid death till medical treatment is gotten [31]. Presently, worldwide research work has concentrated on the serious health problematic with the aim of finding an effective way to envisage the risk of CHF by non-invasive and invasive methods [1, 22,

17, 35]. To avoid immediate SCA, initial stage identification of CHF is prime importance and for this, an efficient artificial intelligence (AI) is required to diagnose fast and accurate.

The heart rate variability (HRV) is the most conspicuous non-invasive physiological indicator that is employed to identify heart anomalies, and is highly endorsed for both clinical and non-clinical uses. HRV may simply be defined as the time difference between consecutive R beats on Electrocardiogram (ECG) morphological. Morphological reflection of the ECG signal is fundamental but not satisfactory to discern the existence of CHF as subtle variations are difficult to be identified by ECG alone in CHF subjects [26]; this is often must to characterize these ECG signals into a HRV time series data [41]. In order to maintain homeostasis and activity of organs system, a healthy cardiac system can readily identify and adapt according to evolving demands of metabolism of systems [36]. However, invariant heart rate is related to cardiac disease, for example as CHF and SCA [27]. Hence, HRV variations provide information to determine overall heart fitness, along with the status of the autonomic nervous system (ANS) that controls cardiac function [37].

The most recent two decades a wide range of machine learning (ML) strategies, features ranking and dimension reduction of features have been broadly applied to cardiac diseases prognosis, detection and classification [28, 24]. A large portion of these works utilize ML strategies for demonstrating the detection of CHF and distinguish useful informative causes that are used subsequently in a detection pattern. Major categories of ML methods including artificial neural networks (ANNs) and decision trees (DT) have been utilized for automatic cardiac disease detection [24, 32]. By far most of these authors utilize at least one ML method and incorporated database from multifarious sources like Physionet ATM for the detection of congestive heart failure (CHF), CAD and myocardial ischemia (MI) just as for the prediction and classification of a heart type diseases [7]. An emerging ML is noticed the most recent decade in the utilization of other supervised ML methods, to be specific probabilistic neural network (PNN), support vector machine (SVM), and back propagation neural networks (BPNs) towards cardio vascular diseases (CVDs) detection and prognosis [32, 3]. All such ML algorithms were commonly used by authors in a wide variety of related issues of cardiac diseases but they not discussed about generalization performance of ML, ranking of features and execution validation time for prognosis of CAD disease.

In several areas of biomedical signal investigation, a recent learning method for a single hidden layer of feedforward neural networks (SLFNs) architecture called extreme learning machine (ELM) technique has been widely used for pattern recognition, detection and classification of cardiac diseases [8, 20]. Like ELM, the PNN is also a three-layer feed-forward ML system containing an input layer, a hidden layer, a summation layer with output layer. In contrast to the commonly used PNN algorithm, ELM weight parameters of between input node and hidden node may be allocated arbitrarily independently of training database and need not be tuned, and the output weights are in tune based on applications confines. Due to these properties it also has a fast learning speed and convergence rate. ELM shows better generalization performance compared to PNN due to minimum training error and the smallest norm for anonymous output weights. ELM has been commonly used in prediction and detection of cardiac diseases due to its broad variety of attributes mapping like sigmoid, radial basis function (RBF), multiquadric activation function and it also capable to manage big and small database well. However, an ELM problem is that ELM's classification boundary might not be a perfect one for the input weights of hidden nodes and biases being distributed randomly while remaining unaffected during the training phase [10, 12]. This problems leads to increase substantial number of hidden to accomplish a sufficient level of generalization performance in terms of accuracy. By Bartlett's concept [5]

reducing the norm of the weights parameters of outputs leads to improved generalization performance at reduced number of hidden nodes. In this paper, a modified version of ELM that is known as 1 – *normlinear* programming extreme learning machine (1 – NLPELM) employed to improve generalization performance with less number of hidden node and better classification performance. This result was obtained by solving a set of line equations at a finite numeral of times using 1-norm concept. As per our knowledge through literature survey, this is first time, GDA and 1-NSELM binary classifier method is applied to the clinical data included here.

This paper is organized in the following way: Section 2: Methodology, 2.1: HRV Database and Pre-Processing, 2.2: Features Extraction by Multiscale Wavelet Packet Decomposition, 2.3: Features Ranking Methods 2.4: Kernel Principal Component Analysis, 2.5: 1- Norm Linear Programming ELM, 2.6: Activation and performance parameters for Simulation of 1 – NLPELM, Section 3: Results and discussion 4. Conclusion.

## 2 Methodology

Figure 1 shows a block diagram of proposed model for CHF detection. The free noise and artifact (pre-processed) HRV time series samples are fed to the feature extraction step. The HRV signal has been decomposed up to 5-level using Multiresolution Wavelet Packet (MRWP) decomposition method using Haar mother wavelet. The advantage of this mother wavelet is that the analysis of HRV signals with sudden transitions can be sharply detected as compared to other mother wavelet, such as investigation of cardiac diseases. The sixty three log root mean square (LRMS) attributes were extracted from decomposed HRV signal. All the features are not sensitive to escalation for interpretation and comprehension of healthy and CHF subjects. Therefore, the features were ranked using Fisher score, Bhattacharya space, Wilcoxon, receiver operating characteristics (ROC) and entropy ranking methods. The most important top ten ranked features were applied to attributes space transformation method such as Kernel principle component analysis (KPCA). The KPCA transfer top ten attributes to a new attribute using radial basis function (RBF). The values of new attributes were first regularized in range of  $-1$  to  $1$ , after this, normalized attributes were fed to 1 – NLPELM classifier. The Sigmoid/ Multiquadric activation function has been employed in 1 – NLPELM to introduce nonlinearity into the output of a hidden node.

### 2.1 HRV database and pre-processing

In this study, two repositories have been used; namely, the repositories MIT / BIH SCD both PhysioNet Bank ATM Normal Sinus Rhythm (NSR) databases [11, 33]. Three patients' HRV signals had been removed from the SCD repository as their heart rates were slowed. But no signal was removed from the NSR sample, as it was without any anomalies. A complete description of the two databases is shown in Table 1.

*Pre-processing* The HRV time series obtained from the standard database usually contains ectopic beats (irregular impulse formation in the heart muscle) and non-stationarity, which precludes effective feature extraction and HRV data analysis. To avoid this problem, the pre-processing of the HRV signal is required [34, 40]. After pre-processing, the HRV is known as a normal-normal interval (NN interval). The NN intervals were re-sampled at 4 Hz.

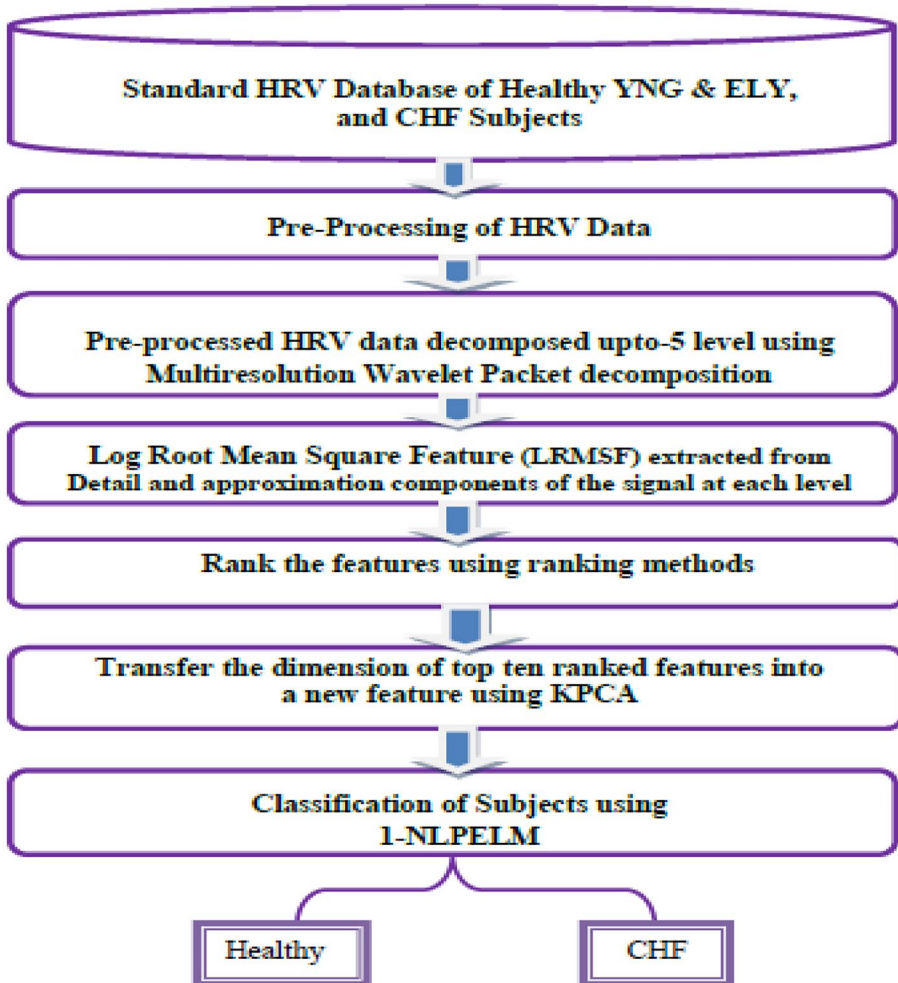


Fig. 1 Represents a block diagram of proposed model for CHF detection

## 2.2 Features extraction by multiresolution wavelet packet decomposition

Multiresolution Wavelet Packet (MRWP) decomposition is derived from decomposition of Wavelets method. It involves numerous frameworks, so different bases can vary in varying output of distinction and will fill the deficiency of time–frequency decomposition in discrete wavelet transform [15]. The decomposition of the wavelet divides the original sample into two feature space,  $V$  and  $W$  that are mutually perpendicular. Orthogonal to each other,  $V$  is the space that contains details on the received data at low frequencies, and  $W$  contains details on the higher frequencies. A wavelet packet (WP) represents a set of linearly grouped wavelet functions produced as defined in Eq. (1) and (2) by the following recursive associations [18].

**Table 1** Details of HRV databases used in this paper

Details	St. Petersburg Institute of Cardiological Technics CHF database	MIT/BIH Normal Sinus Rhythm (NSR) Healthy database	Fantasia Healthy database
Number of subjects	18	18 Healthy	20 Young & 20 Elderly
Subject's description	15 subjects had CHF, 3 was paced	All in sinus rhythm	All in sinus rhythm and laydown
No. of Male and Female	4 Female, age 54 to 63 and 11 Male, age 22 to 71	5 Male, age 26 to 45, and 13 Female, age 20 to 50	Young: (12 Male & 8 Female, age 21–34) & Elderly: (10 male & 10 Female, age 68–85)
Recording type	Long-term Holter recording	Long-term Holter recording	Long-term Holter recording

$$W^{2k}(t) = \sqrt{2} \sum_n S(n)W^k(2t - n) \tag{1}$$

$$W^{2k+1}(t) = \sqrt{2} \sum_n g(n)W^k(2t - n) \tag{2}$$

Here assume that there are the first two components of wavelet packets  $W^0(t) = \varnothing(t)$  and  $W^1(t) = \varnothing(t)$  are defined as scaling operator and wavelet operator. The signal  $s(n)$  and  $g(n)$  are interrelated as  $g(n) = (-1)^n s(1 - n)$  is a pair of *Quadrature Mirror Filters* (QMF) coefficients associated with the *scaling* operator and the *wavelet* operator. MRWP repetitively disintegrated the discrete  $x(t)$  signal into low frequency (LF) well-known as Approximation ( $X_{j+1,2k}(t)$ ) and high frequency (HF) defined as Details ( $X_{j+1,2k+1}(t)$ ) constituents. The input signal  $x(t)$  which is described in Eqs. (3) and (4) may be disintegrated recursively

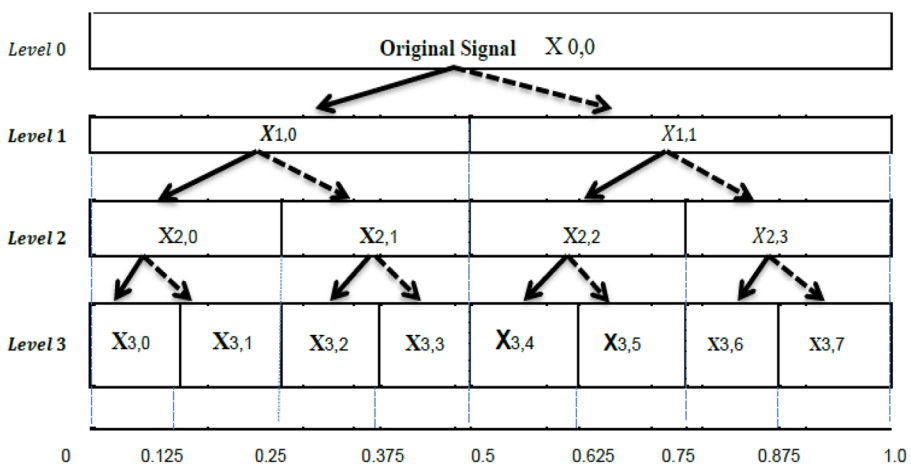
$$X_{j+1,2k}(t) = \sum_m s(m - 2n)x_{j,k}(t) \tag{3}$$

$$X_{j+1,2k+1}(t) = \sum_m g(m - 2n)x_{j,k}(t) \tag{4}$$

Here  $X_{j+1,2k}$  symbolizes the MRWP coefficients at the  $j^{th}$  level,  $k^{th}$  sub frequency group. Hence, the input signal  $x(t)$  can be defined as Eq. (5).

$$X(t) = \sum_{k=0}^{2^j-1} x_{j,k}(t) \tag{5}$$

The 3<sup>rd</sup> level decompositions of HRV signal  $X(t)$  using the MRWP is shown in Fig. 2. In this figure, a bold line denotes LF parts and a spotted line represents HF components. The log root mean square feature (*LRMSF*) of decomposed HRV signal was evaluated. This logarithmic tool is selected as an attribute to detect CHF subject due to its denunciation to nonlinear compoment of HRV signal [25]. The *LRMSF* attribute is expressed in Eq. (6).



**Fig. 2** Demonstrates 3 levels decomposition of HRV signal using MRWP method, horizontal axis denotes frequency variation as a fraction of the sample frequency. The  $X_{1,0}$ ;  $X_{1,1}$ ;  $X_{2,0}$ ..... represents LF and HF components of HRV

$$LRMSF = \log \sqrt{\frac{X_1^2 + X_2^2 + \dots + X_N^2}{N}} \tag{6}$$

where  $X_1, X_2 \dots$  e.t.c represents samples of decomposed signal and  $N$  is the total number of samples in a decomposed HRV. Total number of attributes at each level is calculated by  $2^{level}$ . As, HRV is nonlinear signal, for appropriate analysis of HRV signal, nonlinear features are required. LRMS is also nonlinear methods which show nonlinear behaviour as HRV signal. So, in this article log root mean square (LRMS) features have been extracted to MRWP decomposition of HRV image signal.

### 2.3 Features ranking methods

The LRMSF features extracted from decomposed HRV signal have substantial info about the cardiac disease. Based on the info found in the attributes, it can be graded as being extremely important, weakly important, obsolete and redundant [39]. Obsolete and redundant attributes decreases the classifier performance and increase processing time of the ML system. In this case the ranking scheme is very useful in selecting appropriate attributes. For this, in this paper Fisher score, Bhattacharya space, Wilcoxon, receiver operating characteristics (ROC) and entropy ranking methods were employed to choose top ten rank attributes [2]. The Fisher score is attained by applying Eq. (7). The Fisher score of  $i^{th}$  attribute in  $j^{th}$  level matrix is define as

$$F(i) = \frac{\sum_{j=1}^n N_j (\mu_{i,j} - \mu_i)^2}{\sum_{j=1}^n N_j (\sigma_{i,j})^2} \tag{7}$$

where  $\mu_i$  denotes average value of  $i^{th}$  attribute,  $\mu_{i,j}$  denotes average value of  $j^{th}$  attribute in  $i^{th}$  matrix,  $N_j$  symbolizes number of samples of  $j^{th}$  level matrix of  $i^{th}$  attributes and  $\sigma$  denotes standard deviation.

The Entropy method for ranking of attributes is defined by Eq. (8)

$$En(X^J) = \frac{\left(\frac{\mu_1}{\mu_2} + \frac{\mu_2}{\mu_1} - 2\right)}{2} + (\mu_1 - \mu_2)^2 \times \frac{\left(\frac{1}{V_1} + \frac{1}{V_2}\right)}{2} \tag{8}$$

$En(X^J)$  indicates Entropy score of  $j^{th}$  attributes.  $\mu_1$  and  $\mu_2$  mean of  $1^{st}$  and  $2^{nd}$  group of  $j^{th}$  attributes.  $V_1$  and  $V_2$  of variance of  $1^{st}$  and  $2^{nd}$  group of  $j^{th}$  attributes.

The Bhattacharya space method for ranking of attributes is defined by Eq. (9)

$$Bh(X^J) = \frac{4 \times (\mu_1 - \mu_2)^2}{(d1 + d2)} + 2 \log \left( \frac{(d1 + d2)}{\sqrt{\frac{d1 \times d2}{2}}} \right) \tag{9}$$

where  $d1 = \sqrt{V_1}, d2 = \sqrt{V_2}$  and  $Bh(X^J)$  denotes Bhattacharya score of  $j^{th}$  attributes.  $\mu_1$  and  $\mu_2$  mean of  $1^{st}$  and  $2^{nd}$  group of  $j^{th}$  attributes.  $V_1$  and  $V_2$  of variance of  $1^{st}$  and  $2^{nd}$  group of  $j^{th}$  attribute.

The Wilcoxon method for ranking of attributes is defined by Eq. (10)



$$W_x(X^j) = \text{Absolut} \left( \frac{N_2 \times \text{sum}(\text{Ranks}(\text{group}, :))}{N_1} - 1 \right) \quad (10)$$

Ranks = tiedrank(X) ,  $N_1$  = number of first group attributs,  $N_2$  = number of second group attributs.

## 2.4 Kernel principal component analysis

Principal component analysis (PCA) is a very powerful strategy for the lowering of dimensional space of attributes. PCA is trying to find a reduced-dimensionality linear attributes space than the original attributes space, where new attributes have the greatest variance. Traditional PCA only leads to a reduction of linear dimensions. After all, if the attributes have much more complex shapes and their values are closest to each other that are not well defined in a geometric attributes space, for example the pattern of nonlinear attributes are similar NSR and CHF. In case the conventional PCA is not going to be of great help. Interestingly, KPCA allows us to make generalizations conventional PCA to a decrease in non-linearity dimensionality. In this condition, a transforming the feature space can improve classification [23]. Different strategies have been proposed to reduce the dimension of attributes to classification and detection of cardiovascular diseases [13, 4, 16]. In this paper, KPCA has been applied for reduction of attributes. Its advantages are nonlinearity of eigenvectors and greater number of eigenvectors [6]. The KPCA is nonlinear extension of principal component analysis (PCA). In KPCA, for specified learning data sample is plotted by using a RBF kernel function. It maps high-dimensional attribute space, where disparate classes label of attributes are made-up to be nonlinearly discernible [30]. When there is L class label for given attributes, the dimension of attributes can be lowered by KPCA technique to  $L - 1$ . In this paper, binary classification has been used to classify subjects. Hence the top ten attributes reduced by KPCA to one new attribute.

## 2.5 1-Norm linear programming ELM

### 2.5.1 Fundamental of ELM

The authors [19, 21] has stated that the fundamentals of *ELM* can be described in three phases for given training features (X) and class label (T) set defined as  $\{X_K, T_K\}_{K=1,2,\dots,M}$  and  $X_K = \{x_{K1}, x_{K2}, \dots, x_{Kn}\}^t \in \mathbb{R}^n$ , Here  $M, K$  and  $t$  represent number of samples, number of input features to the  $n$  input layer and transpose of the input features matrix. Notation  $T_K \in \{-1, 1\}$  represents output corresponding to sample  $X_K$  for binary classification. ELM consist of  $m$  number of hidden layers nodes with activation function  $G(\cdot)$  between input and output layer. The construction of ELM is shown in Fig. 3.

Following four phases involved in learning and validation of ELM.

- First take random value of input weights  $(W) = \{W_{m1}, W_{m2}, \dots, W_{mn}\}^t$  and biases  $B = \{B_1, B_2, \dots, B_m\}$  of hidden nodes (N). This value does not change during learning and validation of ELM.
- Compute the hidden layer (N) Output by using activation function  $G(\cdot)$  as  $G(WX + B)$ .

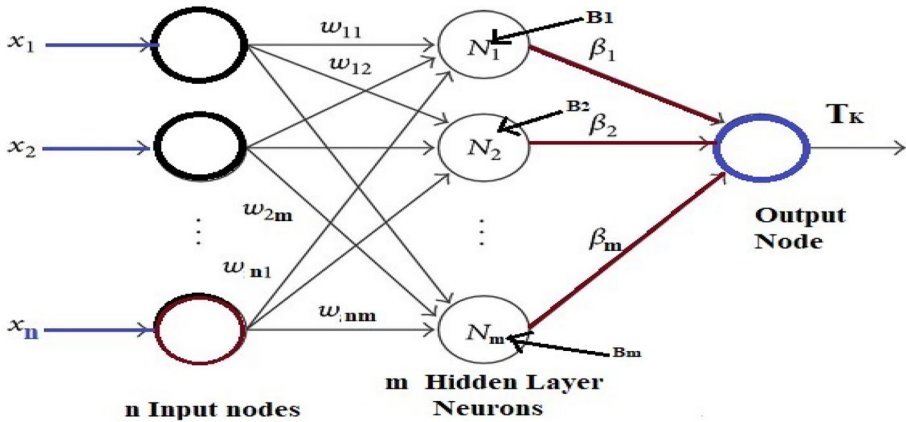


Fig. 3 Represents construction of extreme learning machine

$$N = \begin{bmatrix} G(W_1 X_1 + B_1) & \dots & G(W_m X_1 + B_m) \\ \vdots & \dots & \vdots \\ G(W_1 X_n + B_1) & \dots & G(W_m X_n + B_m) \end{bmatrix}_{n \times m} \tag{11}$$

- Compute the output weight  $\beta = N^\dagger T$ , where  $N^\dagger$  represents Moore–Penrose generalized inverse of  $N$  and  $\beta = \{\beta_1, \beta_2, \dots, \beta_m\}^t$  weight between hidden layers and output node. Minimize the learning error as well as norm of the output weights by  $\|N\beta - T\|^2$  and  $\|\beta\|^2$ .
- For binary classification or detection, the decision is based on Signum function, and calculated as

$$F(X) = \text{Sign}\{(G(W_1, B_1, X), \dots, G(W_m, B_m, X)) \beta\} \tag{12}$$

### 2.5.2 Derivation of 1- norm linear programming ELM

For accurate classification, the unknown output weight  $\beta = \{\beta_1, \beta_2, \dots, \beta_m\}^t \in R^L$  between hidden layers and output node should have smallest minimum learning error property and norm [8] and it is defined by Eq. (13). Here  $\|\cdot\|_1$  and  $\|\cdot\|_2$  specifies 1 – Norm and 2 – Norm of a data matrix.

$$\min \|N\beta - T\|_2 \text{ and } \min \|\beta\|_2 \tag{13}$$

The problem (13) can be expressed in 1-norm and represented as

$$\min_{W \in R^L} \|W\|_1 + \gamma \|HW - T\|_1, \gamma > 0, \tag{14}$$

By applying the process of [3], the 1 – NELM training problem (13) can be converted into a linear program problem (LPP) in this way:  $U, V \in R^L$  and  $P, Q \in R^m$

$$\text{Assume that } W = U - V \text{ and } HW - T = P - Q \tag{15}$$

Since,  $U, V \geq 0$  and  $P, Q \geq 0$  is necessary condition therefore using (14) in (15) can be used to get the 1 – NLPELM resolution in primitive of the method:

$$\min_{U,V,P,Q} E'_L(U + V) + \gamma \min_{U,V,P,Q} E'_m(P + Q)$$

Condition on  $H(U - V) - P + Q = T$ , such that  $U, V, P, Q \geq 0$ , (16)

where  $E_L$  and  $E_m$  represents the column vectors of order  $1 \times L$  and  $1 \times m$  respectively. Therefore, Eq. (17) can be represented in the form of linear programming problem which can be solved and the constraints of the objective function should be  $\leq 0$ . The parameters of 1 - NLPELM can be optimized by toolbox of MATLAB.

### 2.6 Variables and functions for simulation of 1-NLPELM

The 1-NLPELM can employ arbitrary hidden nodes or kernels same as ELM. In order to read easily, the ELM is evaluated from arbitrary *hiddennodes* ( $N$ ) directly.  $N(X) = (G(W_1, B_1, X), \dots, G(W_m, B_m, X))$  where  $G$  is the activation function,  $G$  essentials to placate ELM universal approximation conditions [7].

$$\Omega_{ELM} = NN^T \tag{17}$$

Two kinds of hidden nodes could be employed at middle layer of ELM: additive nodes and radial basis function (RBF) nodes. In the following, the former two are additive nodes and the latter two are RBF nodes

Sigmoid Function is defined as  $G(W, B, X) = \frac{1}{1 + \text{Exp}^{-(WX+B)}}$  (18)

Sinusoid Function is defined as  $G(W, B, X) = \text{Sin}(WX + B)$  (19)

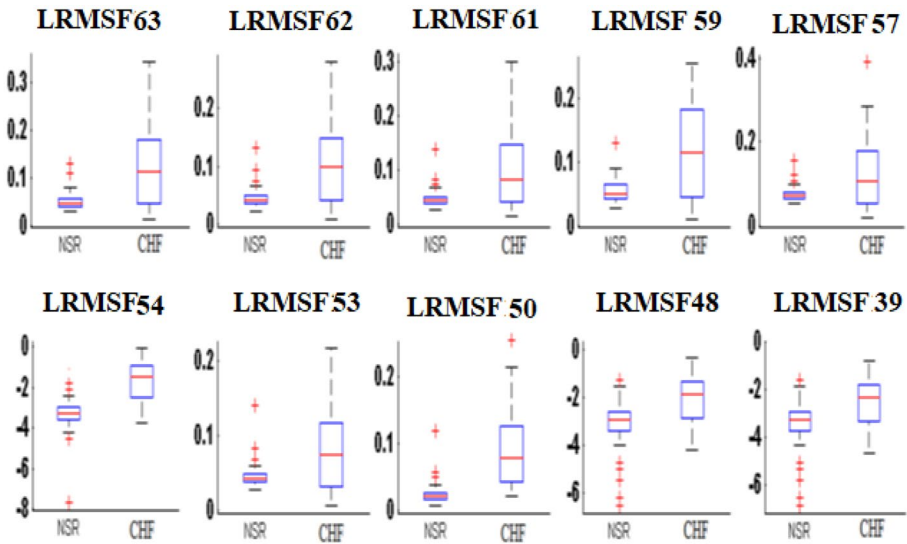
Multiquadric Function is defined as  $G(W, B, X) = \sqrt{\|X - W\|^2 + B^2}$  (20)

Gaussian Function is defined as  $G(W, B, X) = \text{Exp}^{(-\frac{\|X-W\|^2}{B})}$  (21)

Intended for *Multiquadric* and *Sigmoid* activation functions of hidden node, the hidden node layer constraints have been selected randomly with unvarying uniform distribution in  $[-0.6, 0.6]$  to satisfy the ELM condition. The value of  $B = \{B_1, B_2, \dots, B_m\}$  of  $N$  and weights  $(W) = \{W_{m1}, W_{m2}, \dots, W_{mm}\}^t$  were chosen stochastically at the initial of the training for 1 - NLPELM and all parameters remain unchanged for each simulation of datasets. The optimum value of  $\gamma$  (*regularizationparameter*) have been considered by performing 10 - fold cross - validation method. If the value of  $N$  is very high, this leads to higher estimated time and if the value of  $N$  is very low then its effect on classifier performance parameter decreases. The improved generalization performance could be obtained from  $N = 100$  to  $150$  (moderate value of  $N$ ) [11]. The optimum value of  $\gamma \in \{2^{-1}, \dots, 2^{40}\}$  using 10 - fold cross-validation. The average of classification performance like accuracy has been evaluated by conducting 100 random independent trials for every dataset using optimum value of  $\gamma$ .

The classification parameters of classifier were calculated by using confusion matrix and formulated as

$$\text{Accuracy (AC)} = (TP + TN)/(TP + TN + FN + FP) \tag{22}$$



**Fig. 4** Box plot of top ten ranked LRMSF attributes of NSR – CHF dataset before dimension mapping by *GDA* and *KPCA*

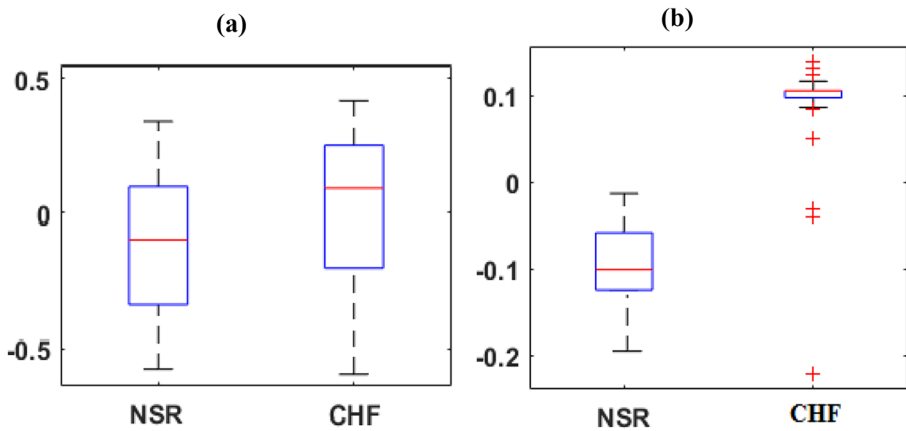
where  $TP = \text{truepositive}$ ,  $FN = \text{falsenegative}$ ,  $TN = \text{truenegative}$  and  $FP = \text{falsepositive}$ .

Area under the curve (AUC) is defined as  $AUC = \int_0^1 ROC(\tau) d\tau \cong \frac{1}{2}(Se + Sp)$ . Where  $\tau = (1 - Sp)$  and  $ROC(\tau)$  is sensitivity.

### 3 Result and discussion

#### 3.1 Performance evaluation of KPCA for NSR-CHF dataset

The KPCA is an attributes size reduction method which maps the dimensions of attributes using nonlinear kernel function same as generalized discriminant analysis (GDA) [46] but KPCA with RBF kernel has more capability to separate completely linearly the two classes (NSR-CHF group) compared to GDA. The box plot of top ten attributes were chosen by using Fisher score method is shown in Fig. 4. To comprehend the separate completely of *KPCA* and *GDA* using *RBF* kernel function after reduction of size of attributes, the box-plots of the top ten ranked LRMSF for NSR-CHF data sets are shown in Fig. 4 and after reduction of top ten LRMSF by *KPCA* are illustrated in Fig. 5a; and by *GDA* are shown in Fig. 5b. The box-plot patterns (median value indicated by red line in box) of ten LRMSF associated with NSR – CHF sets are positioned very near in their vicinity before reduction of attributes. After attributes size reduction by *GDA* and *KPCA* method, the new attributes is well separated within the attributes size. So the new attributes provides not only improvements in the classification capability, it also make suitable tool for a well discrimination of *NSR – CHF* set. As perceived by box plot of Fig. 5a and b that *KPCA* produces more separate linearly pattern in NSR-CHF group compared to *GDA* after attributes size



**Fig. 5** Box plot, after dimension mapping of top ten ranked LRMSF to a new attribute for NSR – CHF dataset using (a) GDA with radial basis function. (b) KPCA with Radial Basis Function

reduction. Hence, KPCA reduction scheme can be used for dimension reduction of top ten LRMSF extracted from NSR-CHF dataset.

### 3.2 Classifier performance using AUC and ranked features

In this section, the AUC value achieved by machine learning (ML) such as KPCA + 1 – NLPELM, KPCA + SVM, KPCA + PNN, 1 – NLPELM, SVM and PNN has been demonstrated at number of ranked attributes. To comprehend the performance of machine learning, a graph between AUC and top ten ranked attributes has been simulated which are presented in Fig. 6 for NSR – CHF (Fig. 6a), for NSR – ELY (Fig. 6b) and for ELY – CHF (Fig. 6c). The rank of attributes was assigned using Bhattacharya method. Ranked top ten attributes were fetched to the ML one after another (like 2, 3, 4, 5—10 attributes). For dimension reduction of attributes by KPCA, initially minimum two attributes are required and KPCA always reduced dimension of attributes to one new attributes in each fetch. In this simulation, the Multiquadric and Gaussian function were employed in 1 – NLPELM and SVM ML. Figure 6a demonstrates that the highest AUC (0.83) value has been achieved by KPCA + 1 – LPELM for NSR – CHF dataset when more than seven attributes are fetched to it. After this ML, the KPCA + SVM achieved very good AUC compared to KPCA + PNN, 1 – NLPELM, SVM and PNN for each datasets. Figure 6b depicts that KPCA + 1 – NLPELM also accomplished highest AUC (0.72) among all considered ML after eight attributes for NSR – ELY dataset. Figure 6c shows that the proposed (KPCA+1-NLPELM) ML attained continuous AUC around 0.76 for ELY – NSR after 7 to 10 attributes while KPCA + SVM acquired constant AUC around 0.72 after nine attributes. Figure 6 shows that KPCA + 1 – NLPELM is an appropriate ML for detection and classification of CHF datasets. This figure also reveals that all considered ML achieved poor AUC if number of attributes fed to ML is less than eight.

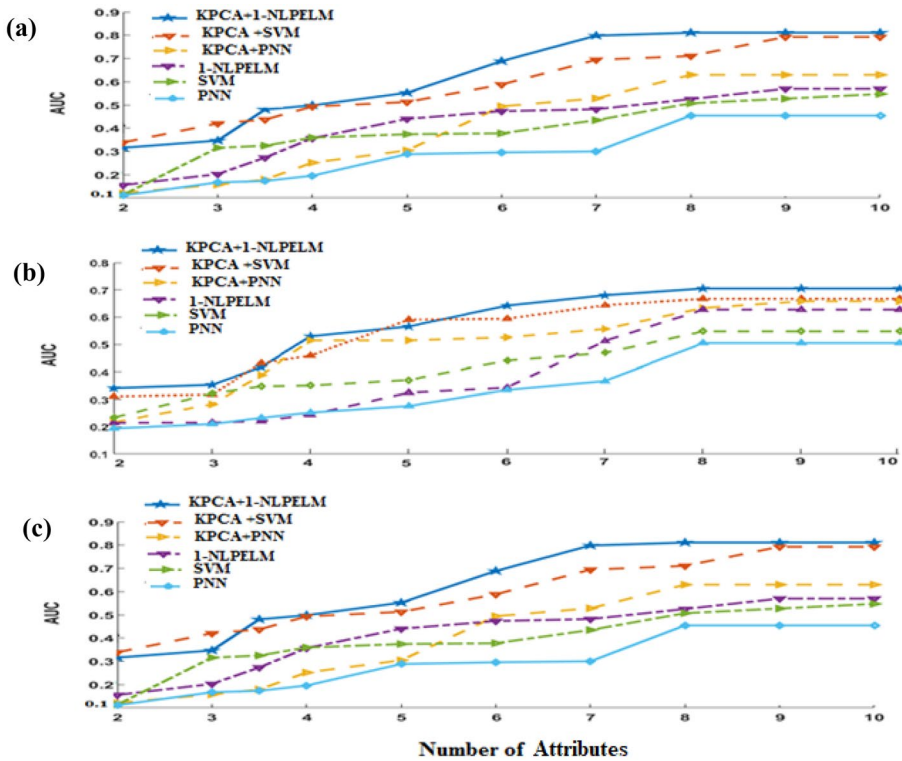
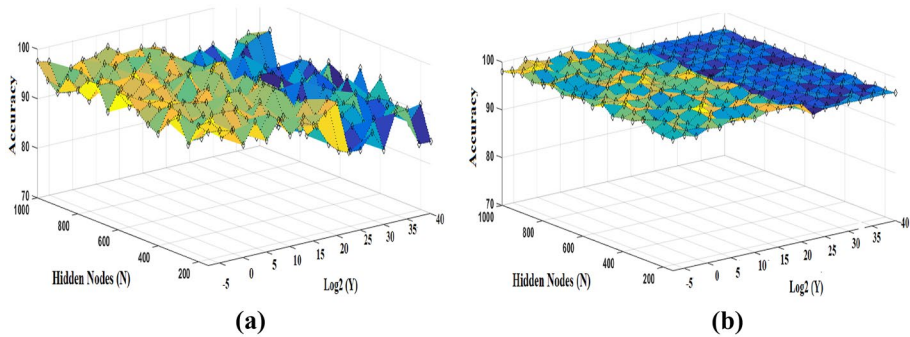


Fig. 6 Demonstrates the AUC value achieved by MLs at number of ranked attributes for (a) *NSR – CHF* (b) *NSR – ELY* (c) *ELY – CHF* datasets

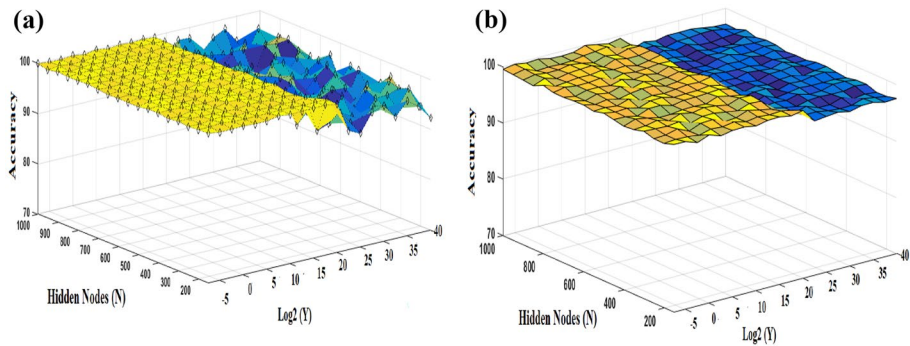
### 3.3 Generalization performance of proposed method

This result section presents the validation of generalization performance of proposed *KPCA + 1 – NLPELM* method and *1 – NLPELM* classifier on *N* and  $\gamma$  in the gravel of the activation function (See Equation no.12) for classification and detection of *NSR – CHF* dataset. For the analysis of generalization performance, the  $\gamma$  varies between  $\{2^{-1}, 2^{40}\}$  and *N* varies from  $\{50, 60 \dots 100, 300, 400, 800 \text{ and } 1000\}$ .

Figure 7 depicts the value of accuracy for variation in every pair of  $\gamma$  and *N* for *1 – NLPELM* with *Sigmoid* additive node and *Multiquadric* radial basis function node. Figure 7a demonstrates that as  $\gamma$  increases from  $2^{-1}$  to  $2^{25}$ , with variation of *N* from 50 to 1000, the accuracy (AC) changes from 87 to 94% for *1 – NLPELM* with *Sigmoid* additive function node for *NSR – CHF* dataset whereas the AC falls below 90% if  $2^{25} < \gamma < 2^{40}$  and  $450 < N < 1000$ . It can be observed from Fig. 7b that as  $\gamma$  increases from  $2^{-1}$  to  $2^{25}$  with variation of *N* from 50 to 1000, the validation AC changes from 93 to 95% for *1-NLPELM* with *Multiquadric* radial basis function node for *NSR – CHF* dataset but AC rapidly falls if  $2^{25} < \gamma < 2^{40}$  and  $600 < N < 1000$ . It can be inferred from Fig. 7 that the validation AC is much more sensitive for every pair of  $\gamma$  and *N*. So it can be seen that the generalization performance of this method is satisfactorily good with the use of *1 – NLPELM* with *Multiquadric* radial basis function node.



**Fig. 7** Generalization performance of 1 – NLPELM on the  $\gamma$  and N for (a) 1 – NLPELM with Sigmoid additive node for NSR – CHF dataset; (b) 1 – NLPELM with Multiquadric radial basis function node for NSR – CHF dataset



**Fig. 8** Generalization performance of proposed model using (a) KPCA with RBF + 1 – N LPELM with Sigmoid additive node for NSR – CHF dataset (b) KPCA with RBF + 1 – NLPELM with Multiquadric Radial Basis Function node for NSR – CHF dataset

Figure 8 illustrates the value of accuracy for variation in every pair of  $\gamma$  and  $N$  for KPCA + 1 – NLPELM with Sigmoid additive node and Multiquadric radial basis function. Figure 8a depicts that as  $\gamma$  lies between  $2^{-2} < \gamma < 2^{25}$  and  $N$  lies between  $50 < N < 400$ , the validation AC increases from 99% to 99.54% for NSR-CHF dataset when Sigmoid additive function node is used with KPCA + 1 – NLPELM. It can also be observed that if the value of  $\gamma$  varies from  $2^{25} < \gamma < 2^{40}$ , the validation AC falls below 99%. It can be observed from Fig. 8b that with the variation in value  $\gamma$  and  $N$  ( $2^{-2} < \gamma < 2^{25}, 50 < N < 400$ ), the validation AC remains constant at 99.96% when Multiquadric RBF node is used with KPCA + 1 – NLPELM. It can be concluded from Fig. 8 that for every pair of  $\gamma$  and  $N$  ( $2^{-2} < \gamma < 2^{25}, 50 < N < 400$ ), the proposed model achieved much better generalization performance as compared to 1 – NLPELM. Hence proposed model can be used for the detection of cardiac diseases with higher degree of accuracy using minimum number of hidden nodes and independent to user specified parameters.



### 3.4 Statistical comparison of attributes

A Z-score test of one-tailed type has been employed to compare and analyze statistically the attributes retrieved by LRMSF from decomposed HRV datasets. The value of P obtained in this test signifies about the attributes whether the attributes are statistically significant for the detection of CHF. For the attributes to be significant statistically, the P-value of the test should be preferably low [58–45]. A low P-value of the test signifies that the probability of reparability of datasets will be considerably high. For the diagnosis of any disease statistically, a p-value  $\leq 0.05$  at 95% confidence limit is said to have considerable significance\*, if p-value  $\leq 0.001$  then it said to be very significant \*\* and if p-value  $> 0.05$  then it said to be statistically insignificant.

The results listed in Table 2 shows that the 4th and 5th level LRMSF attributes extracted from NSR-CHF datasets have p-values  $< 0.001$ , 3rd level have p-values  $< 0.01$  whereas 1st and 2nd level attributes have p-values  $> 0.05$ . Therefore (4th, 5th), 3rd and (1st, 2nd) levels attributes are very most significant, considerably significant and insignificant respectively. It can be inferred that the p-value achieved by 4th and 5th level LRMSF attributes retrieved from NSR-ELY and ELY-CHF datasets have values  $< 0.001$  however the levels 1st, 2nd and 3rd have p-values  $> 0.05$ . Therefore 4th and 5th level of LRMSF attributes are found to be statistically significant for the analysis of CHF disease. A. Kampouraki et al. [23] investigated that statistical significance ( $p < 0.05$ ) may not be suitable method for detection of any cardiac disease, because the  $p < 0.05$  value was carried out by using simple threshold scheme. Hence, efficient classifier and reduction method is required for detection of CHF.

### 3.5 Performance of proposed method with ranking methods

The classification validation AC has been evaluated in order to demonstrate the performance of ranking methods+1 – NLPELM and ranking methods+KPCA + 1 – NLPELM. The Sigmoid additive node and Multiquadric radial basis function node has been used with each of the classifiers. The average AC performance was evaluated using a 100 trials and tenfold cross validation method. In each trials, the data used for training in NSR – ELY, NSR – CHF and ELY – CHF are 100 out of 160, 90 out of 140 and 80 out of 140 respectively whereas the remaining data were used for validation purpose which are listed in Table 3. The AC error rate was measured in the form of standard deviation ( $\pm$ S.d).

Table 3 shows the results achieved by various ranking methods in association of 1 – NLPELM and KPCA + 1 – NLPELM. It can be seen from the Table 3 that the Fisher+proposed method in association with Sigmoid and Multiquadric activation function yields an AC of  $97.32 \pm 1.25\%$  and  $98.16 \pm 1.32$  for NSR – CHF dataset respectively whereas an AC of  $72.54 \pm 5.04$  and  $76.77 \pm 4.7$  has been achieved with Fisher + 1 – NLPELM. With the same proposed method an AC of  $95.93 \pm 1.12\%$  and  $98.44 \pm 1.4\%$  has been achieved when the Bhattacharya ranking method is used whereas ROC + KPCA + 1 – NLPELM achieved an accuracy of  $97.19 \pm 0.61$  and  $98.24 \pm 0.96\%$ . The proposed method combined with Fisher, Wilcoxon, Bhattacharya, Entropy and ROC achieved an accuracy of  $96.76 \pm 3.94$ ,  $95.32 \pm 1.79$ ,  $98.12 \pm 1.85$ ,  $97.11 \pm 1.51$  and  $98.82 \pm 1.50$  respectively. From listed results in Table 3, it can be concluded that when ranking methods combined with proposed method Bhattacharya method is one which provides the highest accuracy among all analyzed ranking methods for considered datasets. Finally, it can be concluded that the KPCA when combined with 1 – NLPELM, the detection accuracy of CHF can be increased to a greater extent.



**Table 2** Reperesents P-values have been achieved by LRMSF features for each datasets

Features	LRMSF1	LRMSF 2	LRMSF 3	LRMSF4	LRMSF 5	LRMSF 6	LRMSF 7	LRMSF 10	LRMSF 11	LRMSF 13
<i>NSR – CHF</i> (p-Value)	1.23	1.10	4.11	1.71	4.21	6.41	5.15	1.11E – 03 **	2.93E – 5 **	1.08E – 5 **
	LRMSF 19	LRMSF 27	LRMSF 39	LRMSF 42	LRMSF 51	LRMSF 58	LRMSF 60	LRMSF 61	LRMSF 62	LRMSF 63
	7.19E – 6 **	1.16E – 6 **	1.99E – 7 **	4.5E – 7 **	1.60E – 7 **	1.5E – 8 **	3.12E – 13**	2.51E – 12 **	2.4E – 14 **	1.7E – 12 **
<i>Features</i>	<b>LRMSF1</b>	<b>LRMSF 2</b>	<b>LRMSF 3</b>	<b>LRMSF4</b>	<b>LRMSF 5</b>	LRMSF 6	LRMSF 7	LRMSF 8	LRMSF 9	LRMSF 10
	4.19	2.10	0.0027 **	1.7	0.16	0.0003 **	0.023	2.16	0.791	0.112
	LRMSF 27	LRMSF 34	LRMSF 51	LRMSF 47	LRMSF 55	LRMSF 57	LRMSF 59	LRMSF 60	LRMSF 62	LRMSF 63
<i>NSR – ELY</i> (p-Value)	0.213	0.0004 **	0.00046 **	0.143	0.0041 **	1.2E – 4 **	1.1E – 5 **	0.00021 **	1.4E – 6	1.7E – 7 **
	<b>LRMSF1</b>	<b>LRMSF 2</b>	<b>LRMSF 3</b>	<b>LRMSF4</b>	<b>LRMSF 8</b>	LRMSF 6	LRMSF 7	LRMSF 13	LRMSF 9	LRMSF 10
	3.16	2.15	0.01021 *	1.10	0.1494	0.0067 **	0.0308 **	1.20	0.629	0.0178 *
<i>ELY – CHF</i> (p-Value)	LRMSF 11	LRMSF 27	LRMSF 34	LRMSF 37	LRMSF55	LRMSF57	LRMSF 60	LRMSF 61	LRMSF 62	MSWPF63
	0.216	0.00368 **	0.0079 **	0.0628	0.0085 **	0.0007 **	0.0004 **	0.000023 **	0.00001 **	0.000018 **

**Table 3** Demonstrates the comparative performance of  $1 - NLPELM$  with ranking methods and  $KPCA + 1 - NLPELM$  with ranking methods in terms of AC (%)  $\pm$  S.D for considered datasets

Database (Trainsize, Testsize)	Fisher + $1 - NLPELM$		Fisher + $KPCA + 1 - NLPELM$	
	Sigmoid	Multiquadric	Sigmoid	Multiquadric
NSR – CHF(90 × 10,50 × 10)	72.54 ± 5.04	76.77 ± 4.7	97.32 ± 1.25	98.21 ± 1.32
NSR – ELY(100 × 10, 60 × 10)	74.36 ± 4.40	88.71 ± 5.31	96.07 ± 4.01	98.84 ± 1.07
ELY – CHF(80 × 10, 60 × 10)	71.64 ± 2.89	79.87 ± 7.51	95.12 ± 4.12	97.16 ± 3.94
Database (Trainsize, Testsize)	Wilcoxon + $1 - NLPELM$		Wilcoxon + $KPCA + 1 - NLPELM$	
	Sigmoid	Multiquadric	Sigmoid	Multiquadric
NSR – CHF(90 × 10,50 × 10)	78.58 ± 8.06	82.41 ± 7.21	<b>96.06 ± 3.17</b>	<b>97.91 ± 1.43</b>
NSR – ELY(100 × 10, 60 × 10)	73.64 ± 7.62	85.83 ± 2.3	94.79 ± 2.77	95.32 ± 1.79
ELY – CHF(80 × 10, 60 × 10)	88.8 ± 8.48	89.36 ± 9.02	95.70 ± 3.31	98.49 ± 1.54
Database (Trainsize, Testsize)	Bhattacharya + $1 - NLPELM$		Bhattacharya + $KPCA + 1 - NLPELM$	
	Sigmoid	Multiquadric	Sigmoid	Multiquadric
NSR – CHF(90 × 10,50 × 10)	72.41 ± 7.9	86.54 ± 4.87	95.93 ± 1.12	<b>98.44 ± 1.4</b>
NSR – ELY(100 × 10, 60 × 10)	82.07 ± 3.49	89.03 ± 7.49	96.27 ± 4.37	<b>97.4 ± 1.45</b>
ELY – CHF(80 × 10, 60 × 10)	71.58 ± 6.87	79.2 ± 4.19	97.06 ± 5.39	99.13 ± 1.85
Database (Trainsize, Testsize)	Entropy + $1 - NLPELM$		Entropy + $KPCA + 1 - NLPELM$	
	Sigmoid	Multiquadric	Sigmoid	Multiquadric
NSR – CHF(90 × 10,50 × 10)	79.35 ± 16.12	89.66 ± 4.47	96.32 ± 1.29	98.55 ± 1.52
NSR – ELY(100 × 10, 60 × 10)	80.78 ± 3.55	86.55 ± 4.35	94.54 ± 5.58	96.34 ± 1.17
ELY – CHF(80 × 10, 60 × 10)	78.11 ± 9.12	84.35 ± 4.94	95.06 ± 2.89	97.11 ± 1.51
Database (Trainsize, Testsize)	ROC + $1 - NLPELM$		ROC + $KPCA + 1 - NLPELM$	
	Sigmoid	Multiquadric	Sigmoid	Multiquadric
NSR – CHF(90 × 10,50 × 10)	78.35 ± 5.19	86.53 ± 4.42	97.19 ± 0.61	98.24 ± 0.96
NSR – ELY(100 × 10, 60 × 10)	82.53 ± 5.34	89.81 ± 5.03	95.55 ± 7.96	96.86 ± 1.92
ELY – CHF(80 × 10, 60 × 10)	79.42 ± 5.18	82.81 ± 8.52	97.96 ± 2.37	98.82 ± 1.50

Bolded text in table shows the best results

Table 4 presents the results of execution performance in terms of validation time (in second) for datasets. In comparison to other classifiers, the proposed  $KPCA + 1 - NLPELM$  with Sigmoid additive activation function consumes the least time. The validation time of  $KPCA + 1 - NLPELM$  is a little longer than that of  $KPCA + LPELM$  with Sigmoid but shorter than that of  $1 - NLPELM$  with both activation function. Out of all learning schemes, PNN consumes highest execution time for detection of CHF. The results of Table 4 reveals that proposed method does not require too much computational time to diagnose CHF disease and is easy to design too.

**Table 4** Validation times (in seconds) compared for all six methodologies for classification of data sets namely NSR-ELY, NSR- CHF and ELY-CHF

<i>Methods</i>	<i>NSR – ELY</i>	<i>NSR – CHF</i>	<i>ELY – CHF</i>
KPCA + 1 – NLPELM (Sigmoid)	5.26	4.223	4.223
KPCA + 1 – NLPELM (Multiquadric )	7.11	5.41	5.41
KPCA + SVM (Sigmoid)	342.52	208.43	208.43
KPCA + SVM (Multiquadric )	389.23	210.07	210.07
KPCA + PNN	192.14	120.17	120.17
1 – NLPELM (Sigmoid)	52.62	39.55	39.55
1 – NLPELM (Multiquadric)	72.64	54.13	54.13
SVM (Sigmoid)	3425.2	2084.99	2084.99
SVM (Multiquadric)	3892.3	2101.27	2101.27
PNN	5386.13	3201.77	3201.77

The validation time calculated after ranking by Bhattacharya

## 4 Conclusion

In this article, a novel algorithm has been presented for detection of CHF. In this algorithm various ranking methods were combined with KPCA and 1 – NLPELM classifier. The analysis results show that the detection based on top ten ranked five-level LRMSF attributes extracted from decomposed HRV signal achieved excellent accuracy as compared to existing decomposed techniques. The investigation of CHF revealed that attributes at 4<sup>th</sup> and 5<sup>th</sup> level of HRV decomposition by MRWP have lowest p – value(< 0.001). Since  $p < 0.001$  attributes have greatest discernment capability so 4th and 5th level of MRWP attributes are considered much suitable than 1st, 2nd and 3rd level MRWP attributes. In addition, the proposed method has achieved very good generalization performance and less execution time as compared to 1 – NLPELM, KPCA + PNN, KPCA + SVM, PNN and SVM. It indicates that HRV signal decomposed by MRWP spectral method are most suitable for the clinical system design and setting for detection of cardiac heart disease and HRV analysis. This proposed method can be employed for detection and classification of cardiac diseases like coronary artery diseases, Acute Infection, arrhythmia disease detection and Myocardial Infarction.

## Declarations

**Conflict of interest** All the authors declared that they have no conflict of interest.

## References

1. Algra A, Tijssen JGP, Roelandt JRTC, Pool J, Lubsen J (1991) QTc prolongation measured by standard 12-lead electrocardiography is an independent risk factor for sudden death due to cardiac arrest. *Circulation* 83(6):1888–1894. <https://doi.org/10.1161/01.CIR.83.6.1888>

2. Aran O, Akarun L (2010) A multi-class classification strategy for Fisher scores: application to signer independent sign language recognition. *Pattern Recognit* 43(5):1776–1788. <https://doi.org/10.1016/j.patcog.2009.12.002>
3. Asl BM, Setarehdan SK, Mohebbi M (2008) Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal. *Artif Intell Med* 44(1):51–64. <https://doi.org/10.1016/j.artmed.2008.04.007>
4. Babaoglu I, Findik O, Ulker E (2010) A comparison of feature selection models utilizing binary particle swarm optimization and genetic algorithm in determining coronary artery disease using support vector machine. *Expert Syst Appl* 37(4):3177–3183. <https://doi.org/10.1016/j.eswa.2009.09.064>
5. Bartlett PL (1998) The sample complexity of pattern classification with neural networks: the size of the weights is more important than the size of the network. *IEEE Trans Inf Theory* 44(2):525–536. <https://doi.org/10.1109/18.661502>
6. Baudat G, Anouar F (2000) Generalized discriminant analysis using a Kernel approach. *Neural Comput* 12(10):2385–2404. <https://doi.org/10.1162/089976600300014980>
7. Bin-Huang G, Chen L, Siew CK (2006) Universal approximation using incremental constructive feed-forward networks with random hidden nodes. *IEEE Trans Neural Netw* 17(4):879–892. <https://doi.org/10.1109/TNN.2006.875977>
8. Bin-Huang G, Wang DH, Lan Y (2011) Extreme learning machines: a survey. *Int J Mach Learn Cybern* 2(2):107–122. <https://doi.org/10.1007/s13042-011-0019-y>
9. Braunschweig F et al (2010) Management of patients receiving implantable cardiac defibrillator shocks. *Europace* 12(12):1673–1690. <https://doi.org/10.1093/europace/euq316>
10. Cao J, Lin Z, Bin Huang G, Liu N (2012) Voting based extreme learning machine. *Inf Sci (NY)* 185(1):66–77. <https://doi.org/10.1016/j.ins.2011.09.015>
11. Castiglioni P, Lazzeroni D, Coruzzi P, Faini A (2018) Multifractal-multiscale analysis of cardiovascular signals: a DFA-based characterization of blood pressure and heart-rate complexity by gender. *Complexity* 2018(1):1–14. <https://doi.org/10.1155/2018/4801924>
12. Ding S, Zhao H, Zhang Y, Xu X, Nie R (2015) Extreme learning machine: algorithm, theory and applications. *Artif Intell Rev* 44(1):103–115. <https://doi.org/10.1007/s10462-013-9405-z>
13. Dua S, Du X, Vinitha-Sree S, Thajudin-Ahamed VI (2012) Novel classification of coronary artery disease using heart rate variability analysis. *J Mech Med Biol* 12(4):1240017. <https://doi.org/10.1142/S0219519412400179>
14. Ebrahimzadeh E, Pooyan M (2011) Early detection of sudden cardiac death by using classical linear techniques and time-frequency methods on electrocardiogram signals. *J Biomed Sci Eng* 04(11):699–706. <https://doi.org/10.4236/jbise.2011.411087>
15. Gao Y, Xie Z, Yu X (2020) A hybrid algorithm for integrated scheduling problem of complex products with tree structure. *Multimed Tools Appl* 79(43–44):32285–32304. <https://doi.org/10.1007/s11042-020-09477-2>
16. Giri D et al (2013) Automated diagnosis of coronary artery disease affected patients using LDA, PCA, ICA and discrete wavelet transform. *Knowl-Based Syst* 37:274–282. <https://doi.org/10.1016/j.knosys.2012.08.011>
17. Hallstrom AP, Stein PK, Schneider R, Hodges M, Schmidt G, Ulm K (2005) Characteristics of heart beat intervals and prediction of death. *Int J Cardiol* 100(1):37–45. <https://doi.org/10.1016/j.ijcard.2004.05.047>
18. Hemamalini B, Nagarajan V (2020) Wavelet transform and pixel strength-based robust watermarking using dragonfly optimization. *Multimed Tools Appl* 79(13–14):8727–8746. <https://doi.org/10.1007/s11042-018-6096-0>
19. Huang G-B et al (2006) Extreme learning machine: theory and applications. *Neurocomputing* 70(1–3):489–501. <https://doi.org/10.1016/j.neucom.2005.12.126>
20. Huang G-B, Ding X, Zhou H (2010) Optimization method based extreme learning machine for classification. *Neurocomputing* 74(1–3):155–163. <https://doi.org/10.1016/j.neucom.2010.02.019>
21. Huang GB, Zhou H, Ding X, Zhang R (2011) Extreme learning machine for regression and multiclass classification. *IEEE Trans Syst Man Cybern Part B* 4(2):513–529. <https://doi.org/10.1109/tsmcb.2011.2168604>
22. Huikuri HV et al (2003) Prediction of sudden cardiac death after myocardial infarction in the beta-blocking era. *J Am Coll Cardiol* 42(4):652–658. [https://doi.org/10.1016/S0735-1097\(03\)00783-6](https://doi.org/10.1016/S0735-1097(03)00783-6)
23. Kampouraki A, Manis G, Nikou C (2009) Heartbeat time series classification with support vector machines. *IEEE Trans Inf Technol Biomed* 13(4):512–518. <https://doi.org/10.1109/TITB.2008.2003323>
24. Karimi M (2006) Noninvasive detection and classification of coronary artery occlusions using wavelet analysis of heart sounds with neural networks. *IET* 20:117–120. <https://doi.org/10.1049/ic:20050342>

25. Khushaba RN, Kodagoda S, Lal S, Dissanayake G (2011) Driver drowsiness classification using fuzzy wavelet-packet-based feature-extraction algorithm. *IEEE Trans Biomed Eng* 58(1):121–131. <https://doi.org/10.1109/TBME.2010.2077291>
26. Lal H et al (2020) Machine learning based congestive heart failure detection using feature importance ranking of multimodal features. *Math Biosci Eng* 18(1):69–91. <https://doi.org/10.3934/mbe.2021004>
27. Li S, Wang H, Wang S, Zhang S (2020) Life detection and non-contact respiratory rate measurement in cluttered environments. *Multimed Tools Appl* 79(43–44):32065–32077. <https://doi.org/10.1007/s11042-020-09510-4>
28. Li K, Daniels J, Liu C, Herrero P, Georgiou P (2020) Convolutional recurrent neural networks for glucose prediction. *IEEE J Biomed Health Inform* 24(2):603–613. <https://doi.org/10.1109/JBHI.2019.2908488>
29. Liping X et al (2020) Computational diagnostic techniques for electrocardiogram signal analysis. *Sensors* 20:6318. <https://doi.org/10.3390/s20216318>
30. Martinez AM, Kak AC (2001) PCA versus LDA. *IEEE Trans Pattern Anal Mach Intell* 23(2):228–233. <https://doi.org/10.1109/34.908974>
31. Patel DR (2011) Sudden cardiac death in young athletes. *Asian J Sports Med* 2(2):120–122. <https://doi.org/10.5812/asjms.34778>
32. Poddar MG, Kumar V, Sharma YP (2015) Automated diagnosis of coronary artery diseased patients by heart rate variability analysis using linear and non-linear methods. *J Med Eng Technol* 39(6):331–341. <https://doi.org/10.3109/03091902.2015.1063721>
33. Rajendra AU, Kannathal N, Krishnan SM (2004) Comprehensive analysis of cardiac health using heart rate signals. *Physiol Meas* 25(5):1139–1151
34. Richman JS, Moorman JR (2000) Physiological time-series analysis using approximate entropy and sample entropy. *Am J Physiol Heart Circ Physiol* 278(6):H2039–H2049. <https://doi.org/10.1103/physrva.29.975>
35. Singh RS, Saini BS, Sunkaria RK (2018) Classification of cardiac heart disease using reduced chaos features and 1-NORM linear programming extreme learning machine. *Int J Multiscale Comput Eng* 16(5):465–486. <https://doi.org/10.1615/IntJMultCompEng.2018026587>
36. Singh RS, Saini BS, Sunkaria RK (2018) Assessment of cardiac heart failure and cardiac artery disease by the higher order spectra. *Biomed Eng Appl Basis Commun* 30(2):1–9
37. Singh RS, Saini BS, Sunkaria RK (2018) Times varying spectral coherence investigation of cardiovascular signals based on energy concentration in healthy young and elderly subjects by the adaptive continuous morlet wavelet transform. *IRBM* 39(1):54–68. <https://doi.org/10.1016/j.irbm.2017.12.004>
38. Tamil EBM, Kamarudin NH, Salleh R, Tamil AM (2008) A review on feature extraction & classification techniques for biosignal processing (Part I: Electrocardiogram). *IFMBE Proc* 21(1):107–112. <https://doi.org/10.1007/978-3-540-69139-6-31>
39. Tang J, Alelyani S, Liu H (2014) Feature selection for classification: a review. *Data Classif Algorithms Appl* 2014:37
40. Vollmer M (2016) A robust, simple and reliable measure of heart rate variability using relative RR intervals. *Comput Cardiol* (2010) 42(6):609–612. <https://doi.org/10.1109/CIC.2015.7410984>
41. Wallen H, Linder R (2013) Heart disease. Essential guide to blood coagulation, 2nd edn. Wiley, Hoboken, pp 97–104

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Authors and Affiliations

Demissie J. Gelmecha<sup>1</sup> · Ram S. Singh<sup>1</sup> · Devendra K. Sinha<sup>2</sup> · Dereje Tekilu<sup>1</sup>

Demissie J. Gelmecha  
gelmechad@gmail.com

Devendra K. Sinha  
ds3621781@gmail.com

Dereje Tekilu  
dereteklu@gmail.com

<sup>1</sup> Electronics & Communication Eng. Department (SIG), School of Electrical Engineering & Computing, Adama Science & Technology University, 1888 Adama, Ethiopia

<sup>2</sup> Mechanical Design & Manufacturing Eng. Department, Center of Excellence in Advanced Manufacturing Engg., School of Mechanical, Chemical & Material Engineering, Adama Science & Technology University, Adama, Ethiopia