Automated Coronary Artery and Congestive Heart Failure Disease Detection Using Nonlinear Features and On-Line Sequential Extreme Learning Machine

By

Dr. Demissie Jabir

Dr. Rama sewakSignh

A Final Research Report Submitted to Adama Science and Technology University

Adama, Ethiopia
May 2023

ABSTRACT

Cardiovascular diseases (CVDs) are major reason of mortality in the world population, and the numeral of cases is up surging every year. The mortality rate due to coronary artery disease (CAD) and congestive heart failure (CHF) is higher than any other type of CVDs. The vast majority of cardiovascular disease fatalities occur in middle and low-income nations, including Ethiopia. According to the latest WHO data published in 2020, coronary heart disease deaths in Ethiopia reached 47,712 or 7.81% of total deaths. The age adjusted death rate is 112.44 per 100,000 of population and it ranks Ethiopia @112 in the world. Arrhythmias are a marker of the heart's abnormal activity and are linked to an increased risk of sudden cardiac death (SCD), one of the most crucial tests in cutting-edge cardiology. Many of the symptoms of CVDs can be relieved or avoided if they are detected and treated early. Consequently, the improvement of investigative procedures might improve the health of many individuals. As a result, the goal of this research is to automate the identification of CVDs. In this proposal, automated algorithms have been developed for the categorization of heart disease to identify cardiac arrhythmias using Heart Rate Variability (HRV) standard database and self recorded data.

For investigation and detection of CHF and CAD, twelve non-linear attributes like correlation dimension (CD), detrended fluctuation analysis (DFA) variants DFA-α1 and DFA-α2, Bubble Entropy (BBEn), sample entropy (SampEn), dispersion entropy (DISEn), Lempel–Ziv complexity (LZ), sinai entropy (SIEn), improved multiscale permutation entropy (IMPE), hurst exponent (HE), permutation entropy (PE), approximate entropy (ApEn) and standard deviation (SD1/SD2) were retrieved from HRV signal. A feature reduction technique known as generalized discriminat analysis (GDA) has been used to reduce the dimension of these attributes. The reduced attributes have been normalized between 1 and -1 and then fed to online sequential extreme learning machine (OSELM) for classification and detection of CHF and CAD. For analysis of these cardiac diseases, the HRV signal was obtained from self recoded and standard database of MIT BIH of participants with normal sinus rhythm (NSR). The St. Petersburg Institute of Cardiological Technics data source provided the CHF and CAD databases used in this study. The numerical results have shown that GDA with Gaussian kernel function

and OSELM with sine activation function achieved accuracy (AC) of 99.34% and sensitivity (SE) of 99.32% for NSR-CAD group, and AC and SE of 100% were achieved for NSR-CHF group. In addition, it has been observed that the algorithm's classification performance was improved with fewer blocks of data, and its generalization performance was excellent for detecting CHF and CAD.

Additionally, the 1-norm extreme learning machine (1-NELM) binary classifier suggested in the preceding part of the study was expanded for various data groups in this section, and the findings were elaborated utilizing ranking methods such as Fisher Wilcoxon, Entropy, Bhattacharya and receiver operating characteristic (ROC) and reduction strategies such as GDA. For this study, the proposed method was validated using self recorded and standard database of MIT-BIH and SPICT. The considered dataset were grouped as CAD-CHF, YOUNG-CAD, YOUNG-CHF, ELDERLY-CAD and ELDERLY-CHF subjects. To diagnose disorders of the heart, such as CHF and CAD, linear discriminant analysis (LDA) and GDA are used as feature reduction techniques in conjunction with the 1-NELM binary classifier. The activation functions used with the classifier are Sigmoid, Hardlim, and RBF whereas LDA and GDA are used with the kernel functions such as Gaussian and RBF. Various nonlinear features were generated from HRV data and utilized in training and validation of proposed algorithm. The analyses were carried out numerically through the combination of database sets YNG-ELY, YNG-CAD and ELY-CAD subjects. The numerical results have shown that ROC with GDA and 1-NELM approach achieved an accuracy of 99.76±0.14, 99.87±0.12 and 100±0 for YNG-CAD, YNG-ELY and ELY-CAD groups respectively. The Fisher with GDA and 1-NELM; and Bhattacharya with GDA and 1-NELM approach achieved an accuracy of 100±0 for all considered datasets. The proposed method also achieved very good generalization performance with the smallest 1-Norm Root Mean Square Error (1-NRMSE) and less execution validation time as compared to support vector machine (SVM) and probabilistic neural network (PNN).

Finally, this work presented an 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. The ten ranked attributes were then mapped to one new feature by KPCA and fed to 1-NLPELM. The simulation results demonstrated that Bhatacharya+KPCA with 1-NLPELM approach achieved an accuracy of 98.44±1.4%, 99.13±1.85% for NSR-CHF and ELY-CHF respectively. Out of all ranking methods, Bhatacharya 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, PNN and SVM.

ACKNOWLEDGEMENT

We would like to express my greatest appreciation to Electronics and Communication Engineering Department, and School of Electrical Engineering and Computing, Adama Science and Technology University, Adama, Ethiopia for giving opportunity and encouragement to propose the research proposal.

We would like to gratefully acknowledge Dr. Tekilu, Dean, SoEEC, ASTU, Adama, Ethiopia for his continuous encouragement throughout the research work. As well-wisher, his insight, observations and suggestions helped us to establish the overall direction of the research and contributed immensely to the success of the work.

We acknowledge Mr. Eshetu Tessema , HOD of ECE Department, and Mr. Tadesse Hailu Ayane, Dean for research & Technology Transfer, Adama Science and Technology University, Adama, Ethiopia for their support and suggestions during the process of the project.

We must acknowledge the hospitals Adama general hospital and medical college, Adama, St. Paulo Hospital, Addis Ababa, American medical center, Addis Ababa for their support in providing data for research.

We acknowledge the academic resource that we received from Adama Science and Technology University, Adama, Ethiopia giving us a comfortable and active environment for pursuing our research work.

Finally, we would like to thank our staff members of the Electronics and Communication Engineering Department, ASTU, Adama for their support and encouragement throughout the research work.

TABLE OF CONTENTS

List of Figures

List of Acronyms

Acronyms	Full Form
AC	Accuracy
AF	Atrial fibrillation
AFL	Atrial flutter
ANN	Artificial neural network
ANS	Autonomic nervous system
AP	Action potential
ApEn	Approximate entropy
APIV	Arterial pressure interval variability
AR	Autoregressive
AUC	Area under ROC curve
AV	Atrio-ventricular
a.u.	Arbitrary unit
BPN	Back propagation neural
bpm	Beats per minute
BIDMC	Beth Israel Deaconess Medical Center
CAD	Coronary artery disease
CBC	Cumulative bi-correlation
CD	Correlation dimension
PE	Permutation entropy
CHF	Congestive heart failure
CVDs	Cardiovascular diseases
DE-ELM	Differential evolution extreme learning machine
DFA	Detrended fluctuation analysis
DFT	Discrete Fourier transform
DTCWT	Dual tree complex wavelet transform
DWT	Discrete wavelet transform

ECG	Electrocardiogram
ELM	Extreme learning machine
1-NELM	1-Norm extreme learning machine
1-NLPELM	1-Norm linear programming extreme learning machine
ELY	Healthy elderly
FZE	Fuzzy entropy
GDA	General discriminant analysis
HE	Hurst exponent
HEG	Healthy elderly group
HF	High frequency
HFp	High Frequency power
HRV	Heart rate variability
HYG	Healthy young group
ICA	Independent component analysis
IMPE	Improved multiscale permutation entropy
K-NNE	Kraskov nearest neighbour entropy
KPCA	Kernel principal component analysis
LBBB	Left bundle branch block
LDA	Linear discriminant analysis
LE	Lyapunov exponent
LF	Low frequency
LF_P	Low frequency power
LRMS	Log root mean square
MF	Mid-frequency
MLP	Multilayer perceptron
MIT BIH	Massachusetts Institute of Technology, Beth Israel Hospital
MSE	Mean of squared Errors
MSWP	Multiscale wavelet packet
N.U.	Normalized unit
N-N	Normal to Normal Interval
nHF	Nomalized power in high frequency

nLF	Normalized power in low frequency
NSR	Normal sinus rhythm
OS-ELM	Online sequential extreme learning machine
PAC	Premature atrial contractions
PCA	Principal component analysis
PE	Permutation entropy
PNN	probabilistic neural network
RBBB	Right bundle branch block
RBF	Radial basis function
ROC	Receiver operating characteristics
RSA	Respiratory sinus arrhythmia
REP	Respiratory
REPU	Respiratory Unit
SA	Sino-atrial
SABPV	Systolic arterial blood pressure variability
SBP	Systolic blood pressure
SCD	Sudden cardiac death
SE	Sensitivity
Self_NSR	Self recorded data as normal sinus rhythm of healthy men
SLFNS	Single hidden layer feedforward neural networks
SNR	Signal to noise ratio
SP	Specificity
STFT	Short time Fourier transform
SVM	Support vector machine
SVT	Supraventricular tachycardia
SVTG	Supraventricular tachycardia group
T-F	Time-frequency
TV	Time-varying
VF	Ventricular fibrillation
VLF	Very low frequencies
VT	Ventricular tachycardia

WHO	World health organization
WN	Wavelet network
WPT	Wavelet packet transform
WT	Wavelet transform
YNG	Healthy young
γ	Adaptive parameter
δ	Delta rule
$\emptyset(.)$	Activation function

List of Tables

CHAPTER-1

1. INTRODUCTION

1.1 Motivation

According to the report of world health organization (WHO) fact sheet[1], cardiovascular diseases (CVDs) like congestive heart failure (CHF) and coronary artery disease (CAD) are the foremost cause of global mortality as compared to other causes. An expected 17.9 million people die every year from CVDs, indicating approximately 32% of all global deaths [2]. Coronary heart disease, ischemic heart disease and stroke are the leading accountable reasons for greater than 80% of CVD deaths. In 2015, patients suffered with CHF and CAD was found to be 500 million. This figure is expected to be doubled by 2025. Over 75% of CVD deaths befall in middle and low income countries among which Ethiopia is also one [3]. According to the latest WHO data published in 2020, coronary heart disease deaths in Ethiopia reached 47,712 or 7.81% of total deaths. The age adjusted Death Rate is 112.44 per 100,000 of population and it ranks Ethiopia @112 in the world. As per the WHO statistical details of heart disease and stroke deaths released in April 2018, it has been able to track the occurrence of CAD and CHF worldwide (Table: 1.1). This figure is predicted to be much greater than the assessments based on statistical data which have taken into account only death certificate reports and do not consider patients who die from CAD and CHF complications in rural area [4].

Coronary angiography,Computed tomography (CT), Electrocardiogram (ECG) stress test, Stress echocardiography and Myocardial Perfusion Imaging are common methods for diagnosis of CADs. Coronary angiography With CT is under heavy development and is expected to play a significant role in CAD diagnosis. Although being invasive it has a relatively high precision, by analysis it was found that the sensitivity (SE) of 96.12 percent and the specificity (SP) of 86.98 percent [5]. The stress test of the ECG is relatively simple assessment for CAD and CHF.

It is a simple and inexpensive method as well as the results of the test specifically relates to heart functionality. This form, therefore, faced another drawback that it has a low precision as compared to 68 percent of SE and 77 percent of SP [6]. The Echoes stress test is an extension of an ECG test phase by evaluating changes in stressed ventricles activity, including exercise [7]. The benefit has increased precision but its disadvantages are due to increased variation in the technicians as well as the system needs highly skilled personnel.

Table: 1.1 The statistical details of Cardiac heart diseases like CAD, CHF and Heart stroke deaths released by WHO in 2020.

Country	Deaths	Percentage (%)	Rate/1 lakh	World Rank/192
		SOUTH AMERICA		
Guyana	924	16.41	151.35	50
Venezuela	21,165	17.66	107.26	102
Basil	133992	13.76	81.18	134
		NORTH AMERICA		
Honduras	6,168	16.01	152.44	48
Nicaragua	3,073	13.93	95.28	119
United states	445,864	21.42	80.48	135
		AUSTRALIYA		
Australia	24,905	21.91	60.34	162
		AFRICA		
Sudan	39,326	10.67	212	24
Ethiopia	47,712	7.81	112.44	112
Morocco	33760	24.40	168	36
Nigeria	71732	4.20	121.6	80
Algeria	14708	10.63	75.17	145
		ASIA		
Turkmenstan	11,665	30.86	405.13	1
Malaysia	22,701	22.18	138.75	57
Korea(north)	27,569	16.35	113.17	91
China	1,040692	11.71	79.72	137
		MIDDLE EAST		
Iraq	26,676	14.12	214.13	22
Saudi Arabia	20,877	23.98	180.58	32
Arba Emirates	1,161	19.74	94.48	120
		EUROPE		
Ukraine	338,108	49.36	399.79	2
Romania	56,727	26.16	155.04	45
Poland	79,036	26.93	122.34	78
Switzerland	9,804	19.91	52.18	175

Despite recent medical and treatment advancements, CVDs remains the foremost common risk factor for mortality accounting for nearly one third of deaths worldwide. Robust and efficient diagnosis is essential for improving CVD detection and saving the life from mortality. Signal processing of heart rate variability (HRV) signal (R-R interval consecutive of PQRS morphological of ECG signal) plays a key role in decision making during diagnosis. Present techniques of HRV analysis rely mainly on qualitative visual interpretation of ECG signal and basic quantitative measurements of cardiovascular function [8]. To maximize the predictive value of cardiac diseases, more developed HRV signal processing methods are required to allow

better quantification of phenotypes in the ECG morphology and remove the noise for accurate detection.

The big data analysis and the availability of high cloud computing in recent years have guided rapid innovation of artificial intelligence (AI) technology in cardiac disease prediction and detection [9]. Machine learning (ML) strategies based on features extracted from HRV, treatment rely on optimization techniques / modeling techniques which learn from the past experiment was conducted to find by recognizing hidden and multiple concerns of HRV signal. Many authors already demonstrated the ECG based cardiovascular diseases diagnosis like detection and prediction of CAD and CHF with ML. The ML detection and prediction efficiency based on features of HRV signal has the greatest potential to reduce the burden of heart disease by enabling faster and more efficient diagnostic decision taking [10].

A new fast learning neural classifier for Single-hidden-layer feed forward networks (SLFNs) modeled Extreme Learning Machine (ELM) [11], [12] has recently been investigated to optimize the effectiveness of SLFN. Unlike neural network machine learning (like Back-propagation (BP) models), which may experience difficulties in automatically tuning process parameter (learning speeds, having to learn epochs, etc.) and/or local minima, ELM is entirely systematically deployed without recursive tuning and in principle, so users do not need to interfere. In fact, ELM's learning speed is incredibly high as compared to other conventional approaches. In the ELM technique, the learning specifications of hidden neurons include input synaptic weights which can be individually allocated randomly as well as the network solved analytically can be mathematically calculated via simple and standardized reverse process. In the training phase the ELM has the ability of learning fast without time-consuming process of learning with a corrected nonlinear activation function [13]. In addition, the ELM methodology can provide a good result in generalization.

Online version of the classification problem is needed due to the rapidly growing need of streaming data [14],[15] . For this, Online Sequential Extreme Learning Machine (OSELM) is modified version of ELM will employ to binary classifies the cardiac heart diseases (either healthy or CAD, CHF) and GDA for features dimension reduction which will improve the classification performance like accuracy,sensitivity, specificity and decrease the detection time .

1.2 Cardiovascular Diseases

Cardiovascular diseases (CVD) include a group of conditions related to the heart, which consists of:

i) Congestive Heart Failure (CHF): The body receives the blood pumped from the heart. If the heart cannot supply the blood vessels with blood to circulate in the body, it can lead to heart failure if the heart muscles are weakened due to a heart attack or cardiomyopathy (heart muscle disease).

ii) Coronary Artery Disease (CAD): Whenever the main arteries supplying the heart become weak and diseased, it results in coronary artery disease. Coronary vein infection is caused by cholesterol-containing deposits (plaques) in the coronary arteries and their subsequent exacerbation. In the heart, the coronary veins provide blood, oxygen, and other support materials. Plaque buildup in these veins can restrict blood flow to the center of the body. Angina (chest discomfort), exhaustion, and other symptoms of coronary artery disease might result from a reduced blood supply. Heart failure might be caused by full obstruction.

iii) Rheumatic Heart Disease (RHD): This form of heart disease occurs in infancy as a bacterial infection caused by streptococcal bacteria that affects the heart valves and joints during rheumatic fever. Bacterial diseases damage the heart's inner tissues, including the valves (endocarditis) and the external tissues surrounding the heart (pericarditis).

iv) Arrhythmia: Cardiac arrhythmia is related to abnormal heart rate or rhythm [16].

v) Coronary Heart Disease (CHD): The blood arteries are called coronary blood vessels. By developing plaques and plugs, atherosclerosis causes the narrowing and hardening of the heart's arteries, which leads to a loss of oxygenated blood (ischemia) to the heart, leading to complications such as angina and dyspnea (shortness of breath).

vi) Congestive Heart Disease: Congenital heart illness is caused by any malformation or irregular formation of the heart's structure at birth. Pregnant women who take medications, alcohol or a bad diet that causes vital nutrient deficiency may cause a baby to have such a disease. This cardiac condition contains irregular valves, heart holes, or abnormal heart chambers that often lead to death unless quickly repaired by surgery.

1.3 Research Gaps

- As the studied in the frequency domain by various method using the Fourier transform like non-parametric method, Autoregressive modelling (choosing the order) and parametric method. These estimators are limited by poor spectrum resolution in VLF, LF and HF frequency range, due to resampling and model misspecification (choosing order). These methods are only suitable for stationary and quasi stationary signals.
- A more appropriate approach for nonstationary processes is the time-frequency distribution, The most studied is the STFT and Wigner distribution and SPWV provide information about Time and frequency spectrum simultaneously but these methods also have poor resolution of spectrum and not reduced the noise content of nonstationary signal (As nonstationary signal content wideband noise due to instantaneous frequency change of signal).
- Comparing with STFT, wavelets can be utilized to analyse multi-scale signals through scaling and translation, and extract time-frequency characteristics of the signals effectively. Therefore, wavelets are more suitable than STFT for analysing non- stationary signals. Wavelets being non-adaptive, however, have its own disadvantage that their analysis results depend on the choice of the wavelet base function. This may lead to a subjective and a prior assumption on the characteristics of the signal. As a result, only the signal characteristics that correlate well with the coefficient of the mother wavelet function have a chance to produce high value coefficients. Any other characteristics will be masked or completely ignored.
- Non linearity of cardiac rhythm losses due to scaling and re-sampling.
- Selection of efficient classifier and reduction of dimensionality of the features extracted from HRV signals.

1.4 Objective

1.4.1 General Objective**:** Automated coronary artery disease and congestive heart failure detection using nonlinear features and Extreme Learning Machine.

1.4.2 Specific objectives:

- ☐ To reduce the artifact and noise from ECG signal.
- ☐ To detect the R-Peaks from Morphological waveform of ECG signal and extract HRV signal from ECG signal.
- ☐ To extract the nonlinear features from HRV signal.

- To develop an automated model for online detection and classification of CAD and CHF patients.
- To Compare the performance of proposed model with ML like support vector machine (SVM), PNN and Extreme learning machine (ELM).

1.5 Framework of Proposed Model

The R-R interval (HRV) data used in this thesis work has been obtained from the ECG signals provided by the normal sinus rhythm (NSR), coronary artery disease (CAD), congestive heart failure (CHF), healthy young, healthy elderly and self recorded data of healthy young subjects. The data of NSR has been obtained from Massachusetts Institute of Technology, Beth Israel Hospital (MIT BIH) database. The databases of normal sinus rhythm (NSR) consisted of eighteen long-term ECG recordings of subjects; they include five men, aged 26 to 45, and thirteen women, aged 20 to 50 years, each sampled at 128 Hz. The CAD data has been obtained from St. Petersburg Institute of Cardiological Technics consists of 75 annotated recordings extracted from 32 Holter records. Each record was 30 minutes long and contains standard ECG leads. Only thirteen subjects (9 men and 4 women, aged 18-80; mean age: 58) suffered from CAD, each sampled at 257 Hz [17]. For analysis and detection of arrhythmic subjects, the database was obtained from MIT-BIH arrhythmia [18] and for analysis of cardiovascular signals of healthy young and elderly subjects, the database was obtained from the Fantasia Database [19]. The Beth Israel Deaconess Medical Center (BIDMC) congestive heart failure (CHF) database of long-term ECG recordings from fifteen subjects (4 women, aged 54 to 63 and 11 men, and aged 22 to 71) with severe CHF, each sampled at 250 samples per second.

The self-recorded data base consists of 13 healthy men subjects, aged between 23 and 32 years. The ECG was continuously recorded for 2 minutes of each subject in the relaxed supine and normal sinus rhythm position in a room free from any kind of disturbance with controlled temperature (22–35°C). The recording of ECG using electrode placement method was done at the sampling rate of 1000 Hz with Avidhrt Sense portable device**.** The flow chart of research structure for the investigation of cardiovascular signals is shown in Figure 1.1**.**

1.6 Validation of proposed method and Collection of database:

- First of all, for real detection, self-recoded Electrocardiogram (ECG) signal have collected using an Avidhrt Sense portable device with concern of cardiac expert doctors from different hospitals like Adam, Addis Ababa hospitals.

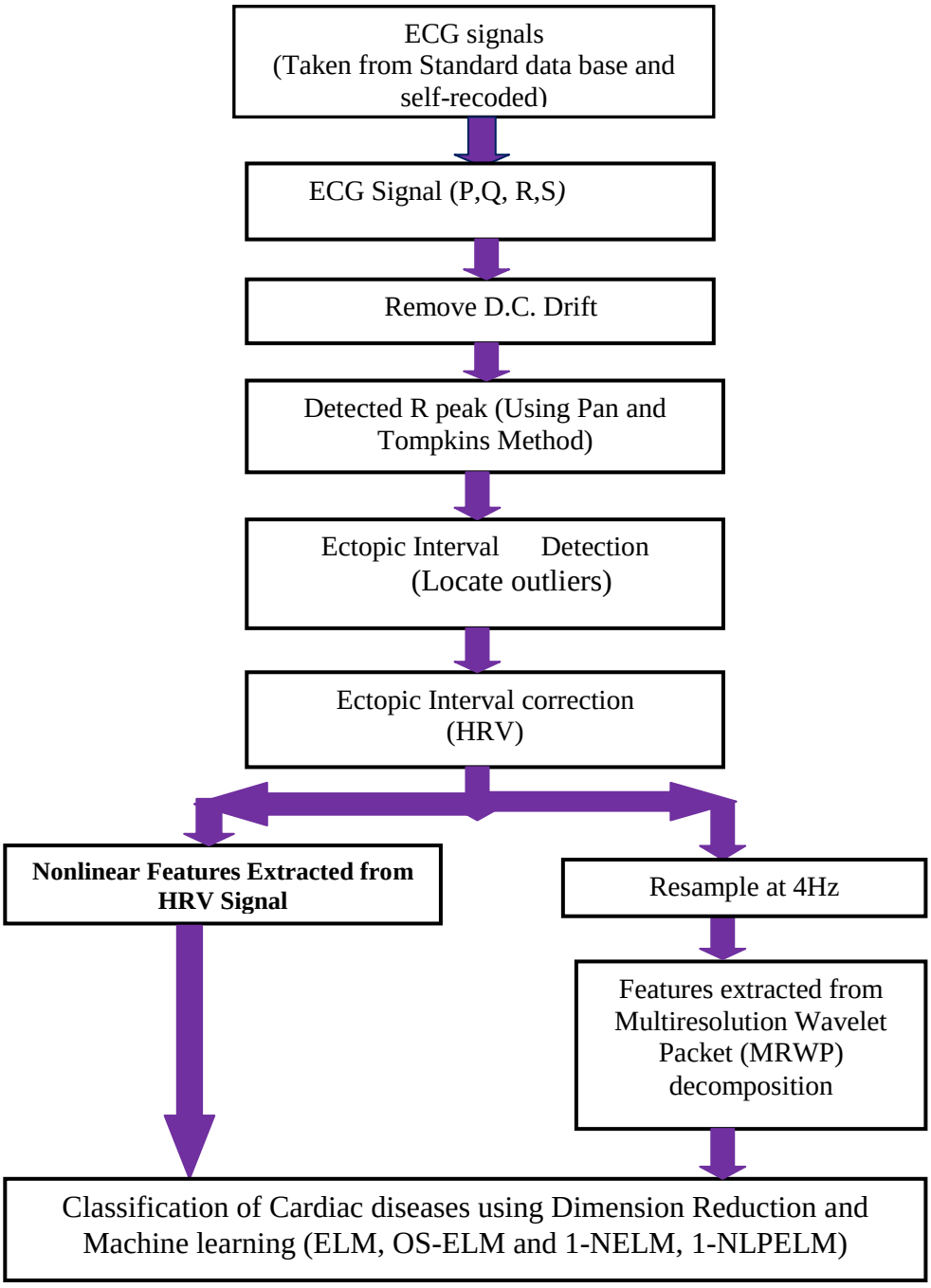


Figure 1.1: Framework of analysis of HRV signal

- Finally, the proposed model has been validated on publicly available datasets (Standard database). This data can be found from website:

https://physionet.org/physiobank/database/sddb/;https://physionet.org/physiobank/database/nsrdb https://physionet.org/physiobank/database/fantasia/.

- MATLAB codes have be used for simulation and emulation of data analysis.

1.7 Evaluation.

In this phase the individual responses has to be evaluated to test the proposed model.

- The analyzed results of the proposed model have been compared with conventional ELM model for technology acceptance.
- The results with publicly available datasets set and collected dataset have compared using the proposed model for addressing the classification problem of CAD and CHF.
- It will have attitude of ease of use of the model by the Ethiopian medical industries and medical practitioners.
- The ease of use refers to the degree to which medical practitioners will believe that the use of online detection and classification will classify the coronary artery diseases and preventive actions can be taken to save the Human Life.

1.8 Significance

- Propose model can be used online for treatment, monitoring, diagnosis of cardiac diseases and distinguish the CHF and CAD affected patients.
- The proposed algorithm can be used for self heart rate monitoring.
- The Physician can suggest the cardiac health and medicine to subjects on the basis of records of analyze online data. In this way, these allow for remote patient monitoring, which is of special necessity to the elderly, who are more prone to CAD and CHF diseases and cannot always visit a healthcare facility.
- The proposed model can provide early warning to patients at risk of SCD
- This model can be useful for those subjects suffered with CAD and CHF who are outside hospitals (rural areas where there are no expert doctors are available).

1.9 Project Report Organization

Chapter 1: describes a brief background of statistical details of Cardiac heart diseases like CAD, CHF and Heart stroke deaths released by WHO and contribution, objective, significance of the study and framework of analysis of Cardiac heart diseases followed by research gaps.

Chapter 2 presents critical literature review on research work reported on different detection and classification techniques of Cardiac heart diseases,

Chapter 3: presents details of extraction of R-Peaks and HRV signal from ECG signal, removal of noise and framework of proposed model with hardware. It also includes fundamental of extreme learning machine, OS-ELM and 1-NLPELmodels and nine different chaos attributes like Correlation Dimension (CD), Detrended Fluctuation Analysis (DFA), Approximate Entropy (ApEn), the results of the Poincare plot as SD1/SD2 ratio, Hurst Exponent(HE), Permutation Entropy (PE), Improved Multiscale Permutation Entropy (IMPE) and cumulative bi-correlation (CBC). This chapter also focuses on detection and classification of CAD and CHF cardiac diseases. This chapter focuses on detection and classification of CAD using eleven different nonlinear and chaos features were extracted from pre-processed and fragmented HRV database. In addition, this chapter also describes for binary classification of CHF disease based on features extracted decomposition from HRV signal using Multiresolution Wavelet Packet (MRWP) method. The HRV signal has been decomposed up to 5-level using MRWP decomposition method using Haar mother wavelet.

Chapter 4 presents the conclusion for the research work followed by reference. It also focuses on contributions of the project and future scope.

Chapter 2

LITERATURE REVIEW

2.1 Physiology of Human Heart

The human heart is a muscular organ and works as electromechanical pulsatile pump. From the anatomic perspective, as shown in Figure 2.1, there are two separated pumps: one at the right that inflates blood through the lungs organ, and one at the left that inflates blood through the exterior organs. The left pump composed of a left atrium and a left ventricle and the right pump composed of a right atrium and right ventricle. The ventricle is bigger heavy-walled chambers that do most of the work. The atria collect blood from the lungs and venous system and then contract and expel the blood into the ventricle. The ventricle then pumps the blood all through the body or into the lungs [16]. There are four valves namely Tricuspid, Mitral, Aortic and Pulmonary valve contained by heart to force the direction of blood. Tricuspid valve links the right ventricle and right atrium. Mitral valve links the left atrium to the left ventricle and has an elliptical orifice. Aortic valve opens between the aorta left and the ventricle [20]. The cardiac myocytes have an inimitable capability of automatic electrical impulse generation, which consequence is automatic rhythmicity. An electrical impulse is responsible of the mechanical stimulation of the muscle that provides periodicity of electromechanical pumping of heart. Each cycle is originated by spontaneous generation of an action potential (AP) in the Sino-atrial (SA) node, as it has the fastest electrical impulse generation capability and therefore initiates the heart. The AP is of long duration about 300 ms, which is longer in comparison with the AP of the rest of the cells in the body. The SA node is located by the right atrium nearby the superior vena cava. The electrical impulse, or AP, trekking through both atria and reaching the Atrio-ventricular (AV) bundle, where is delay by around 0.1s. The AV node transmits the impulse with this delay and spreads over and done with the ventricular myocardium through the AV bundle of His and Purkinje fibres. The delay time permits the atrium to pump blood into the ventricles. Subsequently, the ventricles are filled with blood and ready to be stimulated. This coordination spreads the electrical impulse from the AV node to the entire ventricular muscle very fast, allowing a synchronized stimulation and subsequently an effective pumping of the blood. This cycle is recurrent up to the decease of the heart. Now we will try to relate the electrical and mechanical behavior of the heart described above. The activation of the cardiac muscle

composed of two phases, contraction and relaxation, or in electrical terms as depolarization and repolarization. As the heart function produces an electrical field, the voltage generated can be recorded by the electrocardiograph from the surface of the body. This process described in section 2.2.

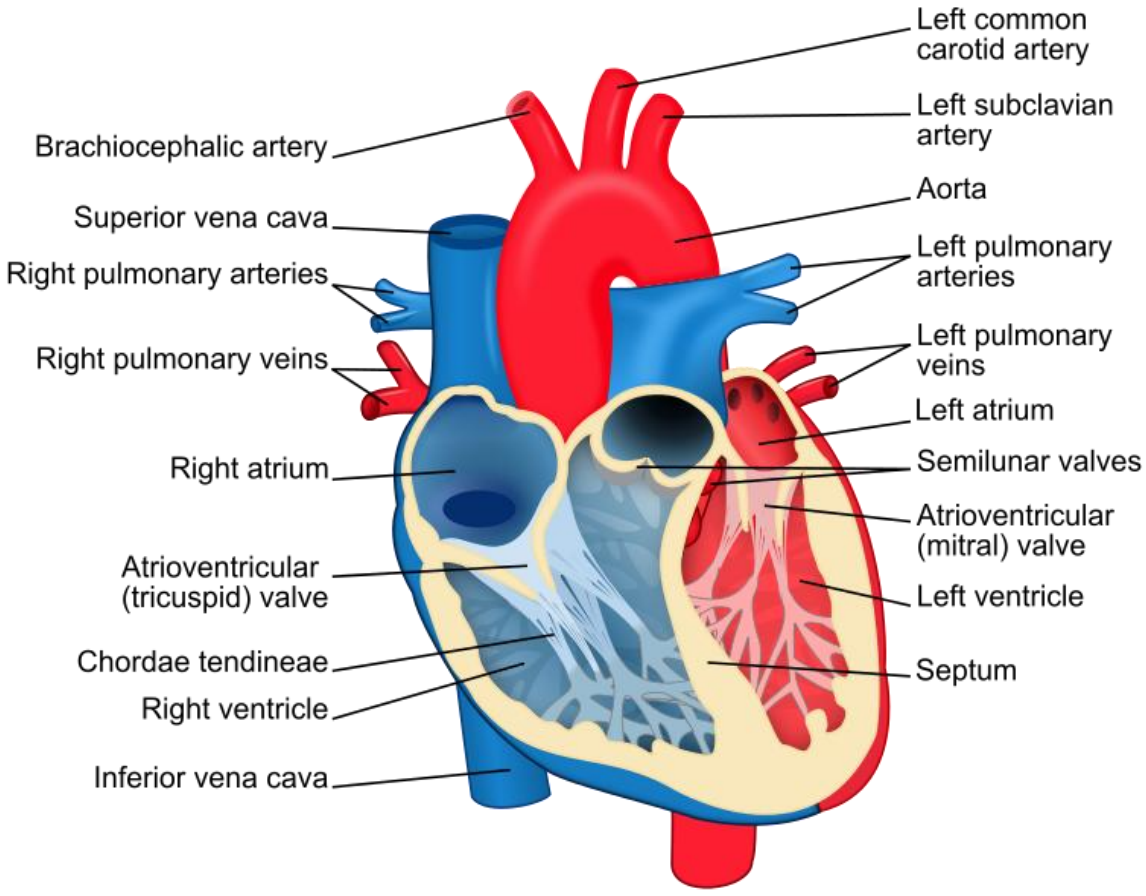


Figure 2.1: Construction of the heart, and course of blood flow through the heart chambers and heart valves. Diagrams based on image http://en.wikipedia ... -en.svg [12].

2.2 Basic Components of the Electrocardiogram

Electrocardiogram (ECG) has normally used to gather a lot of estimations that contain specific evidence in the cardiac signals. The ordinary diagnostic technique is off-line investigation from the recorded data, and utilizing a cardiogram to recognize arrhythmic categories of the subjects. The ECG is a diagnosis device that revealed the electrical signal activity of heart repolarization and depolarization recorded by skin electrode [21]. The electrical signal activity is due to

depolarization and repolarization of Na+ and K+ ions in the blood [22]. It is a noninvasive system that implies the electrical signal activity is measured on the surface of human body, which is used in detection of the cardiac diseases [22]. The ECG complex morphology and heart rate variability reflects the cardiac health of human heart [23]. Any disorder of heart rate or rhythm, or change in the morphological pattern, is an indication of cardiac arrhythmia, which can be detected by analysis of the recorded ECG signals [24]. The duration and strength of the P-QRS-T wave of ECG signals contains useful information about the idea of disease related to the heart. The ECG signal provides the following information of a human heart [25] : (i) heart location and its relative chamber size (ii) origin of impulse signals and its propagation (iii) morphological pattern, heart rhythm and conduction disturbances (iv) Level and position of myocardial ischemia (v) changes in electrolyte concentrations (vi) Drug-effects on the cardiovascular system.

The basic components of the ECG signal are illustrated in Figure 2.2. The first impulse, named as P wave, relates to the depolarization of the left and right atrium: the electrical signal spreads from the Sino-atrial (SA) node trench the atrium [26] [27]. The PR (or PQ) interim is the time period between the onsets of ventricular and atrium depolarization. A short duration isoelectric line is extant within the PR interim. This is known as PR segment and reaches out from the finish of the P wave until the start of the QRS complex. Ventricular depolarization converts into the QRS complex [28]. It ought to be noticed that atrial repolarization happens amid ventricular depolarization and is covered up in the QRS complex. The ventricular depolarization an isoelectric line is perceptible in the ECG record. This is called ST segment and stretches out from the finish of the QRS complex to the start of the T wave. It relates to an electrically neutral time for the heart, between ventricular depolarization and repolarization. Depression or elevation of the ST segment might be indicative of myocardial damage [29]. The T wave always follows the QRS complex because it represents ventricular repolarization. The electrical depolarization is preceded by the corresponding mechanical contraction. The ordinary T wave ought to be a similar way of the QRS complex and is marginally lopsided. The time between the end of ventricular repolarization and onset of ventricular depolarization is named as QT interval [30]. Along these lines, it is estimated from the earliest starting point of the Q wave to the finish of the T wave. The length of this interval changes as per age, sex, and for the most part of heart rate variability.

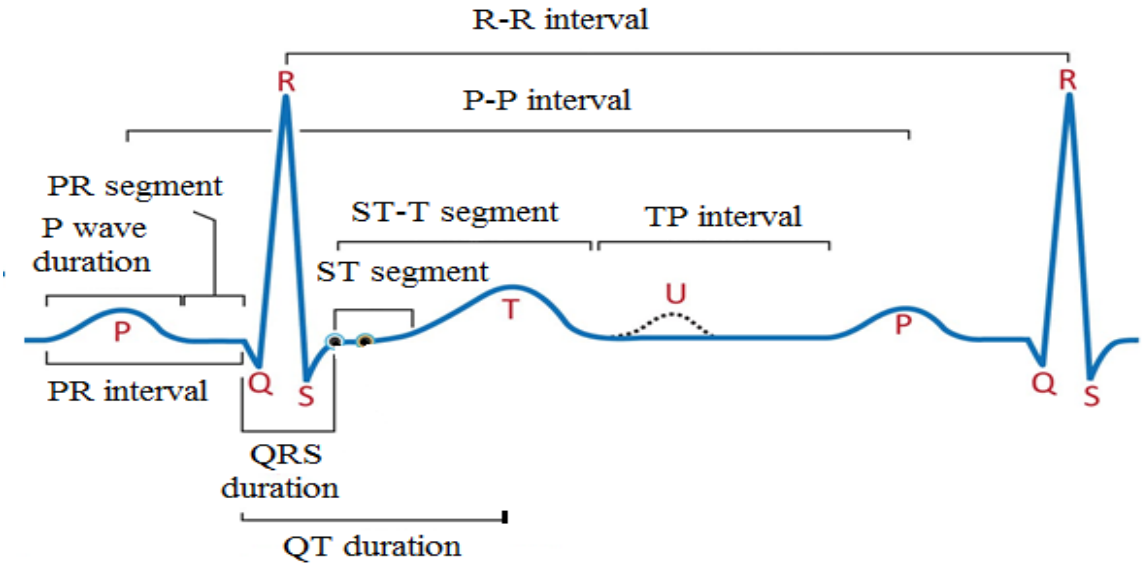


Figure 2.2: Basic components of the ECG complex of healthy human heart [14].

The U wave, with small amplitude follows the T wave, signifies late ventricular repolarization. This wave is not always appearing in the ECG recorded data and its nonappearance is not a sign of cardiac disease. U waves are more easily noticeable with slower heart rates [31] [32]. TP interval is not the reference level for measuring the ST segments. The correct reference level is PR segment. This reference level is known as baseline level or isoelectric level.

2.3 Frequency Components of Heart Rate Variability

Morphological format impression of ECG signal is basic yet not sufficient to recognize the presence of CAD because subtle variations are challenging to identify by ECG alone in CAD patients, hence, it is a must to portray these ECG signals into heart rate variability (HRV) signals [33],[34]. The analysis of (HRV) has become a standard tool for studying a wide variety of clinical and research topics. These areas have ranged from autonomic nervous system (ANS) regulation risk stratification for sudden cardiac death, diabetic neuropathy, pharmaceutical evaluations, to psychological disorders [35], [36], [37], [38]. The variation of consecutive R-R interval on the QRS templates of ECG records is known as HRV. The HRV is influenced by multiple neural and hormonal inputs that generate specific observable rhythms in the series [39]. The dynamic characteristics of HRV are measure of the balance between sympathetic and

parasympathetic mediators (both are branches of ANS) of heart rate. Sympathetic mediator is initiated from influence of epinephrine and norepinephrine chemical messengers (neurotransmitters). These chemical is released from sympathetic nerve fiber and acts on sino-artial (SA) node and atrio-ventricular (AV) node [40]. Which increase the rate of cardiac contraction and accelerate electrical system at the AV node, leading to increase heart rate. Sympathetic mediators appear to exert their effect over longer time periods and are replicated in the low frequency (LF) band (0.04 to 0.15 Hz) of HRV [41]. In this band, power spectral distribution is known by LFp [42]. Parasympathetic mediator is initiated from effect of acetylcholine chemical neurotransmitter. This chemical is released from parasympathetic nerve fiber and acts on SA node and AV node, which slows the rate of cardiac contraction of AV node, leading to decrease in heart rate. Parasympathetic mediators employ their effect more quickly on the heart and are reflected in high frequency (HF) band (0.15 to 0.4 Hz), the spectral power distribution in HF band is represented by HFp of HRV [43]. Thus, the power spectral analysis is performed on the HRV data to show vagal tone and the sympatho-vagal balance as LFp/HFp ratio at any point of instant [44],[45],[46].

Sporadic constituents of HRV tend to amass within several frequency bands. The respiratory frequency is most noticeable in healthy subjects at sinus arrhythmia, in which the sinus rate decreases with expiration and increases with inspiration. The respiratory spectrum band is considered from around 0.15 Hz to 0.4 Hz in humans however it may be extended beneath 0.15 Hz and up to 1 Hz or more for neonates or infants and for healthy persons during work out [47]. This spectrum band is well-known as high frequency (HF; 2.5 to 6 secs cycle length in humans), which are synchronized to the respiratory sinus rhythm [48] and are predominantly modulated by dint of heart parasympathetic innervation [49]. The HRV oscillations also take place at low frequencies (LF) and is considered from about 0.04 to 0.15 Hz, including a 0.1Hz frequency component that is some of the time alluded to as the 10-s rhythm or the Mayer wave [50]. This frequency assortment has been named as the mid-frequency (MF) band by some authors [51]. The LF spectrum components of HRV have been suggested to reveal mainly sympathetic activity. But it is believed by most of the researchers that the LF (> 6 sec cycle length in humans) band are sensitive to change in cardiac presumably parasympathetic and sympathetic nerve activity [52].

Other HRV oscillations occur at frequencies underneath 0.05 Hz. These frequencies have been designated differently, but generally used frequency bands incorporate as very low frequencies (VLFs; about 0.004-0.04 Hz or > 25 sec cycle length), ultra low frequencies (ULFs below 0.004 Hz or > 5 hrs cycle length) [53]. In the VLF band, HRV oscillations have been studied much less than higher and lower frequency rhythms. It may reflect thermoregulatory cycles [54] or fluctuations related to plasma renin activity [55], [56], including changes of activity, breathing, posture, autonomic outflow, temperature regulation, state of arousal [57] [58] and humoral systems [59]. These attributes will not be considered because their mechanisms and origins stay mystery. Even though, VLF or faster rhythms of HRV may have psychophysiological correlates and clinical applications [60].

Like many physiological signals (respiratory, heart rate, blood pressure), HRV signals also show non-stationary and non-linear phenomenon. Its variation may provide information about current disease, or caveats about future cardiovascular diseases [61]. Hence, HRV variation estimation has become a standard non-invasive implement for assessing the cardiac diseases and activities related to ANS. The features extracted from HRV signals have extensive applications in biomedical for distinguishing the cardiovascular diseases. Thus HRV is a noninvasive valuable tool to investigate the sympathetic and parasympathetic function of the ANS [62].

Power spectral investigation of HRV signals in frequency domain shows specific frequency variation in the cardiac record which reflects central neural control (CNC) of sinus node activity (SNA) [41]. Spectral approaches produce a disintegration of total disparity of a HRV data time series into its spectral components. The HRV signals can be articulated in the form of a power spectral density function which illustrates spectral power as a variation of frequency [63]. Power spectral for a given spectrum band can be measured by deriving the region under the spectral density variation contained by the specified frequency band [64]. Kay and Marple [65] presented a broad outline of a considerable lot of the strategies accessible for spectral analysis of HRV signals. The two most common techniques like fast Fourier transform (FFT) technique [52] and autoregressive (AR) modeling [66] are using for spectral analysis of ECG components. The fundamental distinction between the FFT and AR approaches is the way by which the data are viewed. In FFT, (i) the analysis expect that the time arrangement contains just deterministic parts. (ii) The spectrum computed is derived from all the data regardless of how well they fit in a

model based on peaks in the spectral distribution. (iii) The spectrum concentrates on the more significant peaks including with noise. In AR, (i) the analysis expect that the time arrangement as a composite of stochastic and deterministic segments. (ii) The time-domain data are used to identify a best-fit model from which a number of peaks and the final spectrum are derived. (iii) The technique concentrates on the more significant peaks, attempting to exclude noise. Thus, in its most exceedingly terrible fundamental application the FFT approach could be viewed as an expressive strategy and the AR approach would be more predictable with a statistical or stochastic approach for ECG components analysis. In Practical application, this refinement is obscured by the basic use of smoothing calculations or windowing to stabilize variance estimates from FFT analyses. In spite of the fact that there are a number of disadvantages and advantages to each of these strategies, there are also numerous similarities that in practice usually lead to basically equivalent results [67].

2.4 Arrhythmia

Arrhythmias are defined as any cardiac rhythm other than the normal sinus rhythm (NSR). Sinus rhythm originates in the sinus node and subsequently is conducted at appropriate rates through the atria, AV junction, and the intraventricular specific conduction system. At rest the sinus node discharge cadence tends to be regular, although it presents generally slight variations. However, under normal conditions and particularly in children, it may present slight to moderate changes dependent on the phases of respiration, with the heart rate increasing with inspiration. In adults at rest the rate of the normal sinus rhythm ranges from 60 to 100 beats per minute (bpm). Thus, sinus rhythms over 100 bpm (sinus tachycardia) and those under 60 bpm (sinus bradycardia) may be considered arrhythmias. The ECG complexes of arrhythmias are shown in Figure 2.3. However, it should be taken into account that sinus rhythm varies throughout a 24-h period and sinus tachycardia and sinus bradycardia usually are a physiologic response to certain sympathetic (exercise, stress) or vagal (rest, sleep) stimuli. Under such circumstances, the presence of these heart rates should be considered normal. The term arrhythmia does not mean heart beat anomaly, as normal arrhythmias can happen frequently with absolute stability (paroxysmal tachycardia, flutter and so forth.), sometimes the heart showing heart beat in the normal range. On the other hand, some unpredictable rhythms ought not to be thought about arrhythmias (direct inconsistency in the sinus release, especially when related to breath). In addition, a finding of

arrhythmia in itself does not mean manifest pathology. Indeed, in healthy subjects, the irregular presence of specific arrhythmias both dynamic (active) (premature complexes) and passive (escape complexes, certain level of AV block, sinus arrhythmia, and so forth.) as often as observed.

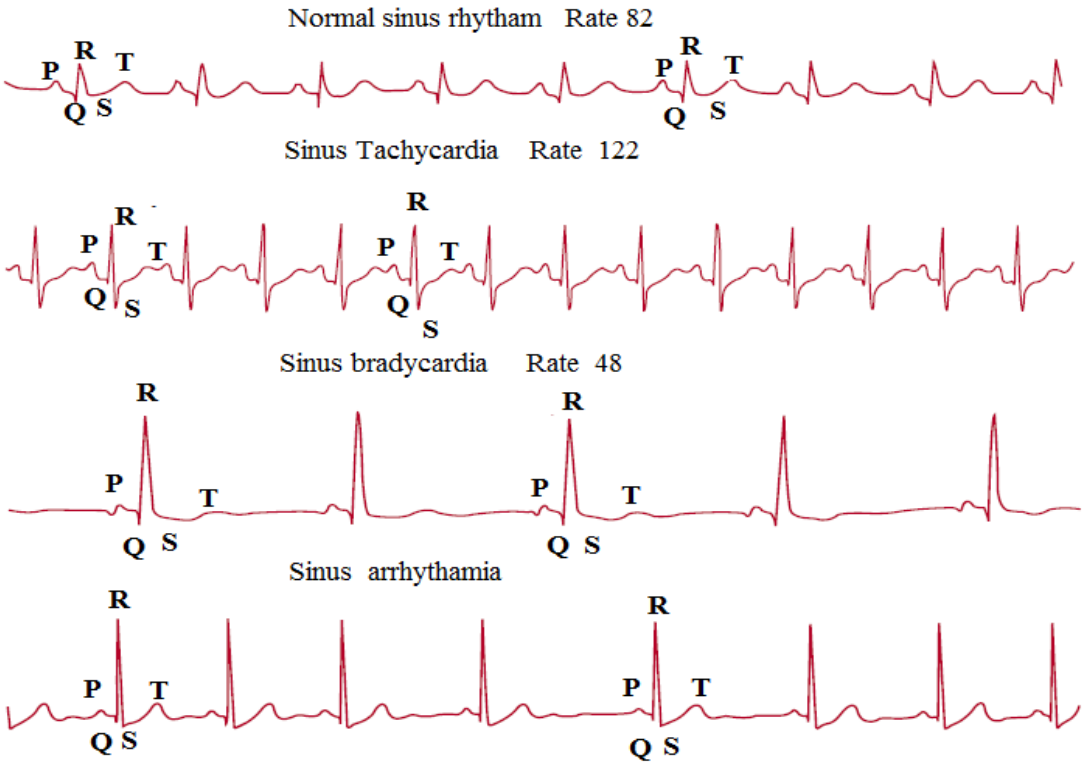


Figure 2.3: Several examples of sinus rhythms

There are different ways to characterize heart arrhythmias [16]

- Location of origin: arrhythmias are alienated into ventricular arrhythmias and supraventricular (containing those having their origin in the atria, AV junction and SA node).
- Underlying mechanism of conduction: arrhythmias may be divided by: 1) irregular formation of electrical impulses, which consist of increased heart automaticity (superfluous parasystolic or systolic mechanism) and triggered electrical impulse activity, and 2) disturbances of conduction and/or decreased automaticity.
- According to clinical point of view: arrhythmias may be incessant, or paroxysmal, or perpetual. In reference to tachyarrhythmia (an instance of a dynamic arrhythmia), paroxysmal tachyarrhythmia befall unexpectedly and generally disappear naturally (for instance AV junction reentrant paroxysmal tachycardia). Incessant tachyarrhythmia is categorized by petite and

repetitive runs of ventricular tachycardia or supraventricular. Permanent tachy arrhythmias are continuously present (an example: chronic atrial fibrillation) [68].

2.5 Time- Frequency Analysis of Heart Rate Variability

In ideal case, the time-frequency plotting provides only information about the frequency occurring at a given instant of time [69] and with excellent resolution and without cross interface, if exist any [70]. The energy concentration related with resolution in the time-frequency analysis is one of its most important and severely studied aspects in time-frequency analysis of HRV signals [71]. Various popular methods like short time Fourier transform (STFT) [72], wavelet transform and Stockwell transform (ST) exist for linear time-frequency analysis of HRV signals [73]. In STFT, temporal resolution improvement is obtained at the expense of frequency resolution due to fixed duration window [74]. The multiresolution technique overcomes the problem of STFT. By using multiresolution strategy, it provides fixed number of cycles per dilation and resolution remains stationary along with the dilation. The standard ST is becoming a popular method for time-frequency analysis of non-stationary signals due to its simplicity and preservance of the phase information of signals [75]. The Gaussian window function of ST depends on frequency; hence the ST produces excellent frequency resolution at lower frequencies and sharper time localization at higher frequencies [76], [77]. The characteristic of window function changes with frequency and not with time, thus the ST method is inappropriate for resolving signals, whose spectral components are fast changing with time [78]. This reality has persuaded the utilization of frequency domain strategies, for example, Fourier transform (FT), for investigation cardiovascular signal [79].

A number of wavelet-based techniques have been proposed for the detection, classification and analysis of arrhythmic ECG signals. Govindan *et al.* [80] have developed an algorithm for classifying four groups as Atrial Fibrillation (AF), NSR, chronic AF and paroxysmal AF. In this analysis, they used a Daubechies (D6) wavelet to extract the features and preprocess the ECG data of the considered groups of arrhythmias before to classification using an artificial neural network. They have found that the classification of average success rate around 77% ± 9% for paroxysmal AF and achieved 94% ± 8% for NSR. However, in this study small numbers of training and validation data were used. Khadra *et al.* [81] were used a raised cosine wavelet transform method for investigation of three arrhythmias like Ventricular fibrillation (VF),

Ventricular Tachycardia (VT) and AF. They proposed an algorithm built on the scale-dependent energy using wavelet decomposition to classify the arrhythmias. In this study also involved low numbers of data: 12 VT, 13 VF, 13 AF and 8 NSR.

Zhang *et al.* [82] have suggested a novel arrhythmia detection strategy; in view of a Morlet wavelet transform for use in implantable defibrillators. Their framework, assist to recognize connections between simultaneous epicardial coronary artery diseases related to the cell action potentials and bipolar electrogram. Their system, also detects the bifurcation point in the ECG complex, where NSR degenerates into a ventricular fibrillation (VF). Al-Fahoum and Howitt [83] have proposed a radial basis neural network for the automatic classification and detection of both VF and VT arrhythmias which employ the Daubechies D4 wavelet transform for preprocessing of the ECG complex. They used a database of 159 arrhythmia files from three distinct sources for classification and reported that the correct classification of arrhythmia is more than 97.5%. Morlet *et al.* [84] presented a Morlet wavelet-based method for the discernment of patients suffered to the onset of VT. They found that the identification of strings of neighborhood maxima of the Morlet wavelet vector at or after 98 ms after the QRS onset point was a reasonable basis for VT risk stratification in post-infarction patients. They achieved specificity 85% and sensitivity 90% for considered patients group. Englund *et al.* [85] studied about patients with hypertrophic cardiomyopathy at increased risk of malignant ventricular arrhythmias or sudden death. For identification of these patients, they used predictive value of wavelet decomposition of the signal averaged ECG complex. They resolved that wavelet decomposition was of limited value in analysis of hypertrophic cardiomyopathy patients [86].

To develop an analytic strategy for arrhythmia detection and classification, signal investigation based on wavelet transform (WT) has been proposed to extract the basic components of ECG signals [87]. The wavelet coefficients indicate measures of similarity of the local shape of ECG complex under various translation and dilation parameters. This investigation is vigorous to non-stationary signal analysis, yet it isn't well-found for classification. With advantages of multiresolution and localization capability of the wavelets and pattern detection capability of the artificial neural network (ANN). A method was developed by Lin and Wang [88], in which ANN combined with wavelet transform that is known as wavelet network (WN) has become important for classification and pattern recognition non-stationary signal. After validation of this method on non-stationary signal, then the Lin *et al.* [89]

proposed a method for ECG ectopic heart beat detection and classification of arrhythmias using adaptive wavelet network (AWN). The ECG ectopic heart beat detection algorithm divided into a series of stages, feature extraction from QRS complexes, and afterward according to characteristics of the cardiac arrhythmias. They detected the bundle branch ectopic beat, supraventricular ectopic beat, ventricular ectopic heart beat and classifying bundle branch, Supraventricular Tachycardia (SVT) and VT arrhythmias. Martinez *et al.* [90] proposed a method based on the WT to assess a robust single-lead ECG delineation scheme. In an initial step, QRS complexes were recognized. After this, each QRS complex is delineated by detecting and classifying the peaks of the individual waves with complex onset. At last step, the detection of T wave and P peaks, onsets and ends was accomplished. They assessed the algorithm on MIT-BIH Arrhythmia database.

Recently, several detection and classification approaches have been proposed, namely the consecutive hypothesis detection technique [91], the multiway successive hypothesis detection [92], the threshold crossing interval [93], approaches based on supervised neural-network systems [94], and Time-Frequency (T-F) method like wavelet transform [95], hve been used for analysis of arrhythmias. New methodologies were used complexity and chaos measure [96] and multifractal investigation joined with a fuzzy Kohonen neural system [97].

Zhan *et al.* [98] proposed a time-frequency coherence using continuous wavelet transform (CWT) together with its confidence intervals to be evaluated the correlation between two non-stationary processes. A systematic comparison between approaches using CWT and short-time Fourier transform (STFT) is carried out. Synthetic data are generated to test the performance of these methods. Benesty *et al.* [99] proposed a method known as Minimum variance distortionless response (MVDR). This method is very popular in array processing and power spectral estimation and finding in coherence value. It is better than Welch method, because it is designed by special filter. Lackner *et al.* [100] studied the synchronization of heart rate, blood pressure and respiration in the sympathetic and parasympathetic branches of the autonomic nervous system during a cancellation test of attention and during mental arithmetic tasks. Fonseca *et al.* [101] described a simple and causal coherence from respiration to HRV was higher during inspiration than expiration. The estimates of gain also differed significantly in the HF band between those obtained from the whole recording of the aspiratory and expiratory

periods. These results indicate that a single linear model fitted to the whole recording and neglects potentially significant differences between inspiration and expiration, and authors describes how such differences can be estimated, without the need to control breathing. Valenza *et al.* [102] studied to investigate the synchronization between breathing patterns and heart rate during emotional visual elicitation, that is, using sets of images gathered from the international affective picture system having five levels of arousal and five levels of valence, including a neutral reference level. Gallet and Julien [103] proposed an approximate value of the coherence threshold when using the Welch's periodogram method with overlapping segments. For the most common case (50%), it allows to estimate the theoretical threshold correctly except when less than seven segments are used. Yoshino and Matsuoka [104] used Causal coherence analysis based on a closed-loop bivariate autoregressive model and applied to HRV and systolic blood pressure (SBP) variability analysis during mental arithmetic tasks to clarify how mental task load affects the linear closed loop interaction between cardiac and vascular systems. Mormann *et al.* [105] used the mean phase coherence as a statistical measure for phase synchronization, they observed characteristic spatial and temporal shifts in synchronization that appear to be strongly related to pathological activity. Keissar *et al.* [106] presented a framework for applying wavelet transform coherence (WTC) for quantitative analysis of coherence in cardiovascular variability research. Computer simulations were performed to estimate the accuracy achieved by WTC method and also a method for determining the coherence threshold for specific frequency band was developed and evaluated.

The TVSC has been used in many field of science, including neuroscience, detection of uterine electrical activity synchronization in labor [107], sleep apnea, time-varying estimation of correlation between cardiovascular signals and baroreflex sensitivity [108]. Further, application of time varying coherence has been mentioned detailed in discussion section. Typically, Parametric and nonparametric estimators has been introduced to estimate the TVSC of the signals. The most of parametric methods are frameworked on the base of Yule-Walker autoregressive modeling [109]. But these methods are subjected to error due to model misspecification. For highly non-stationary signals these methods has performed less accurately compared to nonparametric methods [110]. The estimation of nonparametric methods not required assumption on the order of model. Including of these methods, in analysis of

cardiovascular signals, the short time Fourier transform (STFT), Smoothed Pseudo-Wigner-Ville distribution (SPWVD), Multitaper Spectrogram (MTSP) and standard Stockwell transform (SST) has been recently proposed to detect the transient phenomenon of cardiovascular signals [111]. The use of Wavelet transform coherence in cardiovascular variability research was first introduced by Keissar *et al.* [112] and further Ostlund *et al.* [113] explored the coherence between heart rate and respiration signals.

2.6 Features Extraction of HRV signal

Feature extraction from HRV is a key phase of any general structure for classification and detection of cardiac disease or, all the for the most part for any neural network, machine learning design and system that entails automatic decision making capability. This frequency-domain features [114] include average low frequency power, high frequency power, spectral flatness, chaos, spectral flux and nonlinear features [115]. Hoyer *et al.* [116] presented that quantitative measures of complexity by correlation dimension (CD) and predictability by largest Lyupnov exponent (LLE) chaos method that provide significant information about ANS processes. Detrended fluctuation analysis (DFA) method is an alteration of root-mean-square investigation of random walks has been used to analysis of non-stationary signals [117]. There is well-organized nonlinear behavior of HRV and RESP, which can be interpreted with regard to terms such as nonlinear stochastic, regular deterministic and chaotic. According to authors Porta *et al.* [118], changes in complexity of short-term heart period variability can be reliably related to modifications of the state of the autonomic nervous system, mainly to an increase of sympathetic modulation, thus suggesting the use of unpredictability indexes to estimate cardiac sympathetic modulation. In healthy subjects at rest, the short-term heart period variability is mainly linear and the unbalancing of the autonomic nervous system towards a sympathetic (or parasympathetic) prevalence does not increase nonlinear components. Therefore, the use of nonlinear models might not be supported by data and might not increase our ability to derive information from short-term heart period variability. Faes *et al.* [119] proposed a method to perform time-varying (TV) nonlinear prediction of biomedical signals in the presence of non-stationarity. The method is based on identification of TV autoregressive models through expansion of the TV coefficients onto a set of basis functions and on k-nearest neighbor local linear approximation to perform nonlinear prediction. The approach provides reasonable nonlinear prediction even for TV

deterministic chaotic signals, which has been a daunting task to date. Binachi *et al.* [120] proposed a method for processing of ECG and respiratory signals, which aim to detecting parameters like LF and HF components of respiratory signals and HRV signals for correlate with normal and diabetic subjects. Sunkaria [121] has explained the non-linear methods of HRV analysis such as correlation dimension, largest Lyupnov exponent, power law slope, fractal analysis, detrended fluctuation analysis, complexity measure etc. for HRV analysis. Entropy and correlation methods constitute most popular and powerful tools to estimate the dynamical physiognomies of the HRV signals. Higher entropy normally exhibits higher uncertainty, however lower entropy indicates more certainty and regularity of the cardiac system [122]. There are a number of entropy approaches commonly applied to HRV analysis, such as approximate entropy (ApEn), non-linear fuzzy entropy (FZE) [123], non-linear Kraskov nearest neighbour entropy (K-NNE) [124], permutation entropy (PE) [125], wavelet entropy, multiscale permutation entropy (MPE), improved multiscale permutation entropy (IMPE) [126]. The multiscale Wavelet Packet (MSWP) transform has been proposed by Coifman *et al.* [127] to decompose the HRV signal for detection of cardiac diseases. They extend the link between wavelets and multi-resolution approximations. The approximate entropy **(**ApEn) is a new family of statistics, measuring the chaos and regularity of HRV time-series data. The use of approximate entropy **(**ApEn) for HRV analysis has been described in [128]. Poincare plot analysis is a rising quantitative-visual method, whereby the shape of the plot is sorted into useful classes, which show the level of the cardiovascular illness subjects [129]. The multiscale permutation entropy (MPE) technique cannot provide a consistent investigation for short HRV time series [130]. To overcome this problem, an improved MPE (IMPE) was proposed [126]. In which, many successive coarse-grained versions [131] are set made by averaging the time series data points within non-overlapping windows of expanding length, which is called scale factor. The new method as bi-correlation is not a widely discussed for measurement of chaos of HRV signal, but the cumulative bi-correlation (CBC) has been applied in the way of the statistic of the alleged as Hinich test of non-linearity or linearity, under the constraint that the second delay is twice the first delay (Hinich 1996). The CBC is the accumulative function of the bi-correlation for the specified range of delays [132].

A statistical Student's t-test is method also employed to resolves whether or not two sets of feature are dissimilar. The Student's t-test produces the so called p-value. The p-value is directly

proportional to the similarity of the two features [133]. A box plot was also employed for representing groups of numerical data along with their quartiles values. It may also have straight lines outspreading vertically from the boxes (whiskers) indicative of erraticism separate the upper and lower quartiles with box plot [134]. Outliers are plotted as separate points or plus symbols. Box plots are non-parametric, they show variety in tests of a measurable data without making any presumptions of the underlying statistical dissemination.

Many set of procedures were developed for classification of cardiac diseases using HRV time series data in the last decades. Most likely and most applicable research paper before the start of this thesis are [135], while others were published over the most recent years [136]. However, because of the lack of standardization in database the development and assessment criteria, investigation of results across most of these works couldn't be performed reasonably. In order to mitigate this issue, some methodological aspects in the advancement and assessment of HRV classifiers were followed in recent works [137]. The most relevant key-points are:

- Usage of standard databases, as the ones accessible in Physionet ATM
- Standard data division into training and testing sets, as described in [138].

Concerning to the features used for the classifier, the HRV time series data were used in nearly all published research works. Other typical decisions were based on ECG sampled data (commonly from the QRS or T wave of ECG) [138] or transformed by Hermite method transform [139] or WT decomposition [138], features extracted from the demarcation of the QRS complex and T wave duration, lead to suitable for classification. As a result of the large changes in the HRV time series patterns of a number of cardiac disease classes, there is commonly a significant similarity between some of cardiac disease classes like CHF and CAD, Elderly and CAD subjects in the non-linear feature space. In these classes is the larger similarity with each other, it is challenge to differentiate between the two [140]. In this condition, a feature dimension transformation technique will be very useful. Various techniques have been developed to reduce the features data size for classification problem like kernel principal component analysis (KPCA), independent component analysis (ICA), linear discriminant analysis (LDA) and generalized discriminant analysis (GDA) [141]. Due to, an advantage of GDA and KPCA, both dimensionality reduction of the input feature space and selection of the useful discriminating features are based on nonlinear kernel function [142].

2.7 Machine Learning Classification and Prediction of Cardiac Disease

The machine learning and artificial intelligence methods found a powerful tool for the binary and multi-class classification and prediction of cardiac disease. A number of classifiers were used in the published paper like simple LDA based on the Gaussian assumption of the data [138] to others more expanded, as artificial neural networks (ANN's) [143], One drawback of several patient-adaptable approaches is that they cannot operate without assistance [137]. This is not the case of those developed as an evolution of a previous automatic classifier [138]. It is fascinating to take note of that a few classifiers require from 2 to 5 minutes of manual explanations, which is equivalent to several hundred of annotation labeled on R-R interval [144] while [144] needs the annotation of a few heartbeats, contingent upon the number of arrhythmias present. One limitation of several patient-adaptable methodologies is that they can not work without assistance [144].

Recently, a new learning process for single hidden layer feedforward neural networks (SLFNs) design called extreme learning machine (ELM) strategy has been extensively used in many field of bio-medical signal analysis [12]. The primary advantage of ELM is that the hidden layer of SLFNs require not be tuned and it also has fast rate of convergence [145]. Due to random nature of the hidden layer of ELM, output matrix may reduce its learning accuracy [146]. The support vector machine (SVM) kernels is replaced by ELM kernels in the SVM formulation, as it was discussed in [147] that better generalization can be obtained. This was perceived further that the suggested technique accomplishes similar or better generalization performance in contrast to SVM and is less sensitive to the input vector parameter. For the study of ELM as a combined learning algorithm used in application of classification and feature mapping and its relationship with proximal SVM (PSVM) and least squares SVM (LS-SVM), the reader can follow the article to [148].

In recent years, the application and study of 1-norm regularization or penalty has been discussed in [149], as 1-norm have a tendency to sort certain of the fitted coefficients of the ELM model by making them zero and hence provides sparse ELM models that are simply interpretable. On the premise of literature survey of 1-norm SVM problem, a method defined as a linear programming optimization problem by Mangasarian [150]. The main advantage of the 1-norm linear programming extreme learning machine (1-NLPELM) method is that it is a sparse model representation whose solution is achieved by solving an arrangement of linear equations at a finite number of times [151].

CHAPTER: 3
Methodology

3. R-Peaks and Heart Rate Variability Extraction from ECG Signal

3.1 Heart Rate Variability Signal

The electrocardiogram (ECG) is a tool which provides information related to cardiac diseases in the form of PQRS morphological signal. Morphological reflection of ECG is essential but not adequate to distinguish between CAD and healthy young and elderly subjects because in order to detect small variations in R-peaks are challenging by ECG alone; henceforth, it is necessary to characterize the signal of ECG into heart rate variability (HRV) data [152], [61]. The estimation of HRV is done through computing time interval between two consecutive peaks of R points $(R_{n+1} - R_n)$ where $(n = 1,2, \dots \dots)$ on the PQRS morphology of ECG.

For addressing the 1st and 2nd part of specific objectives following stage performed.

3.2 Noise Cancellation and R-Peak Detection

Stage 1st: Noise Cancellation and R-Peak detection using Pan Tompkins algorithm is shown in Figure.3.1

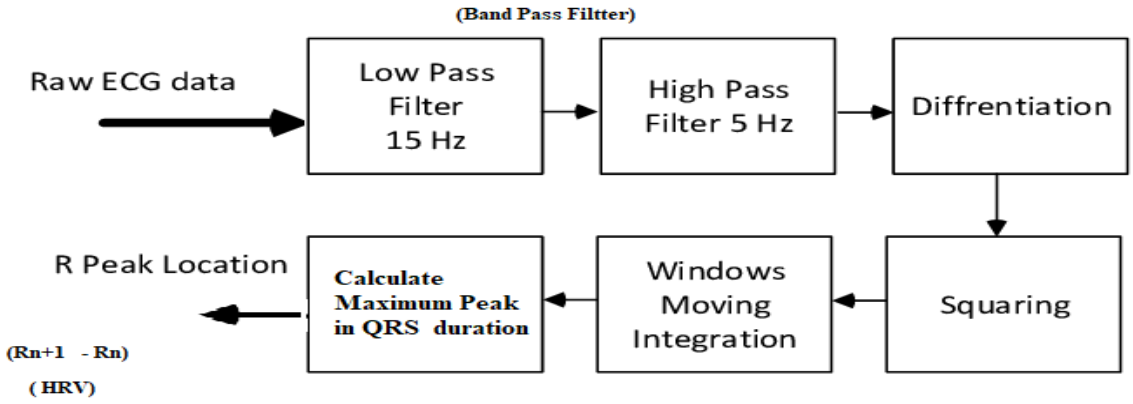


Figure.3.1. Block diagram for noise cancellation and R-peak detection

- As a first and 2nd step, a band-pass filter is applied to increase the signal-to-noise ratio. A filter bandwidth of 5-15 Hz is suggested to maximize the QRS contribute and reduce

muscle noise (above 100 Hz), baseline wander (0.5 to 0.6 Hz) , power line interference (50 to 60 Hz) and the P wave/T wave frequency content (1-4 Hz) .

- As a 3rd step, a derivative filter is applied to provide information about the slope of the QRS are shown in Figure 3.2.
- 4th step: The filtered signal is squared to enhance the dominant peaks (QRS).

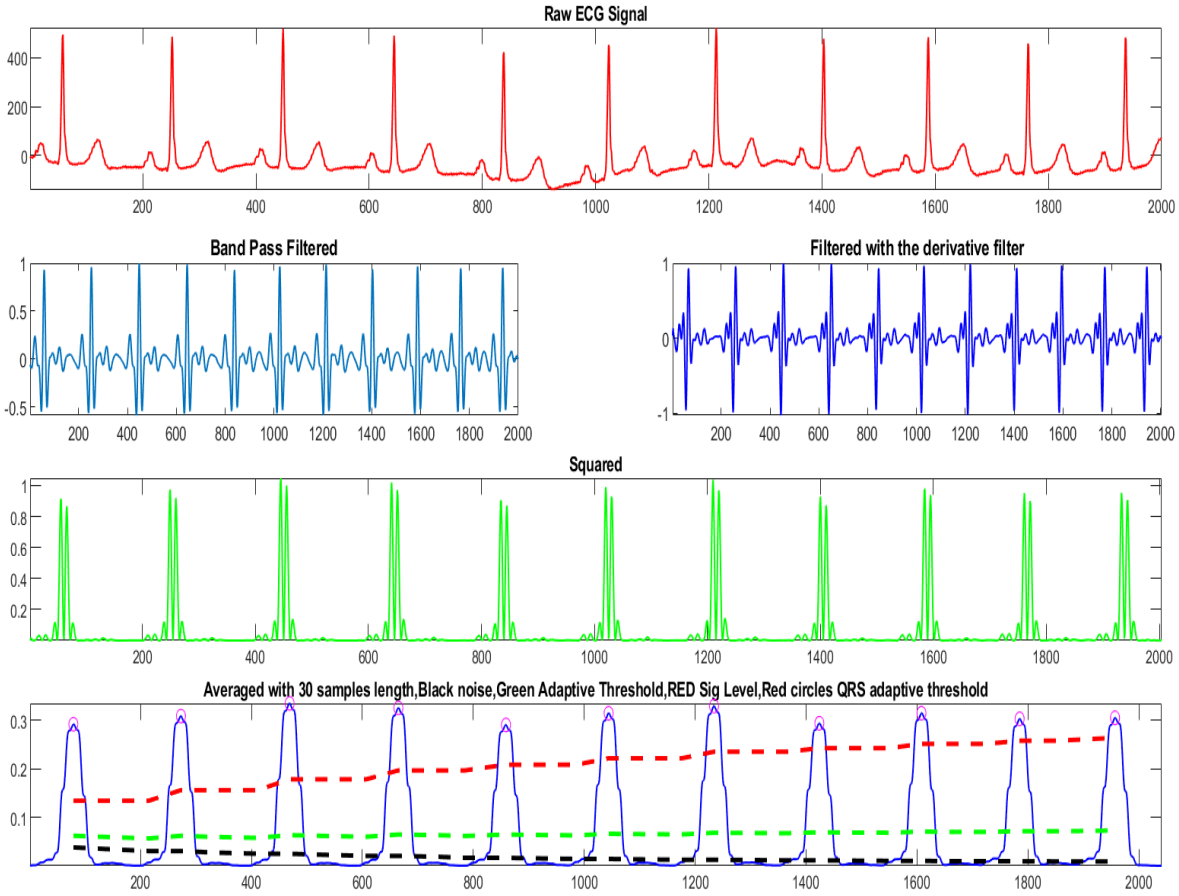


Figure 3.2: Represents (a) Raw ECG signal (b) after and Band Pass filter (c) after derivative filer (d) squaring of derivative signal (e) Average of 30 samples signal.

- 5th step: A moving average filter is applied to provide information about the duration of the QRS complex. The number of samples to average is chosen in order to average on windows of 150 ms are shown in Figure 3.3.

After extraction of HRV signal is shown in Figure 3.4, it will apply to features extraction step and Machine learning stage 2nd

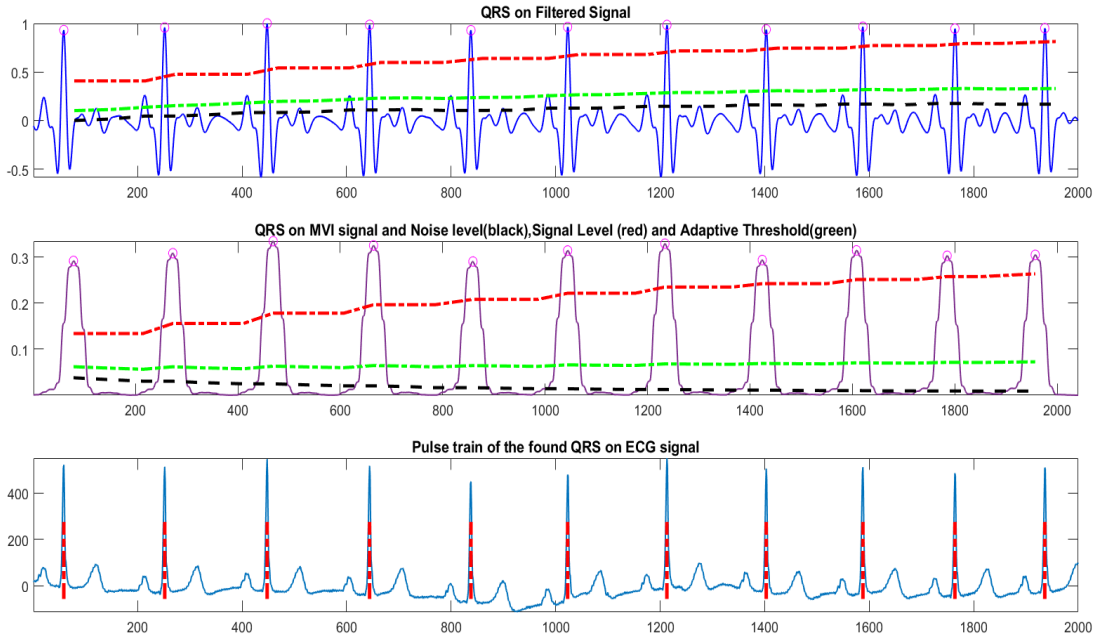


Figure 3.3: R-peaks detection (a) QRS on Filtered signal (b) QRS on windows moving (c) R-peaks

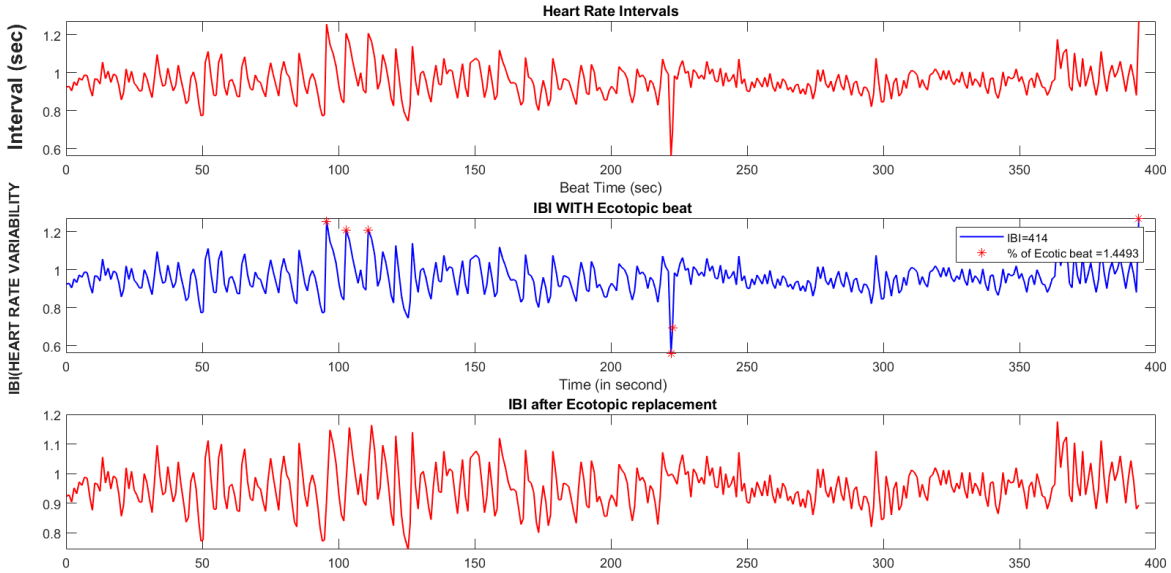


Figure 3.4 HRV extraction (a) R-R interval (b) Ectopic beat detection on R-R interval (c) Inter beat Interval (IBI) after Ecotopic replacement using spline method.

3.3 Detection and Classification

Stage 2nd: For detection and classification of CAD and CHF, nine different nonlinear features have been extracted from pre-processed HRV signal. The nonlinear features such as correlation dimension (CD), detrended fluctuation analysis (DFA), approximate entropy (ApEn), the results of the Poincare plot as SD1/SD2 ratio, Hurst exponent(HE), permutation entropy (PE), improved multiscale permutation entropy (IMPE) and cumulative bi − correlation (CBC) [3, 7, 12] have been extracted from HRV signal. Figure 3.5 shows a proposed adaptive ML model for CAD and CHF detection.

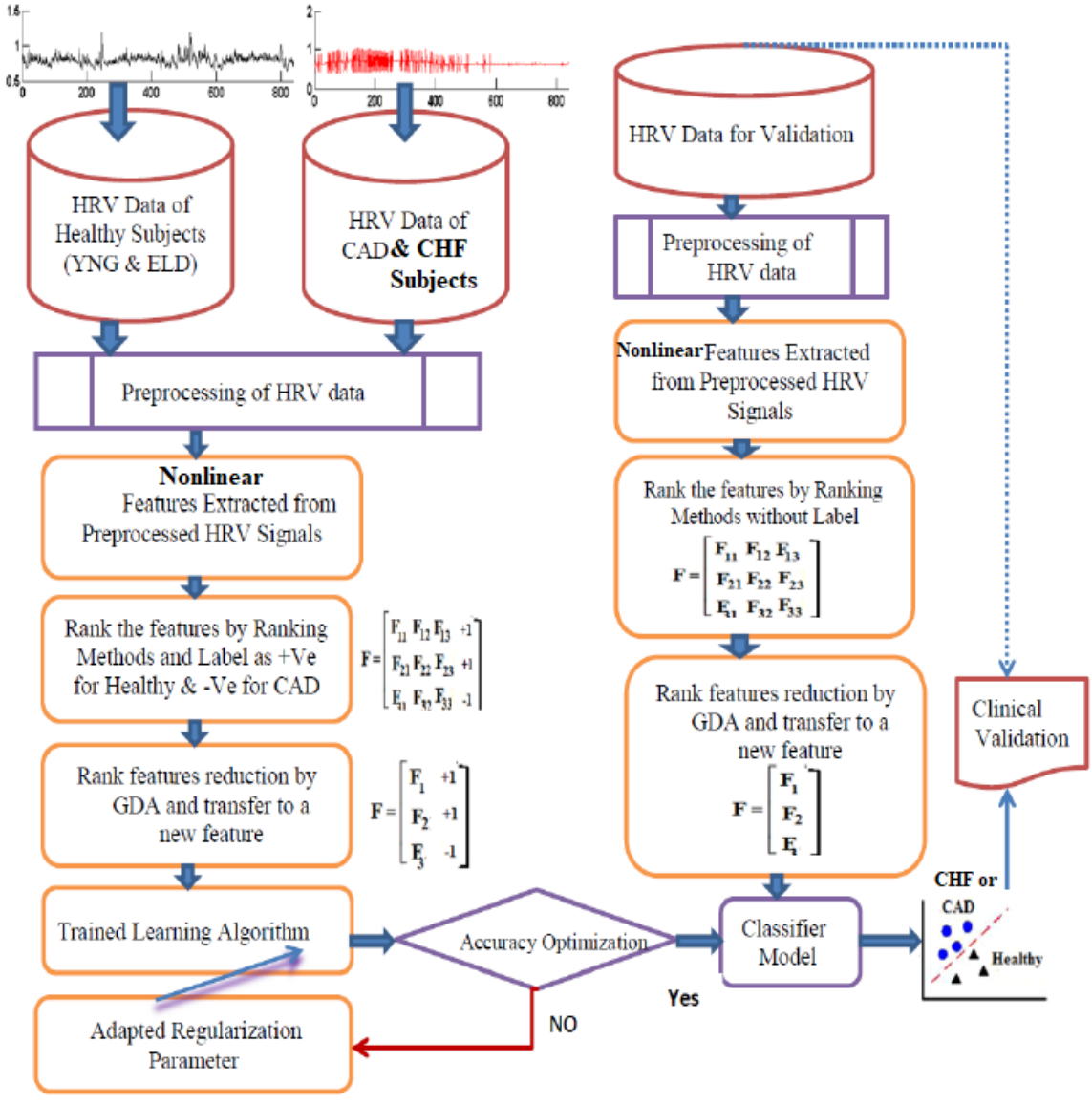


Figure 3.5: Framework of the proposed ML for CAD and CHF detection.

3.4 Hardware Framework for Detection and classification of CHF and CAD

Stage 3rd**:** The Hardware framework of the proposed model is shown in Figure 3.6. This framework consists of one android mobile phone, computer and Avidhrt Sense Device (ECG Recorder). The function of Avidhrt Sense Device is to generate ECG signal. The transfer of mobile phone data take place by wireless network (by internet) in both directions. The 1st stage algorithm will be applied to subject side mobile which will remove the noise, display the ECG signal and also show the heart beats of the subject.

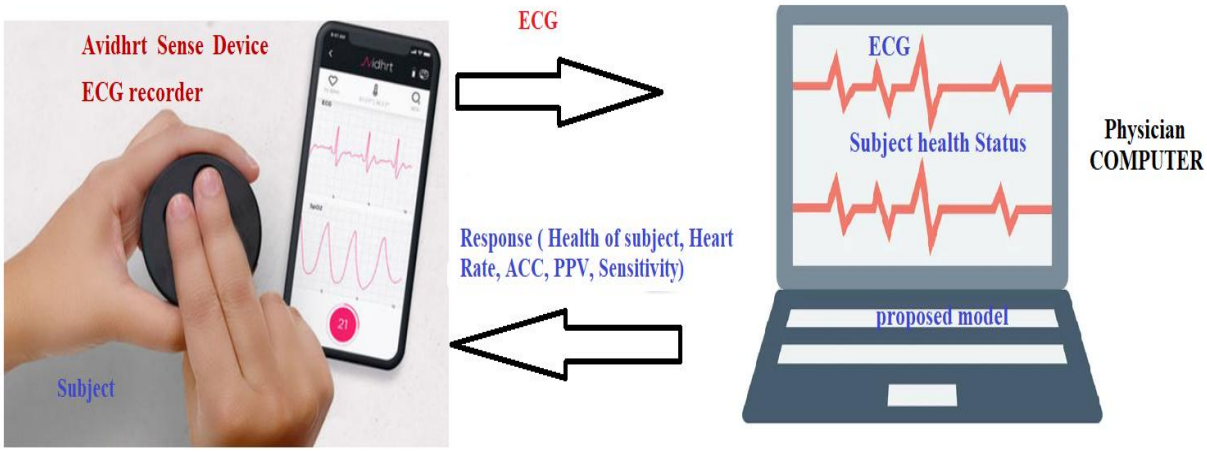


Figure 3.6: Hardware Framework for Detection and classification of CHF and CAD

- The 2nd stage algorithm will be applied to Physician side that is computer which analyzes the cardiac diseases (CAD, CHF). By this process, the Physician can suggest the cardiac health of the subject on the basis of records of analyzed data. In this way, these allow for remote patient monitoring, which is of special necessity to the elderly, who are more prone to CAD and CHF diseases and cannot always visit a healthcare facility.

4. Detection of CAD and CHF Using Ranked Nonlinear Features and OS-ELM

4.1 Introduction of Proposed Model

For detection and classification of CAD and CHF, nine different nonlinear attributes were extracted from pre-processed and segmented HRV database. The chaos and nonlinear features namely Correlation Dimension (CD), Detrended Fluctuation Analysis (DFA), Approximate Entropy (ApEn), the results of the Poincare plot as SD1/SD2 ratio, Hurst Exponent(HE), Permutation Entropy (PE), Improved Multiscale Permutation Entropy (IMPE) and cumulative bi-correlation (CBC) have been extracted from segmented HRV database. Figure 3.5 shows a proposed adaptive ML model for CAD and CHF detection. All the features are not sensitive to escalation for interpretation and comprehension of healthy and CAD database. Therefore, features were ranked using Fisher score ranking method. The most important top five ranked attributes were applied to attributes space transformation technique as Generalized Discriminant Analysis (GDA). The GDA transfer top ten features to a new feature. The values of new feature were first regularized in the range of -1 to 1, after this, fed to online sequential extreme learning machine (OSELM) having Sigmoid, Hardlim, RBF and Sine activation function binary classifier. The simulated results show that GDA with Gaussian or RBF kernel function and (OSELM) classifier with Sine activation function achieved better Accuracy (AC), sensitivity (SE), specificity (SP) and positive prediction value (PPV) as compared to other considered kernel function.

4.2 Fundamentals of ELM

The authors [12] 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 $\{X_K, T_K\}_{K=1,2\ldots,M}$ and $X_K = \{x_{K1}, x_{K2}, \ldots\ldots, x_{Kn}\}^t \in R^n$, 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 (.) between input and output layer. The construction of ELM is shown in **Figure 4.1**.

Following phases involved in learning of ELM

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

$$N = \begin{bmatrix} \emptyset(W_1X_1 + B_1,) & \cdots & \emptyset(W_mX_1 + B_m) \\ \vdots & \ldots & \vdots \\ \emptyset(W_1X_n + B_1,) & \cdots & \emptyset(W_mX_n + B_m) \end{bmatrix}_{n\times m}$$

3. Compute the output weight $\beta = N^{\dagger}T$, where $N^{\dagger}$ represents Moore-Penrose generalized inverse of N and $\beta = \{\beta_{1,}\beta_{2,} \ldots \ldots \ldots \ldots \beta_{m}\}^{t}$ 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 $\|\beta\|^2$.

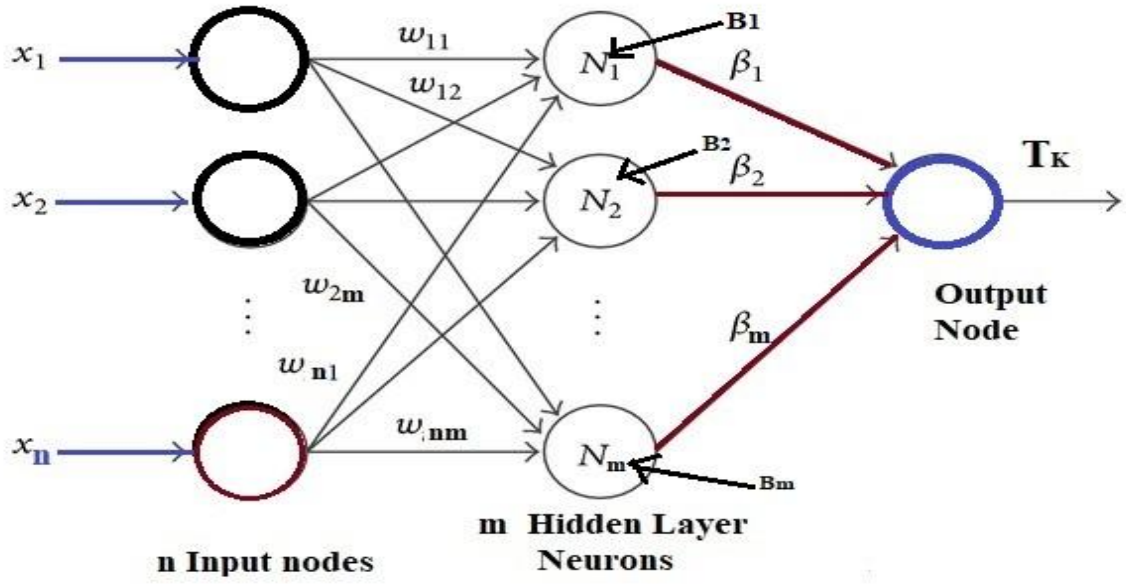


Fig.4.1. Represents construction of Extreme Learning Machine. It consists of input nodes, hidden layer neurons and output node.

4.3 Analytical Concept of OS-ELM

The concept of batch learning of ELM originates the analytical derivation of OSELM [148] . The training data can be viewed sequentially, i.e. one or the other by one or block by block of

constant or varying duration, and this process can be eliminated for training once training has been completed. It is processed in two stages:

First Stage**:** Firstly set $L = 0$, where L denotes Lth block of data. At first, training data (M_0) should not be less than m. For this training data, output of hidden layer (N_0) is defined as $N_0 = [F(W_i, b_i, X_j)]$, where $i = 1$ to m and $j = 1$ to M_0, and F is an activation function and weight between hidden layer and output (β_0) is defined as $\beta_0 = P_0 N_0^T T_0$, where $P_0 = (N_0^T N_0)^{-1}$ and T_0 denotes the class level (M_0) [15].

Second stage Sequential**:** In this stage, the output of partial hidden layer neurons (N_{L+1}) is evaluated using Moore-Penrose Inverse $(N^{\dagger})$ same as ELM where $(L+1)^{th}$ block of new observation $(N)_{L+1}$ denotes number of training data in $(L+1)^{th}$ block. The output weight matrix $\beta^{(L+1)}$ is calculated as follows:

$$\beta^{(L+1)} = \beta^{(L)} + P_{L+1} N_{L+1}^T (T_{L+1} - N_{L+1} \beta^L)$$

$$\text{Where } P_{L+1} = P_L - P_L N_{L+1}^T (I + N_{L+1} P_L N_{k+1}^T)^{-1} N_{L+1} P_L.$$

4.4 Analytical Concept of 1- *Norm Extreme Learning Machine*

$1-\text{Norm ELM } (1-\text{NELM})$ with combination of absolute loss function was proposed by [153] for addressing sparse model representation. But this algorithm agonized with optimization problem caused by not differentiability of the loss function. For overcoming this problem, 1-NELM expressed as a linear program problematic, in which the initial problematic converted into a new problematic using differentiable *constraint*. This approach leads to fast execution and simple estimation with better generalization performance for classification problem.

For a given L number of *hidden nodes*, the ELM decides the unknown output weight matrix $\Omega = \{\Omega_1, \Omega_2, \ldots\ldots\ldots\ldots \Omega_L\}^t$ vector and these weights should have smallest norm training error concept [145], this is known as minimum norm least squares problematic of Ω and denoted as:

$$min||H\Omega - T||_2 \text{ and } min||\Omega||_2 \quad (4.1)$$

MNLS problem of equation (3.1) converted into 1-norm by using equation (3.2)

$$min_{\Omega \in R^L} ||\Omega||_1 + \gamma ||H\Omega - T||_1 \quad (4.2)$$

Where $\gamma > 0$ is constant and denotes regularization parameter. 1-NELM problematic (3.2) can be articulated into a linear program problematic using the method of [66],[154] assume that $\lambda, \beta \in R^L$ and $p, q \in R^m$

Suppose that $\Omega = \lambda - \beta$ and $H\Omega - T = p - q$ (4.3)

Such that $\lambda, \beta \geq 0$ and $p, q \geq 0$ hold. Then, using (3.2) in (3.3) can be expressed into $1-Norm$ linear programming using primitive form of formula as $min_{\lambda,\beta,p,q} E_L^t (\lambda + \beta) + \gamma \, min_{U,V,P,Q} E_m^t (p + q)$

Condition should be $H(\lambda - \beta) - p + q = T$, such that $\lambda, \beta, p, q \geq 0$, (4.4)

Where E_L and E_M denotes the matrix as $L \times M$ size. Hence, equation (4.4) shows that the prolem of single degree (linear) has unique solution, conceivable and further the equation (4.4) is constrained and differentiable.

4.5 Box Plot

The boxplot is a graphic approach representing five numerical statistics of a sample group in its conventional form in order to visualize its scattering and skewness [155].

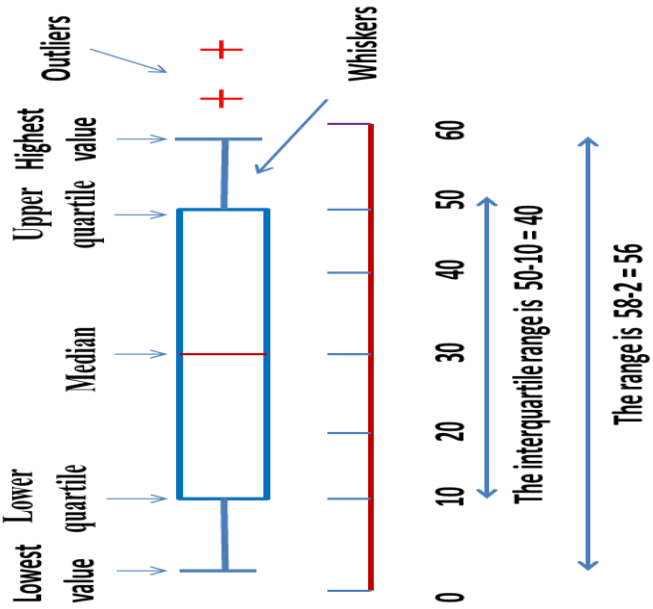


Figure 4.2. A brief labeling with interquartile range, outliers median and structure of a box plot

All descriptions are concentrated on the median and refer to the smaller measurement, a first half part of the sample median (1st quartile,Q1), the median (2nd quartile, Q2), the 2nd half of the sample median (3rd quartile, Q3), and the largest assessment. The range between the 1st and 3rd quartiles is defined as the interquartile range (IQR) and offers an indicator of data distribution (IQR = Q3-Q1). The IQR refers visually to the display's only box and it occupies about 50 per cent of findings closest to the mean. The highest and the lowest findings are those which are

beyond the lines (or whiskers) that link the IQR to a smallest or largest non-outlier value (e.g., within$1.5\ IQR$). In addition, the conventional boxplot may also have a probability distribution of approximately 95 percent around the median. The labeling and structure of a box plot is shown in Figure 4.2.

4.6 Features Dimension reduction by Generalized Discriminant Analysis

Dimension Reduction relates to the procedure of translating a collection of large-scale data into smaller-dimensional data ensuring it's eloquently conveys similar information. Usually these methods are used to solve ML problems in order to achieve better attributes for a classification, detection and prediction of CDs. The boxplot of nine features of NSR-CAD and NSR-CHF are shown in Fig. 4.3 (a, b) which reveals that the some considered attributes conveys similar information in the form of median IQ value. In these cases, differentiating between the two classes (group) is a challenge. An attributes dimension transformation technique would be very useful in these cases [140]. Various methodologies have been applied to minimize the classification size of data attributes [141]. In this chapter, Generalized Discriminant Analysis (GDA) has been employed for dimension reduction of attributes. An advantage of this method is that space reduction of the input attributes dimension and discriminating of features is based on nonlinear kernel function [142].

The formulation of GDA has been explained as: assume that given training sample set X has α attributes vectors out of β classes' level. Let us assume that X_{ij} signifies the j^{th} HRV attributes vector in the i^{th} class β_i is the class size of the i^{th} class and ψ is a non-linear mapping function to facilitate the features data of X which is plotted into upper dimensional attributes space. $\psi: X_k \in R^f \rightarrow \psi(X_k) \in R^F,\ F \gg f$. The interpretations of $\psi(X_k)$ are assumed to be concentrated in space F. Before defining the projection of the input training sample set X into a new set Y by worth of the GDA, two matrices of the within class scatter matrix ω and the between class scatter matrix ρ in space F are formulated as

$$\omega = \frac{1}{\alpha}\sum_{i=1}^{\beta}\sum_{j=1}^{\beta_i}\psi(X_{ij})\psi^t(X_{ij}) \tag{4.5}$$

$$\rho = \frac{1}{\alpha}\sum_{i=1}^{\beta}\beta_i\left(\sum_{j=1}^{\beta_i}\psi(X_{ij})\left(\sum_{r=1}^{\beta_i}\psi(X_{ir})\right)^t\right) \tag{4.6}$$

The GDA's main purpose is to find the projections matrix ∂ so that it can maximize inter-class attrition and reduce intra-class attrition in space F, which is similar to executing the complex Fisher criteria maximizing issue.

$$\partial = \arg\max_{\partial}\left(\frac{\omega^t \rho_{\partial}}{\omega^t \omega_{\partial}}\right) \quad (4.7)$$

The projection vector ∂ is the eigenvector of the matrix $\omega^{-1}\rho$ related with the eigenvalue $\lambda = \omega^t \rho_{\partial} / \omega^t \omega_{\partial}$. All solutions of the ∂ lie in the range of (X). As a result, there exist extension coefficients μ_k such that

$$\partial = \sum_{k=1}^{\beta} \mu_k \, \psi(X_k) \quad (4.8)$$

Assume that the kernel function is represented as $K(X_k, X_l) = K_{kl} = \psi(X_k)\psi(X_l)$ and acting the Eigen matrix disintegration on the kernel vector as $ED = (K_{kl})_{k,l=1,2,\ldots\ldots,;\alpha,}$ hence, α normalized extension coefficients for each projection vector $\mu = \mu/\sqrt{\mu^t K_{\mu}}$ is calculated. Now, for an attribute matrix X from the test HRV data set, the projection on the k^{th} Eigenvector ∂_k can be computed as

$$Y^k = \partial^{k^t} \psi(X) = \sum_{j=1}^{\alpha} \partial_j^k \psi(X_j) \, \psi(X) = \sum_{j=1}^{\alpha} \partial_j^k \, K(X_j, X) \quad (4.9)$$

Where ∂_j^k indicates the j^{th} extension factor of the i^{th} Eigenvector. For the measurement of the reductions of the function element, β-1 Eigenvectors vectors connected to the first highest non-zero β-1 Eigenvalues values are assigned to form the mapping function $(TM)^t = [\partial^1, \partial^2, \ldots, \partial^{\beta-1}]$. Therefore, every one HRV feature vector is projected into a new coordinates using the $\beta - 1$ projection vectors [142]. It implies also that optimal solution of Eigenvectors for the data is transformed is usually equal to β-1. In this chapter, assuming that the class level is 2 i.e. binary class thus GDA transform the dimension of nine attribute to a new attribute.

4.7 Nonlinear Features

4.7.1 *Correlation Dimension*

The *correlation dimension* (CD) is a measurement of the fractal *dimension* of the space taken up by a set of possible points. For M sample points $X(1), X(2), \ldots, X(M)$, with an embedding space (m) $\in R^m$, the CD is defined as

$$CD = \lim_{r \to 0} \frac{\log (C(r))}{\log (r)} \quad (4.10)$$

A $correlation\ function\ C(r)$ is represented as

$C(r) = \frac{1}{M^2} \times (\ Number\ of\ pairs\ of\ (i,j)\ with\ S(i,j) < r)$. Where $S(i,j) = |X_i - X_j|$ in millisecond and $1 \leq i < j \leq M$ [156]. In this work, $m = 10$, $filtering\ level\ (r) = \{0.0050 \ldots 0.1\}$ are taken for the HRV sample analysis. The step to calculate of r value using data extracted from random walker.

4.7.2 *Detrended Fluctuation Analysis*

Detrended fluctuation analysis (DFA) is an alteration of root − mean − square (RMS) and based on a random walk which is implemented to assess the chaos of HRV signal. The RMS fluctuation of combined and detrended of N-N time series is called DFA. It is defined as

$$F(n) = \sqrt{\frac{1}{M} \sum_{k=1}^{M} [Y(k) - Y_n(k)]^2} \quad (4.11)$$

Where $Y_n(k)$ is the k^{th} value of integrated $N - N$ time series data of total length k [117]. The α represents the slope of the line relating $\log F(n)$ to $\log (n)$. Often two particular linear regions on the $\log - \log$ plot are utilized to represent the short term α_1, for this $n_1 = 4$, calculated $F(n_1)$ and the long term scaling α_2, for this $n_2 = 300$ calculated $F(n_2)$.

4.7.3 *Sample Entropy*

The application of Sample Entropy (SampEn) for HRV analysis has been discussed to detect cardiac disease in [128], [157]. For an HRV time series $X(1), X(2), \ldots, X(M)$, assume $X_m(i)$ denotes the m points X_i, X_{i+1}, ------,,X_{i+m-1} which is known as a template and can be reflected a matrix of length m. An example where all the components of the matrix $X_m(j)$ are within a distance r of $X_m(i)$ is known as a template match. Let D_i indicates the number of template matches with $X_m(i)$ and C_i indicates the number of template matches with $X_{m+1}(i)$. The number $P_i = \frac{C_i}{D_i}$ is an estimate of the conditional probability that the point X_{j+m} is within r of

X_{i+m-1} given that $X_m(j)$ matches $X_m(i)$. The SampEn is defined for m points.

$$SampEn(m,r,M) = -\log(\sum_{i=1}^{M-m} C_i / \sum_{i=1}^{M-m} D_i) \quad (4.12)$$

Where m is an integer number and r is a positive real number. In this work, standard value m=2, r= 0.2×std.deviation (NN) were used for HRV analysis.

4.7.4 *Poincare Plot as SD1/SD2 Ratio*

The Poincare plot investigation is a non-linear and graphical method to assess the chaos of NN interval (HRV). This plot scatters the graph of present NN intervals and previous NN interval. Two neighboring NN intervals signify a single point in the scattered plot. The 1st NN interval (NN_i) denotes the X-abscissa, the 2nd interval as NN_{i+1} denotes Y-ordinate. Standard deviation one (SD_1) for short NN interval and standard deviation two (SD_2) for long NN intervals were calculated for chaos analysis of HRV signal. For calculation of these intervals we have taken 150 samples for short and 350 samples for long out of 500 NN intervals. The value of $SD1/SD2$ may likewise be listed to describe the relationship between short and long terms [129].

4.7.5 *Hurst Exponent*

Hurst exponent (HE) is a method of measuring the smoothness of the data from a fractal time series. It is based on the asymptotic execution of the rescaled observation range.

$$HE = \frac{\log(R(n)/S(n))}{\log(T)} \quad (4.13)$$

Here n is the analysis time period (like $n = M, M/2, M/4 \ldots\ldots$ where 2,4,8) which is referred to as blocks size, R(n) is the range of the 1st n values and S(n) is their standard deviation (S.D). R(n)/S(n) is the respective rescaled range value, and T is the population sample length. M is observations of a time series [61].

4.7.6 *Permutation Entropy*

Permutation entropy (PE) is a suitable test for exposing the non-linearity of HRV time series. Due to its easy concept, the PE has also been studied most extensively of physiological signals, based on fewer features, computationally and fairly immune to interference and artifacts. For the data series $\{X(i), i = 1,2,3 \ldots\ldots M\}$, the PE can be represented as the entropy for both the sequences and ordering of the K symbols (O) [122].

$$H = -\sum_{L=1}^{K} \pi_L \ln(\pi_L) \quad (4.14)$$

Where $\pi_1 \pi_2 \ldots\ldots \pi_K$, signify the probability distribution of every symbol series [158] and $\sum_{L=1}^{K} \pi_K$ =1. In this work, standard value $O = 4$ and time delay of 1 are chosen for HRV analysis.

4.7.7 *Improved Multiscale Permutation Entropy*

The $multiscale\ permutation\ entropy$ (MPE) method cannot provide a consistent investigation for short HRV time series. An improved MPE $(IMPE)$ has been presented to overcome this problem [159] to define a real-valued time series $\{X(i), i = 1, 2, 3 \ldots\ldots M\}$ of length M for $IPEM$ formulation with such an embedding space $\{(m) \in R^m\}$. Many subsequent coarse-grained versions [160] are set by averaging the sets of data series within equally spaced extending period S frames, which is termed the scale-factor. Therefore, in this process chosen separate time series as $\eta_i^S | (i = 1,2 \ldots\ldots S)$. For a given scale factor S and embedding $dimension\ m$, there is a different calculation of the $probability\ entropy$ of each of $\eta_i^S | (i = 1,2 \ldots\ldots S)$, this then measures the average values of probability entropy. It is defined as

$$IMPE\ (X, S, m) = \frac{1}{S}\sum_{i=1}^{S} PE(\eta_i^S) \quad (4.15)$$

For HRV analysis, standard value $m = 4, scale\ (S) = 4$ and time delay of 1 have been used.

4.7.8 $Cumulative\ Bi - Correlation$

The $bi - correlation$ is the joint moment of three $NN(i);\ NN(i + \tau)\ and\ NN(i + 2\tau)$ samples and defined as $E[X(i);\ X(i + \tau);\ X(i + \tau_0)]$, where E denotes expected value of three samples, τ and τ_0 delays between samples. We have taken $\tau_0 = 2\tau$ so that the

measurement of joint moment can be function of τ [132] The bi-correlation is not commonly known for the calculation of HRV signal disorder, but the cumulative bi-correlation (CBC) was introduced for the non-linearity test, with the constraint that the second delay (τ_0) is double the first delay (τ) [161] . For the definite set of delays the CBC is a composite function of bi-correlation and for each interval, the CBC is the summation of the bi-correlation's cumulative values up to the interval delay τ. The delay of 1 is chosen for NN analysis.

4.8 Parameters Used for OS-ELM

The detection and classification of CAD and CHF has been performed by OS-ELM using Sigmoid activation function and Multiquadric functional additive node. The multiquadric function was considered for RBF hidden nodes. The sigmoid function and Multiquadric function are defined as $\emptyset(\mathrm{A}, \mathrm{B}, \mathrm{X}) = 1/1 + exp\ -(A^t X + B)$ and $\emptyset(\mathrm{A}, \mathrm{B}, \mathrm{X}) = \sqrt{(||X - A||_2^2 + B^2}$. The optimal values of user-specified parameters like regularization parameter γ for 1-NLPELM were selected using 10-fold cross-validation method. Analyzing classifier output for datasets, following performance parameters calculated:

$$\text{Accuracy (AC)} = (\mathrm{TP} + \mathrm{TN})/(\mathrm{TP} + \mathrm{TN} + \mathrm{FN} + \mathrm{FP}) \quad (4.16)$$

$$\text{Sensitivity (SE) or true positive rate (TPR)} = \mathrm{TP}/(\mathrm{TP} + \mathrm{FN}) \quad (4.17)$$

$$\text{Specificity (SP) or true negative rate (TNR)} = \mathrm{TN}/(\mathrm{FP} + \mathrm{TN}) \quad (4.18)$$

$$\text{False positive rate (FPR)} = (1 - \mathrm{SP}) \text{ using confusion matrix.} \quad (4.19)$$

Where $TP = true\ positive$, $FN = false\ negative$, $TN = true\ negative$ and $FP = false\ positive$.

4.9 Results

To understand the reduction capability of GDA using RBF kernel function, the box-plots of the chaos attribute for NSR-CAD and NSR-CHF data sets are shown in Fig. 4.3 (a,b) and after reduction of chaos features by GDA is shown Fig. 4.4 (a, b). The box-plot patterns of nine chaos features associated with data sets are positioned close to each other (median value indicated by red line in box) before reduction of attributes. After attribute space reduction by GDA method, a new attribute is generated which is well separated within the attribute dimension. So, the new feature provides not only improvements in the detection capability but also make a suitable tool for good discrimination of NSR-CAD and NSR-CHF

data set. For this reason, GDA reduction scheme is an appropriate tool for better discrimination of NSR-CAD and NSR-CHF dataset.

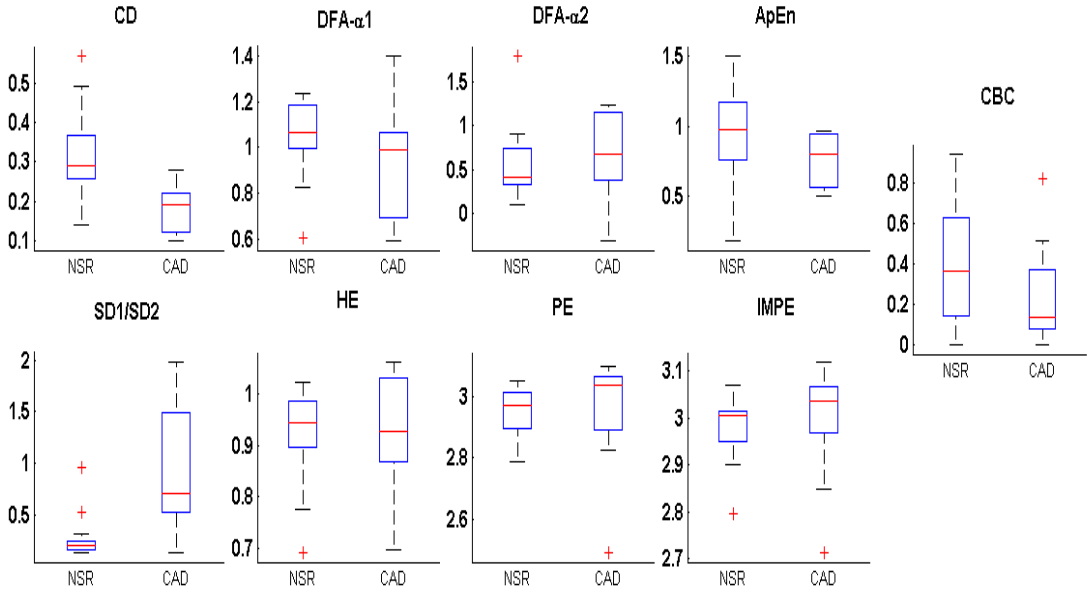


Figure 4.3 (a): Box plot of nine chaos attributes for NSR-CAD data set before reduction by GDA.

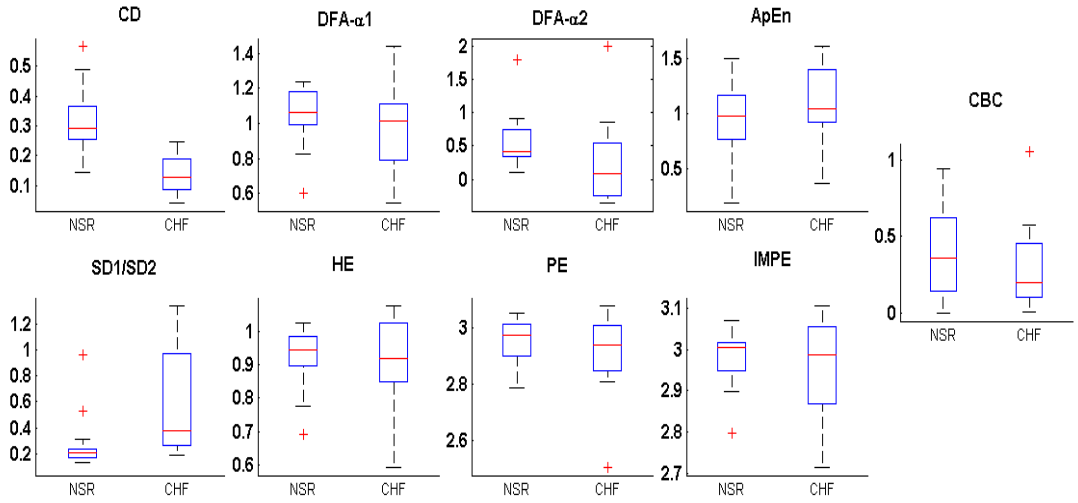


Figure 4.3 (b): Box plot of nine chaos attributes for NSR-CHF data set before reduction by GDA

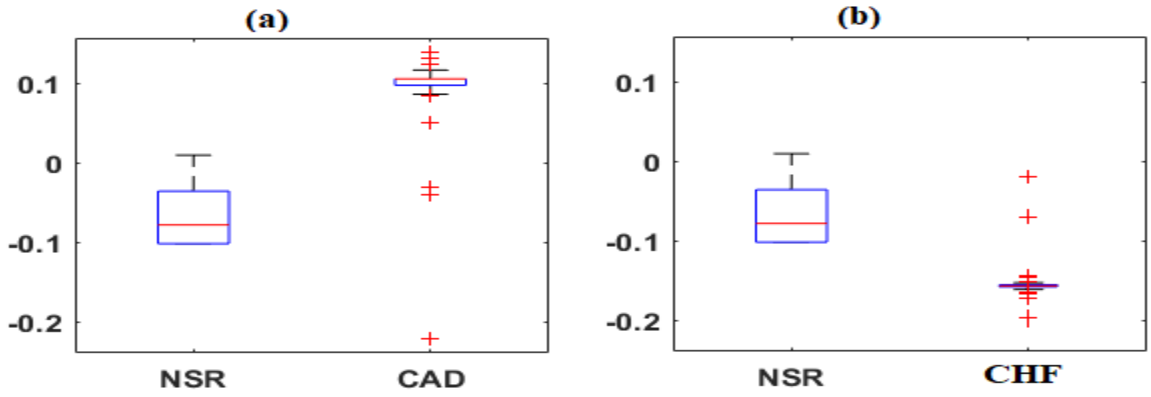
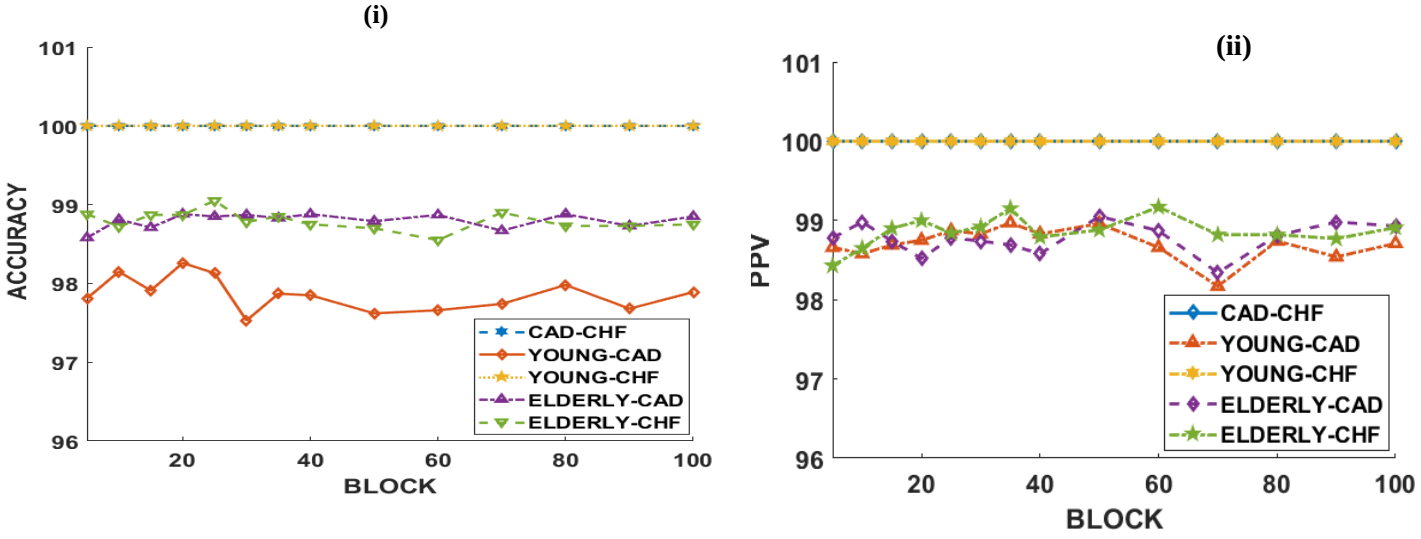


Figure 4.4. Box plot, after attributes reduction by GDA with RBF kernel function for (a) NSR-CAD datasets (b) NSR-CHF datasets

The variation of classification parameters likePPV, SEN and SP with respect to block size (5, 10, 15----- 100) for OSELM and GDA as feature reduction scheme with different kernel and activation functions are shown in Fig.4.5 For evaluation of each parameter, mean of 10 cross validation with 100 times randomly subjects are selected from datasets for each block. To understand the effect of varying block size on a parameter, following data set combination of $CAD - CHF$, $NSR(YNG) - CAD$, $NSR\ (YNG) - CHF$, and $NSR\ (ELDERLY) - CHF$ were used. The parameters: AC, PPV, SEN and SE of different data sets are plotted for OSELM classifier with Hardlim activation function and GDA with Gaussian and RBF kernel as feature reduction scheme. It is observed that even after increasing the block size, there is no significant change in value of classification parameters. The simulation graph reveal that less than 40 blocks are appropriate for detection and classification of CHF and CAD subjects. So, the block size is chosen in such a way that optimum value of testing and validation parameters is achieved. The parameters for different activations function and kernel are tabulated in Table 4.1.



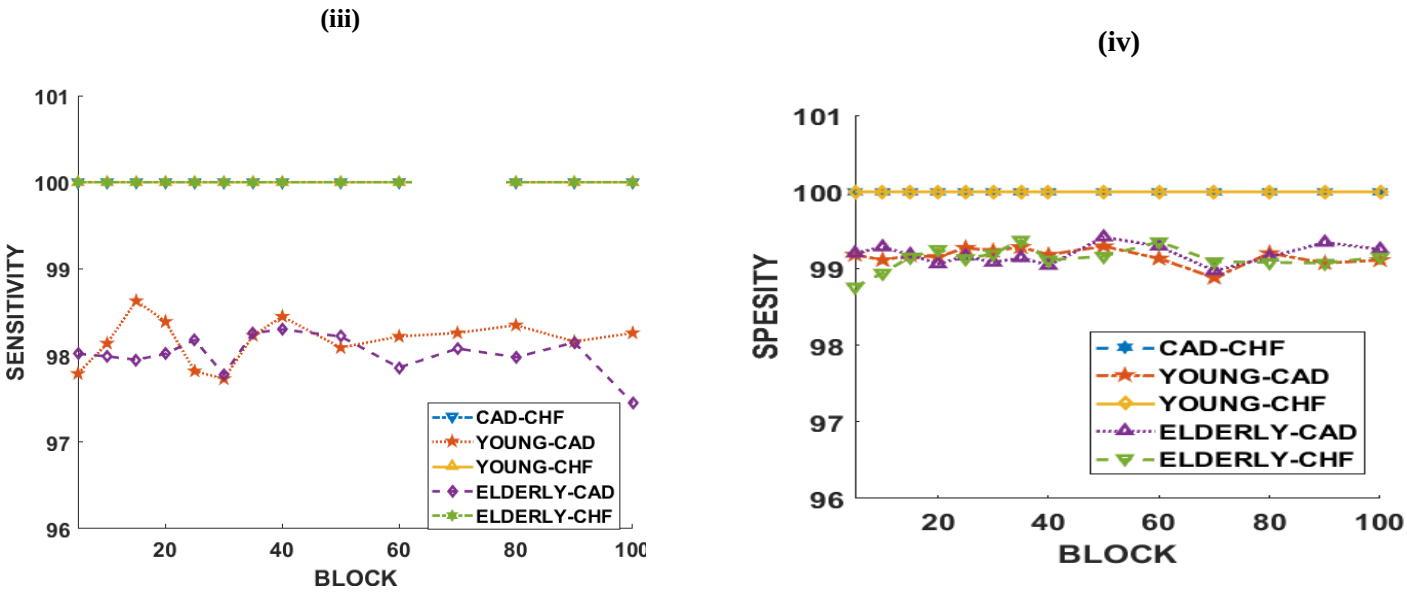


Fig. 4.5 Variation of classification parameters Vs. Block size when using OSELM with Hardlim activation and GDA with Gaussian Kernel for parameters (i) Accuracy (ii) PPV (iii) Sensitivity (iv) Specificity

In order to determine the classification parameters and validate the performance of proposed method in terms of AC, PPV, SE and SP, simulation analysis has been conducted on the data set of various subjects. 50 samples for training and 62 samples for validation have been taken for $CAD-CHF$, 84 and 47 have been taken for $NSR-CAD$, 80 and 60 for $NSR-CHF$, 80 and 52 $NSR-CAD$ and 80 and 60 $NSR-CHF$ data sets have been used for training and validation purpose respectively. These standards databases have been obtained from Physionet Website.

The performance parameters achieved by OSELM and OSELM+GDA are listed in Table 4.1. The numerical simulation results shows that $OSELM$ with $Hardlim$ Activation function achieved an AC, PPV, SE and SP of 81.83, 79.12, 76.6 and 87.08 respectively for $NSR-CAD$ datasets while OSELM with Sine Activation function attained 90.94, 91.17,86.72 and 94.42 respectively. When GDA with kernel function combine with OS-ELM, the classification performance has significantly increased. For instance, GDA with RBF Kernel +OS-ELM with Sine activation function achieved classification performances 98.34, 99, 97.38 and 99.32 respectively for NSR-CAD group. GDA with **Gaussian** Kernel + OS-ELM with Sine activation function produced 100% classification performances for NSR-CHF dataset. This result reveals that proposed method achieved better results as compared to OSELM +various type of activation function.

Table 4.1: Represents results achieved by OSELM and OSELM combine withGDA. The best classification parameters are signposted by boldface in the table.

Data sets (Training and validation size)	Training				Validation			
	AC	PPV	SEN	SE	AC	PPV	SE	SP
OS − ELM with Hardlim Activation Function								
NSR − CAD (84 × 1, 47 × 1)	100	100	100	100	81.83	79.12	76.6	87.08
NSR − CHF (80 × 1, 60 × 1)	100	100	100	100	77.07	76.17	69.84	82.65
OS − ELM with Sine Activation Function								
NSR − CAD (84 × 1, 47 × 1)	100	100	100	100	90.94	91.17	86.72	94.42
NSR − CHF (80 × 1, 60 × 1)	100	100	100	100	81.53	81.22	75.35	87.06
OS − ELM with Sigmoid Activation Function								
NSR − CAD (84 × 1, 47 × 1)	100	100	100	100	85.11	83.23	82.32	89.26
NSR − CHF (80 × 1, 60 × 1)	100	100	100	100	79.2	80.2	71.98	86.78
GDA with Gaussian Kernel + OS − ELM with Sigmoid Activation Function								
NSR − CAD (84 × 1, 47 × 1)	100	100	100	100	98.26	98.08	97.88	98.81
NSR − CHF (80 × 1, 60 × 1)	100	100	100	100	**100**	**100**	**100**	**100**
GDA with Gaussian Kernel + OS − ELM with Sine Activation Function								
NSR − CAD (84 × 1, 47 × 1)	100	100	100	100	98.23	98.7	97.22	99.17
NSR − CHF (80 × 1, 60 × 1)	100	100	100	100	**100**	**100**	**100**	**100**
GDA with RBF Kernel + OS − ELM with Sine Activation Function								
NSR − CAD (84 × 1, 47 × 1)	100	100	100	100	**99.34**	**99**	**97.38**	**99.32**
NSR − CHF (80 × 1, 60 × 1)	100	100	100	100	97.87	97.41	97.99	97.96

4.10 Discussion

Table 4.2 revealed that the cardiac heart diseases like CAD and CHF have also been proven to become one of the leading causes of death, which is the major reason why precise heart failure (HF) risk prediction is vitally important for preventing and treating it [162]. A quick and efficient CAD and CHF diagnosis is essential for evading a life-threatening event [163]. In addition to timeliness, AC, PPV, SE and SP play an incredibly important role in the medical sphere, because it is linked to a person's life. There has been done a terrific and substantial number of works on the diagnosis and detection of cardiac diseases using ML techniques. In contrast, there is yet to be a gap regarding comptability of methodology or classification algorithm suitable with datasets. However, it is demonstrated that the methodology of selection and reductions of features increased the performance and accuracy of classification models. In addition, it is also evident that classifier ensembles have demonstrated to have enhanced classification AC [164]. There are some factors that could not be ignored in health - related issues, such as time taken to execute the procedure and technical challenge that depends on the number of functions, precision and sweeping generalization. Several researchers have focused on methodology to avoid the integration of unnecessary traits, as per assessment. Using state-of-the-art methods such as genetic algorithm (GA) in [165], [166] and [167], LS-SVM in [168], information gain in [169], $F-score$ in [164] for Attributes selection has helped to counteract the question of so many other attributes and thus decrease the processing time and the size of the computing. When an insight assessment of the articles was conducted, it was realized that certain researches have lessened the dimensionality utilizing advanced detection

methods. For instance, Principal components analysis (PCA) in [168], [170], [171] chaos features in [172], and minimal Max significance Redundancy in [173] are incredibly effective in reducing dimensions. Many research focused on obtaining the maximum accuracies through implementing methods that are exceptionally efficient and consistent, for example, GA with SVM to accomplish more than 98 % AC in [165] PCA and SVM to achieve …

Table 4.2: Summary of existing methods applied for CAD and CHF disease detection and classification.

Reference .	Detection of subjects	Proposed Scheme	Database	Accuracy (in %)
[165]	CHF	GA with SVM	CHF2DB and NSR	98.79
[168]	CAD.	PCA and SVM	HRV	99.2
[170]	CAD.	GA and Neural network	Z-Alizadeh Sani Database	93.85
[174]	CHF and CAD.	DT	A dataset of 2346 Patients' records	94
[175]	Paroxysmal Atrial Fibrillation (AF)	SVM	AF Database	87.7
[176]	CHF risk factor.	Rules-based classifier.	Cleveland	86.7
[177]	CAD	Rules and fuzzy experts approach.	Advanced Medical Research Institute	84.2
[163]	CHF	Ensemble of bagged DT	PhysioNet database	98.8
[178]	sudden cardiac deaths (SCD)	t-test for feature ranking and	MIT-BIH and NSR	94.7
[179]	CHF and CAD	Multilayer Perceptron NN.	NSR and Long-term ST.	97.83
[173]	CHF	DTC wavelets transform.	MIT − BIH NSR, , and BIDMC CHF	99.86
[180]	Chronic heart failure diagnosis.	Least-Square SVM	A dataset of 152 heart sound	95.39
[181]	HF patients' mortality.	Random Forest.	Intelligent Monitoring in ICU	82.1
[182]	Heart failure risk	Nural Network	-	66.55
[183]	CAD	Random Forest, C5.0, and fuzzy modeling.	UCI	90.50
[184]	Heart disease (CAD,CHF,MI)	Fuzzy rules − based	Personal Health Record	69.22
	This paper CHF and CAD	with Kernel and with	PhysioNet database	100 for CHF and 98.34 CAD

more than 99 % AC in [168] , PCA to gain more than overall AC in [171], and a set of bagged decision trees (DT) to obtain approximately 99%. AC in [163], some of those methods and their prediction accuracy are shown in Table 4.2 What's troubling is that most of the methods were applied to small database and there's still a desperate need to test them on a broad dataset. The most research haven't addressed whether attributes reduction affects the knowledge of systems or not.

4.11 Summary

In this part, a new approach has been discussed for classification of coronary heart diseases such as CAD and CHF. For this, nine nonlinear features have been extracted from HRV signal. In order to enhance the detection capability of proposed model, the dimension of nine features were reduced using GDA technique. The results of the comparative study have indicated that the GDA with OSELM binary classifier outperforms as compared to other existing techniques. The GDA with Gaussian Kernel + OS-ELM with Sine activation function achieved classification performances of 100% for NSR-CHF datasets. The GDA with RBF Kernel + OS-ELM with Sine Activation Function attained an AC of 99.34%, PPV of 99% and SP of 99.32 % for NSR-CAD. Hence, this proposed scheme can be employed for online coronary diseases monitoring and diagnosis of CAD and CHF subjects.

5. Detection of CAD Using Ranked Nonlinear Features and 1-NELM

5.1 Overview

This part focuses on detection and classification of CAD using eleven different nonlinear and chaos features were extracted from pre-processed and fragmented HRV database. The chaos and nonlinear features namely correlation dimension (CD), Poincare plot such as SD1/SD2 ratio, multivariate largest Lyapunov exponent (MLLE), Hurst exponent (HE), Lempel-Ziv (LZ), sample entropy (SampEn), dispersion entropy (DispEn), Improved permutation entropy (IPE), adaptive multiscale PE (AMPE), multifractal detrended fluctuation analysis (MFDFA) and cumulative bi-correlation (CBC) were employed to input of ML algoritham. **Fig. 3.5** shows a proposed adaptive ML model for CAD detection. All the features are not sensitive for escalation in interpretation and comprehension of healthy and CAD database. Therefore, the features were ranked using Fisher score, Bhattacharya space, Wilcoxon, receiver operating characteristics (ROC) and entropy ranking methods. The most important top seven ranked features were applied to features space transformation technique such as generalized discriminant analysis (GDA). The GDA transfer top seven features to a new feature. The value of a new feature was first regularized in range of -1 to 1, after this, it is applied to 1-norm extreme learning machine (1-NELM) classisfier. The simulated outcomes have shown that GDA and 1-NELM classifier with considered ranking methods achieved better accuracy, generalization performance and less computational validation time as compared to SVM and PNN. It has also been observed that the CD, MFDFA MLLE, , AMPE, DispEn, LZ and $SD1/SD2$ features have lowest p-value (< 0.001)compared to all remaining features.

5.2 Database and Pre-Processing

5.2.1 *Database of Heart Rate variability*

The HRV standard database of normal young (YNG) and elderly (ELY) subjects were acquired from Fantasia database. This databse contains number of 20 YNG (22 - 35 years) and 20 ELY (63 - 86 years) subjects [24]. The Coronary artery Disease database were obtained from St. PIC Cardiological Technics [25]. It has 72 annotated ECG records extracted from 30 Holter ECG templets. Length of each ECG record is one-hour long. Only thirteen subjects (9 Male and 4 Female, aged 19-78) were suffered from Coronary artery Disease [25].

5.2.2 *Pre-Processing of Heart Rate variability*

The HRV time series obtained from the standard database usually contains ectopic beats (irregular impulse formation in the heart muscle) and non-stationary, which precludes effective feature extraction and HRV data analysis. To avoid this problem, the pre-processing of the HRV signal is required. In the pre-processing, first of all non-stationary and ectopic beat using a nonlinear filter and wavelet-based

technique has been detached [26] [27]. A cubic spline interpolation method was employed to supplant ectopic beats intervals. After pre-processing, original HRV is known as a normal-normal interval (NN interval).

Initially, 40,000 samples of NN interval of each database of subjects have been selected. To increase the lenth of training and testing datasets dimension, 40, 000 samples of NN interval were alienated into fragments of 8,000 NN intervals. Finally, the nonlinear and chaos features were extracted from fragmented NN intervals.

5.3 FEATURES

5.3.1 *Correlation Dimension*

The correlation dimension (CD) is a degree of dimensional space occupied by a number of random points and its measurement is based upon the principle of chaos concept. For M data points $X(1), X(2), \ldots, X(M)$, having an embedding space $(m) \in R^m$, the CD is represented as

$$CD = \lim_{r \to 0} \frac{\log (C(r))}{\log (r)} \tag{5.1}$$

The correlation $C(r)$ is denoted as $C(r) = \frac{1}{M^2} \times (Number\ of\ pairs\ of\ (i,j)\ with\ S(i,j) < r)$. Where $S(i,j) = |X_i - X_j|$ in millisecond and $1 \leq i < j \leq M$ [156]. In this work, $m = 10$, filtering value $(r) = \{0.0050 \ldots 0.1\}$ are chosen for the HRV database analysis using random walker steps.

5.3.2 *Multifractal Detrended Fluctuation Analysis*

The multifractal detrended fluctuation analysis (MFDFA) is known to be an effective approach for calculating whether chaos characteristics exist in HRV data [185], [186]. For a specified length (M) of NN time series data of variable $X(P)$ were calculated over $1 \leq P \leq M$ for analysis of CAD. For mathematical expression, fisrt calculated the profile time series using function as $Z(i) = \sum_p^i (X(P) - \epsilon)$ for 1 ≤ i ≤ M, where ϵ is mean of NN time series data. After this, divided the NN data into a chunk dimension d, then $Z(i)$ was splited into N blocks holding a data segment of d sapmle of NN time serise. The MFDFA for multifractal parameter (l) was calculated for each d by the following equation

$$MFDFA(l,d) = \sqrt[l]{\frac{1}{N} (\sum_K^N \rho_d^2\ (k))^{\frac{l}{2}}} \tag{5.2}$$

Where $N = integer(M/d)$ and the standard deviation $(\rho_d(k))$ was calculated in each chunk k for $1 \leq k \leq N$. Accoprding to authures of [186] mutifractal parametes $(l < 0)$ are suitable for smaller amplitude time seriese data., In this work, $l = -4$, $d = 40$ and 5000 beats (HRV sample) were employed for analysis of data sets .

5.3.3 *Sample Entropy*

The Sample Entropy (SampEn) is an entropy family used to estimate the chaos or irregularity of NN time-series database of subjects. The SampEn is fundamentally the same as the Approximate Entropy (ApEn), however, in ApEn, the comparison between the template vector and whatever is left of the vectors likewise incorporates correlation with itself. This ensures probabilities are not ever zero. Thus, it is constantly conceivable to take a logarithm of probabilities. Since template correlations with itself bring down ApEn values [34], the signals are translated to be more general than they are really. The use of SampEn for HRV analysis has been described in [128]. For an HRV time-series $X(1), X(2), \ldots, X(M)$, assume $X_m(i)$ denotes the m points X_i, X_{i+1},------,X_{i+m-1} which is known as a template and can be reflected a vector of length m. An example where all the components of the vector $X_m(j)$ are within a distance r of $X_m(i)$ is known as a template match. Let D_i indicates the number of template matches with $X_m(i)$ and C_i denote the number of template matches with $X_{m+1}(i)$. The number $P_i = \frac{C_i}{D_i}$ is estimate of the conditional probability that the point X_{j+m} is within r of X_{i+m-1} given that $X_m(j)$ matches $X_m(i)$. The $SampEn$is defined for m points is defined by following equation

$$SampEn(m, r, M) = -\log\left(\sum_{i=1}^{M-m} C_i / \sum_{i=1}^{M-m} D_i\right) \quad (5.3)$$

Where m = positive integer and $r \in R^+$. In this work, standard value $m = 2, r = 0.2 \times std.deviation\ of\ (NN)$ were chosen for HRV analysis.

5.3.4 $Lempel - Ziv\ Complexity$

The Lempel − Ziv (LZ) complexity measures the complexity of finite length (L) of the HRV signal. The sample of NN signal represented as $\{X(1), X(2) \ldots\ldots\ldots\ldots X(i)\}$.Where $i = 1,2,3 \ldots\ldots\ldots\ldots L$. In the complexity measure, one needs to transform thesample of HRV signal into a finite symbol series. In this transformation, HRV signal samples are converted into a binary sequence $B = S(1), S(2), \ldots\ldots\ldots\ldots, S(i)$ using threshold (Th). It is defined as

$$S(i) = \begin{cases} 0, & X(i) < Th \\ 1, & X(i) \geq Th \end{cases} \quad (5.4)$$

The Th value was computed using median value of HRV signal, owing to its robustness to outliers [187], [188]. In order to obtain complexity, first, the sequence B is scanned from left to right and the complexity counter $(C(n))$ is added by one unit every time. During this process a new subsequence of successive characters is found. For a metric free of the subsequence length, $C(n)$ must be then normalized. It has been shown by [189] that the upper bound of C (n) should be $C(n) < \frac{n}{(1-\delta_n)\log_\mu(n)}$. Where δ_n is small quantity and $\delta_n \to 0$ when n $\to \infty$. Base (μ) of the logarithm represents symbol set. Then,

$\lim_{n\to\infty} C(n) = P(n) = \frac{n}{log_{\mu}(n)}$. For a binary sequence (0 and 1), $\mu = 2$, therefore $P(n) = \frac{n}{log_{\mu}(n)}$ and $C(n)$ can be normalized by $P(n)$

$$NC\ (n)\ = \frac{C(n)}{P(n)} \tag{5.5}$$

Where $NC\ (n)$ is normalized LZ complexity, it indicates the arising rate of new assortments in the sequence of NN time series.

5.3.5 *Dispersion Entropy:*

Dispersion entropy (DispEn) also measures the complexity of HRV time series [190],[191]. For the formulation of DispEn, First, HRV signal is passed through a normal cumulative distribution function (NCDF). It makes available filtered HRV time series X_f, which variesbetween 0 to 1. This technique is envisioned to better treat outliers. Next, X_f is mapped into D classes (1 to D), by using a function, which is defined as $Z^D(j) = round(D \times X_f(j) + 0.5)$, this function linearly maps the range [0, 1] to [1, D].

Now, let $Y_K(i)$ be the set of successive samples in Z^D from i to $i + K - 1$,that is $Y_K(i) = \{Z^D(\mathrm{i}), Z^D(\mathrm{i}+1), \ldots\ldots\ldots\ldots\ldots\ldots\ldots, Z^D(\mathrm{i}+\mathrm{K}-1)\}$, where $i = 1, 2, 3, \ldots\ldots\ldots\ldots\ldots\ldots N - K + 1$. Each oneelement of vector of $Y_K(i)$ signifies a dispersion sample. The probability of occurrence of every single dispersion sample$Y_K(i)$ in $Z^D(j)$ can be obtained as the number of times the sample $Y_K(i)$ appears on $Z^D(j)$, divided by the total number of samples in $Z^D(j)$, that is, $N - k + 1$ .probability distribution for all possible dispersion samples is $P[Y_K(i)]$. Finally, the $DispEn$ is defined as

$$DispEn(K, D) = -\sum_{i=1}^{D^K} P[Y_K(i)]\ \log(\ P[Y_K(i)]) \tag{5.6}$$

Parameters of DispEn selected as K = 2, and D = 6. NCDF was generated with M indicates mean and SD denote standard deviation of the HRV time series, respectively.

5.3.6 *Poincare Plot*

The Poincare plot investigation is a non-linear and graphical method to assess the chaos of NN interval (HRV). This plot scatters the graph of present NN intervals and previous NN interval. Two neighboring NN intervals signify a single point in the scattered plot. The 1stNN interval (NN_i) denotes the X-abscissa, the 2nd interval as NN_{i+1} denotes Y-ordinate. Standard deviation one (SD_1) for short NN interval and standard deviation two (SD_2) for long NN intervals were calculated for chaos analysis of HRV signal. For calculation of these intervals 3000 samples for short and 5000 samples for long NN interval have been taken. The value of $SD1/SD2$ ratio was calculated to describe the complexity of HRV signal [129].

5.3.7 *Hurst Exponent*

Hurst exponent (HE) is a calculation of the smoothness of data from a fractal NN time series dependent on the exponential behavior of the re-scaled technique. In order to estimate the Exponent of Hurst one needs to calculate the re-scaled range over the observed time period. To do so, a total-length NN time series is split into a number of shorter NN series, and the re-scaled period for each of the shorter time series is measured. It is defined as

$$\mathrm{HE} = \frac{log\ (R(n)/S(n))}{log\ (T)} \tag{5.7}$$

Where n is the length of period of the observations spam (like $n = M, M/2, M/4 \dots \dots$where 2, 4, 8 are known as number of chunks), $R(n)$ is the range of the first $n\,n$ time period and $S(n)$ stands its standard deviation. $\mathrm{R(n)/S(n)}$ ratio indicates rescaled range and T is the length of duration of a NN time seriese. M is observations spam of a NN series [192].

5.3.8 *Improved Permutation Entropy*

Improved permutation entropy *(IPE)* is an efficient appraoch to investigate the change the information of HRV time series. Because of its basic principle like depend on less parameters, computationally efficient towards noise and fairly resilient to artifacts. For the given a NN series $\{X\ (i), i = 1, 2, 3 \dots \dots .. M\}$, the IPE can be defined by using Shannon entropy for a K symbol categorizations and order (O) [122].

$$IPE = -\sum_{L=1}^{K} \pi_L \ln(\pi_L) \tag{5.8}$$

Where $\pi_1 \pi_2 \dots \dots \pi_K$, denotes the probability distribution of each symbol [158] and $\sum_{L=1}^{K} \pi_K$ =1. In this work, $O = 4$ has been taken for NN time series analysis.

5.3.9 *Adaptive Multiscale Permutation Entropy*

The multiscal permutation entropy (MPE) method cannot provide a consistent investigation for short HRV time series. To overwhelmed this problem, an adaptive multiscal permutation entropy $(AMPE)$ has been suggested by [126]. For expression of AMPE, fisrt consider a NN time series database $\{X\ (i), i = 1, 2, 3 \dots \dots .. M\}$ of dimension M, with an embedding space (m) $\in R^m$. Then, the multiple coarse grained NN series developed by [131] combining the data points of scale factor (S) across non-overlapping frames calculated using $z_K^S = \frac{1}{S}\sum_{i=(k-1)S}^{KS} X\ (i)$ for $1 \le K \le int(M/S)$. After this, For S and m, probability entropy (PE) of every one of $\eta_i^S | \ (i = 1,2 \dots \dots \dots S)$ is independently evaluated . The average permutation entropy for the coarse grained NN series that is known as $AMPE$ is calculate by following equation.

$$AMPE\ (X, S, m) = \frac{1}{S}\sum_{i=1}^{S} PE(\eta_i^S) \tag{5.9}$$

In this work, $m = 4$, $S = 4$ and delay of 1 are selected for NN time series.

5.3.10 *Cumulative bi-correlation*

The bi-correlation is the joint moment of three $NN(i)$; $NN(i + \tau)$ and $NN(i + 2\tau)$ samples and defined as $E[X(i); X(i + \tau); X(i + \tau_0)]$, where E denotes expected value of three samples, τ and τ_0 delays between samples. The value of $\tau_0 = 2\,\tau$ has taken so that the measurement of joint moment can be function of τ [132]. The bi-correlation is not commonly used for the calculation of NN variations, but the cumulative bi-correlation (CBC) have been introduced for the non-linearity test of NN interval with the condition that the 1st delay (τ) is half of the 2nd delay (τ_0) [161]. For definite set of delays the CBC is a composite function of $\mathrm{bi-correlation}$ and for each interval, the CBC is the summation of the $\mathrm{bi-correlation's}$ cumulative values up to the interval delay τ.

5.3.11 *Multivariate largest Lyapunov exponent (MLLE)*

The LLE can be used to explain the nonlinear properties of NN time seriese data which is based on chaos concept. It shows the mean rate of divergence of between two losest trajectories of neighboring point of databse. But, the accuracy and precision of this method of measuring was not adequate enough for diagnostic application [193]. To improve these parameters, a new modified approach known as multivariate largest Lyapunov exponent (MLLE) has been suggested by [194]. in this approach, multivariate stage space, prearranged that each point $\mathrm{X_J}$ has a nearest neighbor point $\mathrm{X_{\check{J}}}$, there should be a smallest separation between such two points of data to insure that the these points operate on different tracks. Once the calculation of the L points done, calculated the distance to its closest neighbor by $\mathrm{D_J(L) = \min \left\|X_J - X_{\check{J}}\right\|}$, The MLLE approach defined as

$$MLLE = \frac{1}{\sum L^2}\left\{L \times \frac{1}{H}\sum \ln\left(D_J(L)\right)\right\} \tag{5.10}$$

Whwere H is the lenth of $D_J(L)$.

5.4 Features Dimension Reduction

The higher the number of attributes, the more difficult it is for the learning datasets to be visualized again and focused on. Sometimes, attributes of datasets often become highly correlated, for instance, featurs of ELY-CAD datasets and thus redundant, and also ML becomes over fitted. This is where strategies for the reduction of features dimension play an appropriate role [140]. Various algorithms have been proposed to reduce the dimension of features for improvement of the classification capability of ML [141], [195], [196]. In this paper, the generalized discriminant analysis (GDA) and Kernel principal component analysis (KPCA) have been applied the for reduction of top seven ranked features into one new feature dimension. The benefits of both algorithms that input features space reductions is dimensional and feature

selective are based on dynamical kernel function [142]. The dimesion reduction of GDA is based on linear discriminant analysis (LDA). In GDA, for given input features is decoded to a high dimensional feature by a kernel function , in which each groups will be separated from each other linearly. Subsequently LDA algoritham is employed to map the attributes from maximum to minimum domain space on the basis of eigenvector matrix, where it looks for those vectors which best distinguish between dataset rather than those vectors which better characterize the database [197]. In this paper, top ten ranked features were employed for reduction of dimensionality of features space by GDA and KPCA to new features as shown in Fig. 5.1.

The KPCA is a modification of principle component analysis (PCA) dimension reduction technique in which kernel strategies are utilized. In KPCA, initially linear operations are done in a procreating Hilbert kernel space in which kernel is used [198], [196], but its removes redundant and distinguise features capability is less compared to GDA for ELY-CAD dataset. To comprehend the discernment capability of $KPCA$ and GDA with $kernel$, the box $(whisker)$ −plots of the top ten ranked features for ELY-CAD datasets are used, shown in Fig. 5.1 and after reduction of these features by $KPCA$ shown as in Fig.(5.2.a) ; and by GDA as in Fig. .(5.2.b) The box-plot label (median value) of ten top ranked features associated with data sets are positioned close to each other (median value indicated by the red line in box) before reduction of features. After feature space reduction by GDA and KPCA method, the new feature is well separated in the box plot. So the new feature provides not only improvements in the detection proficiency but also make a seemly apprroach for good discrimination of ELY-CAD dataset. However, as perceived by box plot after dimension reduction by GDA and $KPCA$, the GDA makes a more distinct median value for $new\ features\ compared$ to $KPCA$. Therefore, $GDA\ reduction$ technique is better and an appropriate approach for features dimension reduction for CAD detection.

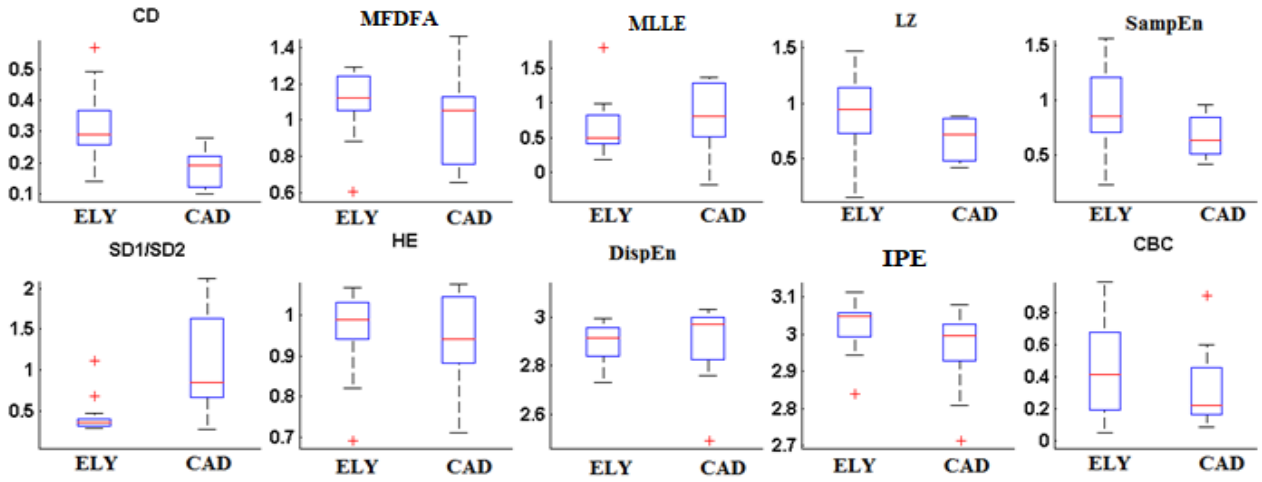


Figure 5.1: Box plot of top ten ranked features for ELY-CAD dataset before reduction by GDA and KPCA

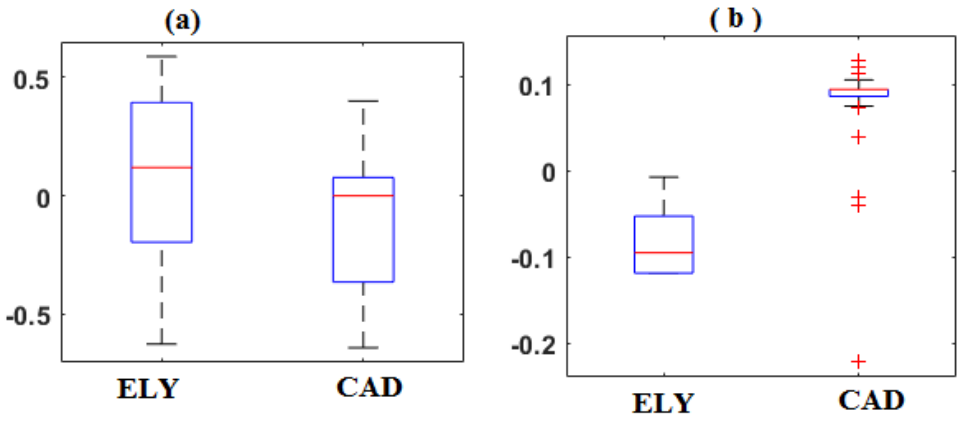


Figure 5.2 Box plot of after dimension reduction of top ten ranked chaos and nonlinear features to a new feature for ELY-CAD dataset using (a) KPCA with RBF kernel function (b) GDA with RBF kernel function.

5.5 Parameters Used for Simulation

The performance of 1-NELM scheme has been tested on Sigmoid activation function and Multiquadric functional additive node. The multiquadric function was considered for RBF hidden nodes. The sigmoid function and Multiquadric function are defined as $\emptyset(\mathrm{A}, \mathrm{B}, \mathrm{X}) = 1/1 + exp\ -(A^t X + B)$and $\emptyset(\mathrm{A}, \mathrm{B}, \mathrm{X}) = \sqrt{(||X - A||_2^2 + B^2}$.

In the SVM, the dimension of the original features was mapped to a higher-dimensional vector space by using an RBF kernel. RBF kernels function depends on two parameters: regularization parameter or error/trade-off (C) parameter and the width of the RBF. These parameters regulate the split-up error in the formation of the separation surface [150]. For SVM implementation MATLAB toolbox of LIBSVM has been used [199]. In the PNN, a Gaussian activation function was used for the pattern layer and each pattern layer has 15 neurons. A linear activation function and a linear basis function were adopted for the summing nodes [200]. The optimal values of user-specified parameters like regularization parameter γ for 1-NELM, C and width of the RBF for SVM, and scale for PNN were selected using 10-fold cross-validation method.

To investigate the performance of ML classifier for group of database, following performance parameters such as Accuracy $(AC) = (TP + TN)/(TP + TN + FN + FP)$, Sensitivity (SE) or true positive rate $(TPR) = TP/(TP + FN)$ and specificity (SP) or true negative rate $(TNR) = TN/(FP + TN)$ and false positive rate $(FPR) = (1 - SP)$ through confusion matrix have been measured. Where TP = true positive, FN = false negative, TN = true negative and FP = false positive. The area under the curve (AUC) for the ROC curve was determined by the formula $AUC = \int_0^1 ROC(\tau)\, d\tau \cong \frac{1}{2}(\mathrm{SE} + \mathrm{SP})$. Where

$\tau = (1 - SP)$ and $ROC\ (\tau)$ is sensitivity. A 1-norm root mean square error $(1 - NRMSE)$ was also calculated to measure the generalization performance.

5.6 Simulation Results

5.6.1 Comparative Generalization Performance

The 1-NRMSE was employed to measure the generalization performance and insensitive to user-specified parameters of the classifier for YNG-CAD. It was calculated using mathematical formula as $RMSE = \sqrt{\frac{1}{Le}\sum_{j=1}^{Le}(d_j - \bar{d}_j)^2}$, where d_j and $\bar{d}_j$ are the observed and predicted values,

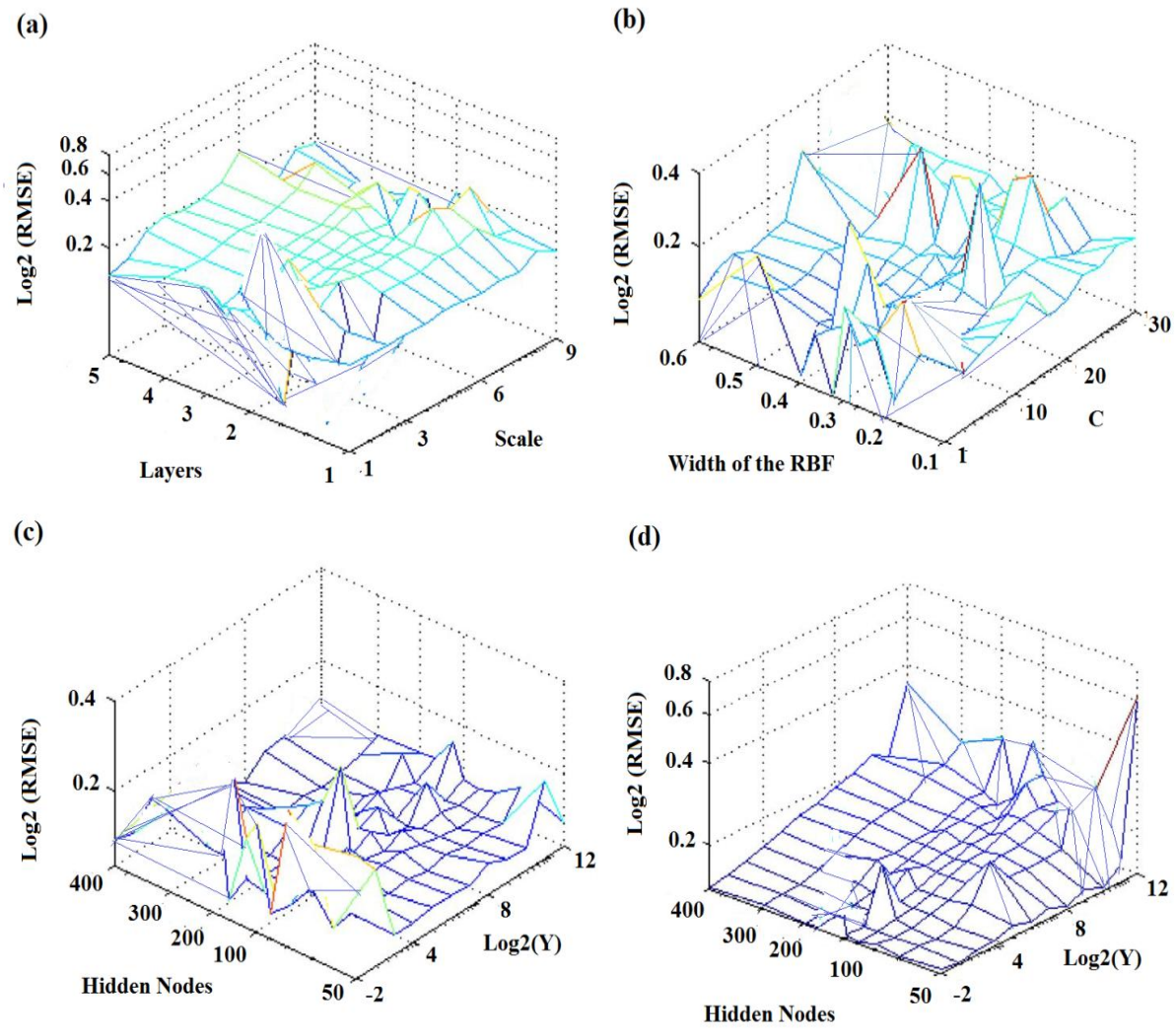


Figure 5.3 Insensitivity performance of (a) PNN (b) SVM (c) 1-NELM with Multiquadric function and (d) GDA+1-NELM Multiquadric function to the regularization parameters on YNG-CAD datasets.

and Le indicates length of test samples. Fig. 5.3 demonstrates the performance of considered classifier PNN, SVM with RBF kernel, 1-NELM with Multiquadric activation function and proposed GDA+1-NELM with Multiquadric activation function. In which PNN (see Fig. 5.3.a) and SVM (Fig. 5.3.b) are sensitive to the user specified parameters. However, it can be observed by Fig. 5.3.d that proposed classifier GDA+1-NELM can achieved best generalization performance with the smallest 1-norm RMSE and it is not very sensitive to the user specified parameters fascinatingly even for small values of hidden nodes (L)compare to 1-NELM (Fig.5.3 c). If numeral of values of L becomes high, it leads to increase in execution time. It reveals that better generalization outcomes can be attained using number of modest/ low values of hidden nodes such as $L = 150$ using GDA with 1-NELM classifier.

5.6.2 *Area Under ROC Curve for detection of CAD*

An area under ROC curve (AUC) is a output metric a specific threshold setting for classification problem. ROC is a probability graph, and AUC is a level or separability metric and effective approach to précis the general diagnostic accurateness of the test. AUC also calculates the probability that a random true positive rate would ahead of a random false negative rate. It says how much ML one can differentiate among classes. Higher the AUC, the easier for ML is to differentiate between subjects with CAD and healthy subjects. Fig. 5.4 shows the AUC values with respect to top ten features ranked by Fisher score approach achieved by considered ML algorithms as GDA +1-NELM, 1-NELM and GDA+PNN, SVM, GDA+SVM and PNN. The position (rank) of the topmost ten features based on highest Fisher score is presented in Table.5.1. The AUC values were obtained by highest ranked features fed to the considered ML one after another (like 2,3, 4……10 features) untill the maximum AUC was reached. Keep in mind that GDA limited each feed to one new feature and originally wanted a total of two features to do so. In this analysis, the multiquadric activation function employed as additive hidden for 1-NELM classifier and SVM were simulated using a RBF kernel function. Fig.4.4 (a) illustrations that the best AUC value (0.98) was achieved by GDA+1-NELM classifier after feeding 7 to 10 features and AUC curve becomes flat. The GDA+SVM acquired the highest AUC value (0.81) which is better to GDA+PNN,

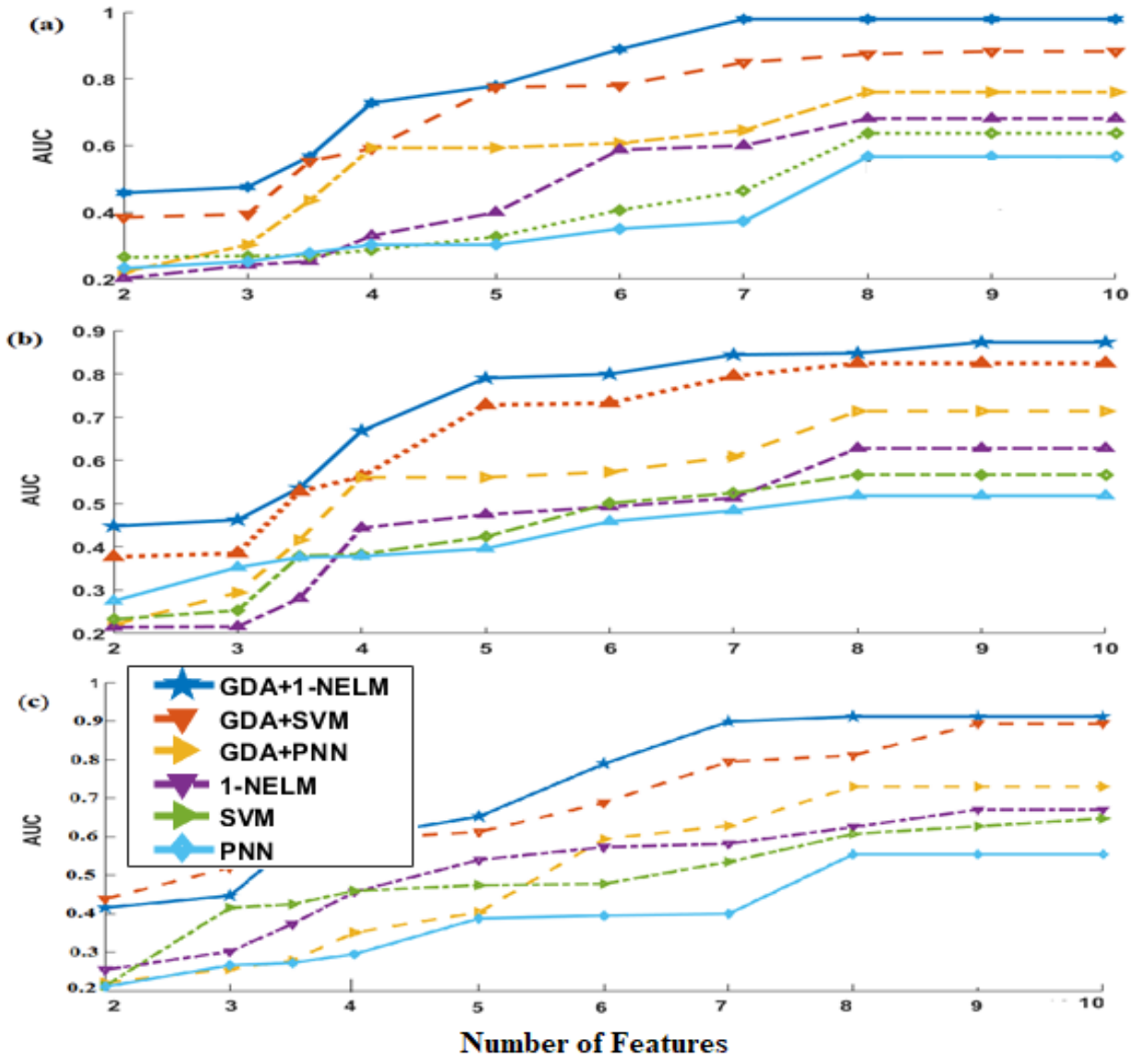


Fig. 5.4 Represents AUC values acquired by ML for (a) YNG-CAD (b) ELY-CAD (c) YNG-ELY datasets. The ranks of top ten features were calculated using Fisher approach.

1-NELM and remaining ML for this dataset and range of attributes. This curve also indicates that AUC value is reduced in case ranked attributes are fed ≤ 7 and all ML algorithams achieved AUC less than 0.6. As, AUC < 0.6 corresponds to the likelihood inequality that curves in ROC space on a diagonal axis. It means that ML algorithm wrongly classifies some diseased CAD subjects as negative and some healthy YNG subjecs as positive which is extremely unlikely to occur in clinical practice. Fig.5.4 (b) represents the AUC values attained by all MLs algorithm for ELY-CAD dataset. In which GDA+1-NELM ML classifier achieved highest AUC (0.88) at or above 9 features compared to other ML for this dataset. Fig.5.4 (c) shows that GDA+1-NELM classifier acquired constant AUC (≈ 0.9) value ranges 7 to 10

features for YNG-ELY dataset. However for this dataset, GDA + SVM classifier achieved constant AUC (≈ 0.887) value at or above 9 features. The AUC values of Fig.5.4 reveals that the GDA+1-NELM classifier performs better to other considered ML for all dataset beacause AUC (≥ 0.9) value of the this classifier is higher than that of the remaining ML algorithms. From the plot analysis, it is also observed that the increment of AUC occurs when GDA is combined with ML relative to when GDA is not combined with ML. It happened because of application of GDA in conjuction of RBF kernel in the form of attributes dimension transformation tool and it transformed to a new attributes by selecting the excellent selective attributes of the top ten rated attributes by Fisher approach. Hence, on the basis of simulation results, It is founded that GDA+1-NELM with multiquadric activation function classifier can be used for classification and detection of CAD and healthy subjects.

Table 5.1 Shows the rank number with Fisher score. Suppose: Y (Fisher score), here Y designates the index of feature and in (..) Fisher score value. Example for ranking order of features for $YNG - CAD$**:** 4 (SampEn) means at 1st rank, 5 (LZ) at 2nd rank, 6 (DispEn) at 3rd rank…..

Features	CD	MFDFA	MLLE	SampEn	LZ	DispEn	SD1/SD2	HE	IPE	CBC
Index	1	2	3	4	5	6	7	8	9	10
Rank	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
YNG-CAD	4 (2.9936)	5(2.67886)	6(2.61784)	1(0.51122)	10(0.2569)	3(0.06843)	9(0.05751)	7(0.018805)	2(0.01611)	8(0.00179)
ELD-CAD	5(1.15658)	4(1.0658)	6(0.67515)	1(0.36481)	10(0.0737)	8(0.01665)	9(0.01267)	7(0.00958)	2(0.00457)	3(0.00170)
YNG-ELY	1 (0.3561)	3 (0.21123)	9 (0.14217)	6 (0.08039)	4 (0.06761)	8 (0.04362)	2 (0.01077)	5 (0.00842)	10 (0.00432)	7 (0.00300)

The ROC analysis can be used to compare the classifier when results from multiple classifiers have been acquired for the same dataset, which is shown in Fig.5.5. The comparatively of the ROC graph specify the relative AC of the proposed GDA+1-NELM and other considered classifier combined with GDA. The simulation of ROC was created on the assumption of classifier scores resultant to the positive condition and the scores consistent with the negative condition can each be characterized by a Binomial posterior distribution. The ROC graph lying upstairs and to the leftward of another graph achieved from classifier reflects better observed AC and appropriate for detection of CAD. Top ten rank features were used for ROC analysis and feature rank by using Fisher score. Fig. 5.5(a) illustrates that the GDA+ 1-NELM is apparently more accurate than other considered methods in detecting-CAD group. This Figure also indicates that the proposed approach has a lower FP fraction at any given TP fraction, another hand higher TP fraction at any specific FTP fraction. It is perceived visually from Fig. 5.5(a) that the proposed method, GDA+ SVM and GDA+ PNN has the same TPR after o.6 FPR in the detection of YNG-CAD group. In the case of ELD-CAD dataset, as in Fig. 5.5.(b), the ROC curve acquired by GDA+ 1-NELM show better diagnostic capabilities compared to GDA+ SVM and GDA+PNN. Fig. 5.5(c) also demonstrates the proposed method captured higher ROC compared to other considered classifier.

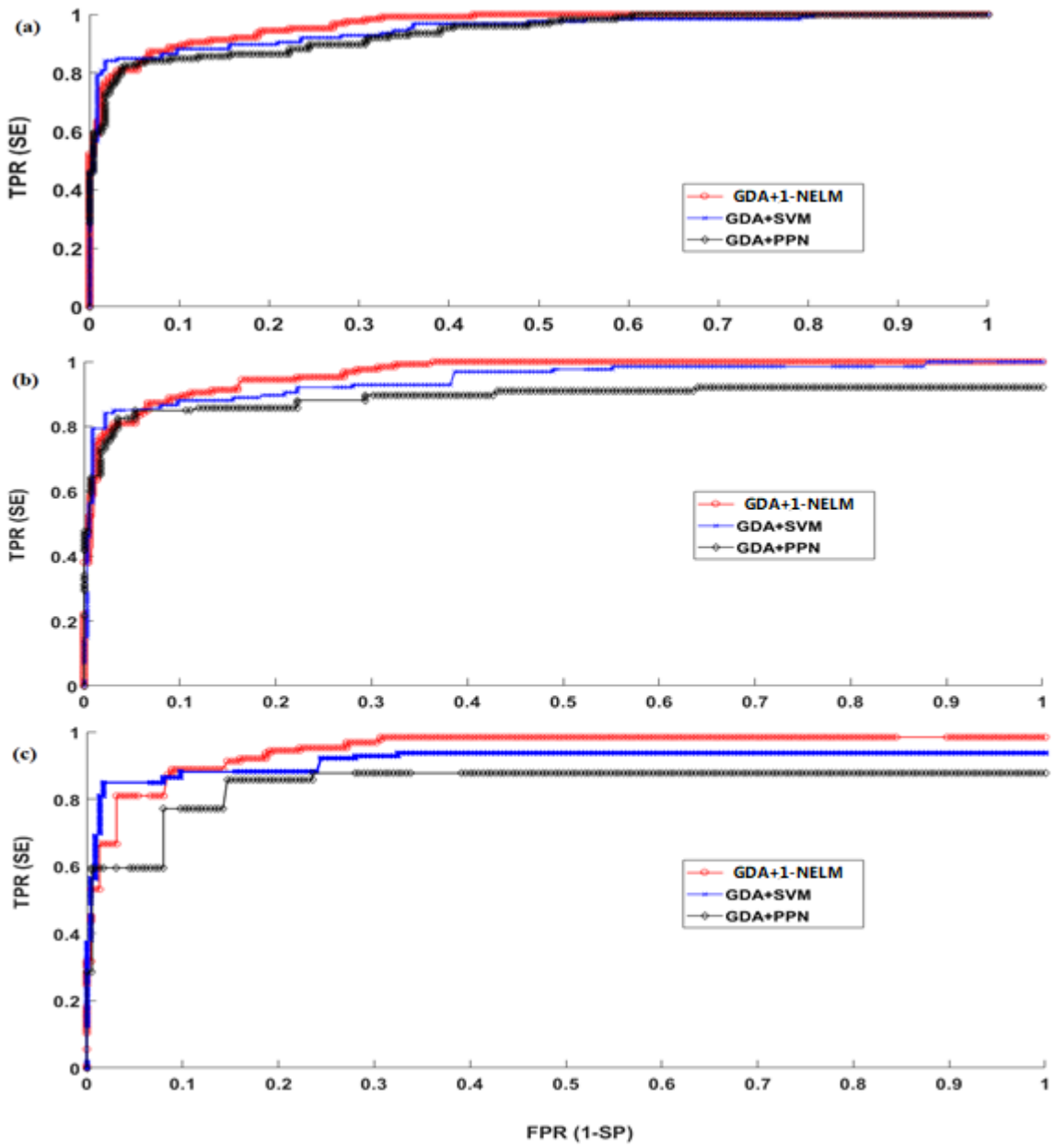


Fig. 5.5 Comparative ROC curves for datasets (a) YNG-CAD (b) ELD-CAD (c) YNG-ELD. For this top ten rank features were used.

Fig. 5.5 illustrate that GDA+ 1-NELM attained greater AC than others consider classifier. It reveals the GDA+ 1-NELM is appropriate classifier for detection of CAD.

The Wilcoxon signed-rank test was conducted to compare the performance values of 1-NELM, SVM and PNN for data set YNG-CAD, ELD-CAD and YNG-CAD. For this study, the first top ten features are transferred to a new feature using GDA. The results are reported in Table 5.2. For all the comparisons of

classifier used in this study, statistical significance level (α) was set to 0.05. The p-values were assessed to focus on the significance of differences in the classifier.

The numerical results reveal that the 1-NELM outperformed SVM (p<0.006) significantly and PNN (p< 0.0001) very significant for YNG-CAD dataset, while SVM and PNN did not have significant (P> 0.05) different performance. The 1-NELM and SVM outperformed PNN (p< 0.0001) very significantly for ELD-CAD, whereas 1-NELM and SVM does not show significantly (p> 0.05) different performance. However, for YNG-ELD dataset, the 1-NELM outperformed SVM and PNN (p< 0.0001) very significantly.

Table 5.2. Wilcoxon Signed Rank test comparison for the performance of 1-NELM, SVM and PNN using the selected 500 samples of the new feature. These samples were taken after the reduction of the top ten features using GDA. R^+ corresponding to the sum of the ranks for the method on the left and R^- for the right.Vs. indicates versus, * indicates significant and ** indicates very significant.

Datasets	Method	Accuracy			Sensitivity			Specificity		
		R^+	R^-	p-value	R^+	R^-	p-value	R^+	R^-	p-value
YNG-CAD	1-NELM Vs. SVM	451	39	0.0045*	479	11	0.0012*	412	78	0.0089*
	1-NELM Vs. PNN	490	0	0.0001**	490	0	0.0001**	490	0	0.0001**
	SVM Vs. PNN	349	141	0.115	363	127	0.105	327	163	0.213
ELD-CAD	1-NELM Vs. SVM	333	157	0.218	381	109	0.101	292	198	0.287
	1-NELM Vs. PNN	490	0	0.0001**	490	0	0.0001**	490	0	0.0001**
	SVM Vs. PNN	490	0	0.0001**	490	0	0.0001**	490	0	0.0001**
YNG-ELD	1-NELM Vs. SVM	490	0	0.0001**	490	0	0.0001**	490	0	0.0001**
	1-NELM Vs. PNN	490	0	0.0001**	490	0	0.0001**	490	0	0.0001**
	SVM Vs. PNN	490	0	0.0001**	490	0	0.0001**	490	0	0.0001**

5.6.3 *Statistical Analysis of Features*

Statistical analyses of all features were analyzed by using unpaired t-test for all considered datasets. This Statistical approach yields p-value that is suitable for estimating the importance of the attributes for identification CAD subjects. A small p-value illustrates high chance to distinguish individual classes of sets of data. In a quantitative test setting, a p-value (< 0.05) indicates quantifiable significance and (< 0.001) quantifiable very significance. The p-values of the top ten features: $CD, MFDFA, MLLE, SampEn, LZ, AMPE\ \ DispEn, SD1/SD2, HE, , IPE$ and CBC for each data are listed in Table 5.3. The p-value attained by $CD, LZ, DispEn$ and $SD1/SD2$ features are (< 0.001) for each dataset. Hence these features can be used for analysis for all considered dataset. The feature $SampEn, MLLE$ and $IMPE$ are significant $(p < 0.05)$ only for $YNG - ELD$ dataset. The $SampEn$ and

CBC features are significant $(p < 0.001)$ for both $YNG-CAD$ and $ELD-CAD$ datasets. Some features like $MFDFA, HE$ and PE did not distinguish between CAD and normal subjects significantly$(p > 0.05)$.

Table 5.3. Indicates P-values acquired by top ten features for f CAD and healthy subjects datasets. if $P > 0.05$: not significant, $p \leq 0.05$: quantifiable significant (*) and $p < 0.001$: quantifiable very significant (**).

Features	*CD*	*MFDFA*	*MLLE*	*SampEn*	*LZ*	*DispEn*	*SD1/SD2*	*HE*	*PE*	*IPE*	*CBC*
YNG-ELY (p-Value)	2.22E-05**	0.15845	7.96E-03**	5.48E-03**	4.12E-03**	1.32E-03**	6.8 E-03**	0.6675	0.0147*	5.12E-02	0.0161*
Features	*CD*	*MFDFA*	*MLLE*	*SampEn*	*LZ*	*DispEn*	*SD1/SD2*	*HE*	*PE*	*IPE*	*CBC*
YNG-CAD (p-Value)	1.88E-06**	0.2523	0.0378	5.20E-05**	1.88E-05**	1.20E-04**	3.14E-04**	0.0265	0.1854	0.0017*	9.34E-02*
Features	*CD*	*MFDFA*	*MLLE*	*SampEn*	*LZ*	*DispEn*	*SD1/SD2*	*HE*	*PE*	*IMPE*	*CBC*
ELY-CAD (p-Value)	8.96E-05**	0.6331	0.49	1.03E-05**	1.24E-03**	2.22E-05**	1.16E-03**	0.0807	0.4547	0.1906	1.72E-02*

5.6.4 *Performance of 1-NELM and GDA With Ranking Approach.*

In this part, we analyzed classification accuracy achieved by various ranking approach with $1-\text{NELM}$ and with GDA and $1-\text{NELM}$ and tabulated in Table 5.4 The performance of classification accuracy (AC) was calculated using a $10-\text{fold}$ cross-validation scheme with 50 trials to check the robustness and improve the efficiency of these ML algorithms. The datasets were randomly alienated by 90 out of 160 from $\text{YNG}-\text{ELY}$, 84 out of 132 from $\text{YNG}-\text{CAD}$ and 80 out of 132 from $\text{ELY}-\text{CAD}$ datasets for testing and remaining data fo validation. The error rate was then calculated on the 10th fold in terms of standard deviation $(\pm \text{S. d})$. The learning process, thus, has carried out for a sum of 500 folds on various testing and validation for datasets which culminated in the 500-fold implementation of the learning method on datasets. With each datsets, the average validity accuracy was measured, and also the average of 500 error rate estimated to generate a accuracy in terms of $(\text{AC\%} \pm \text{S. d})$.

Table 5.4 lists training and validation data dimension which were used for analysis, simulation and accuracy achieved by ranking approaches with $1-\text{NELM}$ and GDA and $1-\text{NELM}$ ML classifier using such as Sigmoidal,± Multiquadric activation functions and RBF kernel incunjuction with GDA. Simulation data of Table 5.4 reveals that $1-\text{NELM}$ classifier with GDA and ranking technique accomplished better than $1-\text{NELM}$ with ranking approach. The Fisher with GDA and $1-\text{NELM}$ and Bhattacharya with GDA and 1-NELM produced an accuracy of 100% for each group of data sets and both activation function. While Bhattacharya with $1-\text{NELM}$ yielded an accuracy of 94.09±3.33 using Multiquadric activation function. The Fisher with GDA and $1-\text{NELM}$ and Entropy with GDA and $1-\text{NELM}$ produced an accuracy of 99.25±0.45 and 99.75±0.15for YNG-ELY dataset using sigmoidal activation function. The ROC with GDA and $1-\text{NELM}$ attained an accuracy of 99.87±0.12 and

99.76±0.14 for YNG-CAD dataset using both activation function; and an accuracy of 100% for remaining dataset. However, ROC with $1-$ NELM achieved an accuracy of 94.81±3.38 using multiquadric activation

Table 5.4: presents Accuracy (AC) attained by the ranking approach with 1-NELM and with GDA and 1-NELM for group of database. The ML classification of datasets was carried out using *Sigmoid* and *Multiquadric* activation nodes. The $AC\ (\%)\ \pm\ S.D$ is shown in the table for the values $L = 150$.The better result is marked by bold dark in the Table.

Database Size (Train, Validation)	**Features of nonlinear Method**			
	Fisher + 1 − NELM		**Fisher + GDA + 1 − NELM**	
	Sigmoid	*Multiquadric*	*Sigmoid*	*Multiquadric*
YNG-ELY (90×10,70×10)	72.78±4.66	79.44±3.47	**100±0**	**100±0**
YNG-CAD (84×10, 48×10)	93.14±5.61	91.02±8.69	**100±0**	**100±0**
ELY-CAD (80×10, 52×10)	89.74±3.33	84.62±4.19	**100±0**	**100±0**
YNG+ELY-CAD(114×10,50×10)	86.12±5.28	89.12±6.28	**98.64±1.1**	**99.12±0.45**
Database Size (Train, Validation)	**Wilcoxon + 1 − NELM**		**Wilcoxon + GDA + 1 − NELM**	
	Sigmoid	*Multiquadric*	*Sigmoid*	*Multiquadric*
YNG-ELY (90×10,70×10)	75.16±4.22	76.12±12.05	**99.25±0.45**	**100±0**
YNG-CAD (84×10, 48×10)	90.31±11.31	87.23±10.2	**100±0**	**100±0**
ELY-CAD (80×10, 52×10)	90.17±3.11	83.76±3.74	**100±0**	**100±0**
YNG+ELY-CAD(114×10, 5×10)	88.45±7.18	79.23±4.11	96.04±2.1	98.10±1.23
Database Size (Train, Validation)	***Bhattacharya* + 1 − *NELM***		**Bhattacharya+GDA+ 1-NELM**	
	Sigmoid	*Multiquadric*	*Sigmoid*	*Multiquadric*
YNG-ELY (90×10,70×10)	74.91±7.6	75.22±3.16	**100±0**	**100±0**
YNG-CAD (84×10, 48×10)	95.51±3.2	94.09±3.33	**100±0**	**100±0**
ELY-CAD (80×10, 52×10)	91.24±2.5	85.47±7.39	**100±0**	**100±0**
YNG+ELY-CAD(114×10, 5×10)	87.15±4.28	85.11±3.12	99.03±0.42	98.32±1.03
Database Size (Train, Validation)	***Entropy* + 1 − *NELM***		**Entropy+ GDA+1-NELM**	
	Sigmoid	*Multiquadric*	*Sigmoid*	*Multiquadric*
YNG-ELY (90×10,70×10)	75.91±3.98	83.27±6.75	**99.75±0.15**	**100±0**
YNG-CAD (84×10, 48×10)	94.09±4.48	94.8±2.84	**100±0**	**100±0**
ELY-CAD (80×10, 52×10)	87.82±3.33	87.61±4.19	**100±0**	**100±0**
YNG+ELY-CAD(114×10, 5×10)	84.12±5.18	82.41±8.13	97.45±1.8	99.11±0.84
Database Size (Train, Validation)	***ROC* + 1 − *NELM***		***ROC* + *GDA* + 1 − *NELM***	
	Sigmoid	*Multiquadric*	*Sigmoid*	*Multiquadric*
YNG-ELY (90×10,70×10)	75.16±3.15	78.29±6.36	**100±0**	**100±0**
YNG-CAD (84×10, 48×10)	93.14±7.35	94.81±3.38	**99.87±0.12**	**99.76±0.14**
ELY-CAD (80×10, 52×10)	88.89±6.06	83.76±5.28	**100±0**	**100±0**
YNG+ELY-CAD(114×10, 5×10)	85.64±7.13	87.89±10.04	96.79±1.6	98.13±1.04

functions. Globally, results depicts that enhancement in the accuracy occurred when GDA combined with considered ML in comparisons to when ML is not combined withGDA. The simulation data also shows

that ranking methods with GDA and 1-NELM model produced an accuracy of 100% for all datasets taken into account using Multiquadric function.

Table 5.5 presents the results of execution performance in terms of validation time (in second) for datasets. Comparing with another classifier, the proposed GDA and $1-$NELM with Sigmoid activation function takes the least time. When using Multiquadric activation with it, the validation time of GDA+ 1-NELM is a little longer than that of GDA and $1-$NELM with Sigmoid but shorter than that of $1-$NELM with both activation function. Though, out of all ML schemes, PNN takes highest execution time. The results of Table 5.5 reveals that proposed method does not require too much computational time to diagnose CAD disease and is easy to design.

Table 4.5: Validation times (in seconds) compared for all six methodologies for classification of data sets namely YNG-ELD, YNG-CAD and ELD-CAD. The validation time calculated after ranking by Fisher score.

Methods	YNG-ELD	YNG-CAD	ELD-CAD
GDA+ 1-NELM (Sigmoid)	4.56	3.962	3.962
GDA+1-NELM (Multiquadric)	6.24	5.41	5.41
GDA+SVM (Sigmoid)	240.41	208.43	208.43
GDA+SVM (Multiquadric)	242.35	210.07	210.07
GDA+PNN	338.62	320.17	320.17
1-NELM (Sigmoid)	45.62	39.55	39.55
1-NELM (Multiquadric)	62.43	54.13	54.13
SVM (Sigmoid)	2404.84	2084.99	2084.99
SVM (Multiquadric)	2423.61	2101.27	2101.27
PNN	3386.13	3201.77	3201.77

5.7 Discussion

In this report, the use of 1-NELM combined with GDA and ranking methods to accurately detect the CAD disease has been evaluated for the first time. The methodology is modified from a set of learning algorithms which have been successful in detecting important cardiac disease in ELY and YNG subjects. The simulated results have shown that proposed learning with both Sigmoid and Multiquadric activation function, achieved higher accuracy and has taken less time than the other considered learning algorithm to detect CAD subjecs. The PNN classifiers are very slower than SVM and 1-NELM. It may be due to a number of pattern layers which has been used to process the attributes (features).

Table 5.6 expresses a summary of research carried out by various authors for detection of CAD by extracting the features from HRV using several techniques, different features dimension reduction approach and several ML binary classifier such as SVM, $k-$nearest neighbors (KNN) algorithm, artificial neuron network (ANN). The authors of [201], [196], [142] and [202] have applied linear (single degree) and non-linear (multi degree) domain techniques for attributes extraction from HRV time seriese and t-test, principal component analysis (PCA), GDA for features diminishment and SVM and least

squares support vector machine $(LS-SVM)$ for detection of heart disease, an accuracy of 90%, 99.72%, 95.77%, and 100% were achieved respectively.

Table 4.6: A brief details of comparison of the performance of the investigation made as compared to the earlier investigations.

Reference	Techniques	Attributes	ML	Accuracy %
[203]	DWT, HSS and WPT	Several statistical features	ANN	90
[201]	Nonlinear methods, Linear, and reduced features by t-test	5 nonlinear features and 6 Linear	SVM	90
[142]	Linear, non linear methods and features reduced by GDA and LDA	7 time domain, 1 frequency and 7 non-linear	SVM	95.77
[204]	EST and features reduced by PCA	18 Features	SVM	79.17
[141]	Features reduced by PCA	6 Nonlinear features	MLP	89.5
[196]	TQWT and features reduced by PCA	2 Entropy features	$LS-SVM$ with Morlet wavelet kernel	99.72
[202]	FAWT and ranking method like Entropy, ROC and Bhattacharya	FzEn and K-NN Entropy estimator	$LS-SVM$ with RBF &Morlet wavelet kernel	100
[205]	Linear and non-linear, features reduced by PCA	Various time domain , Frequency domain and non-linear domain	PNN, KNN and SVM	Without PCA : 68.33, 76.67 and 90.00 With PCA : 68.33,85.00 and 91.67
This work	Chaos investigation methods, ranking method and features reduced by GDA	11 features by chaos method.	1-NELM	100 (only for Multiquadric function)

Limitations of proposed method:

(a) Prior to the deployment of the proposed method for clinical purpose, it needs to be trained by a substantial data set of HRV.

(b) The proposed algorithm also requires selection of the activation function parameter as γ, and kernel parameters using cross-validation and trial criteria which upsurges the computational time.

5.8 **Summary**

A novel automated expert system approach for the diagnosis of CAD subject has been demonstrated in this research proposal. The proposed ML was based on assortment of ranking approaches, GDA and $1-NELM$ in which five different ranking approaches and one features space reduction technique has been used to enhance the classification performance. The excellent classification accuracy, very good generalization performance and least execution time have been achieved through proposed scheme. The ROC with GDA and $1-NELM$ approach achieved an accuracy of 99.76±0.14, 99.87±0.12 and 100±0 for

$YNG-CAD, YNG-ELY$ and $ELY-CAD$ groups respectively. The Fisher with GDA and 1-NELM; and Bhattacharya with GDA and $1-NELM$ approach achieved an accuracy of 100±0 for all considered datasets. The GDA+ 1-NELM (Sigmoid) method achieved lowest validation time 4.56, 3.962, 3.962 seconds among all other considered ML for YNG-ELD, YNG-CAD, ELD-CAD datasets. This automated technique will be useful for treatment, intensive care, analysis of hart diseases and distinguish between CAD and healthy subject.

6. Detection of CHF Using Multiresolution Wavelet Packet Features and 1-NLPELM

6.1 Overview

This section describes for binary classification of CHF disease based on features extracted decomposition from HRV signal using Multiresolution Wavelet Packet (MRWP) method. 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-norm linear programming extreme learning machine (1-NLPELM) classifier. The Sigmoid/ Multiquadric activation function has been employed in 1-NLPELM to introduce nonlinearity into the output of a hidden node. The block diagram of proposed is shown in Figure 6.1.

6.2 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 [186], [34]. 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 6.1.

Pre-processing: The HRV time series obtained from the standard database usually contains ectopic beats (irregular impulse formation in the heart muscle) and non-stationary, which precludes effective feature extraction and HRV data analysis. To avoid this problem, the pre-processing of the HRV signal is required [128], [157]. After pre-processing, the HRV is known as a normal-normal interval (NN interval). The NN intervals were re-sampled at 4Hz.

Table 6.1: 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

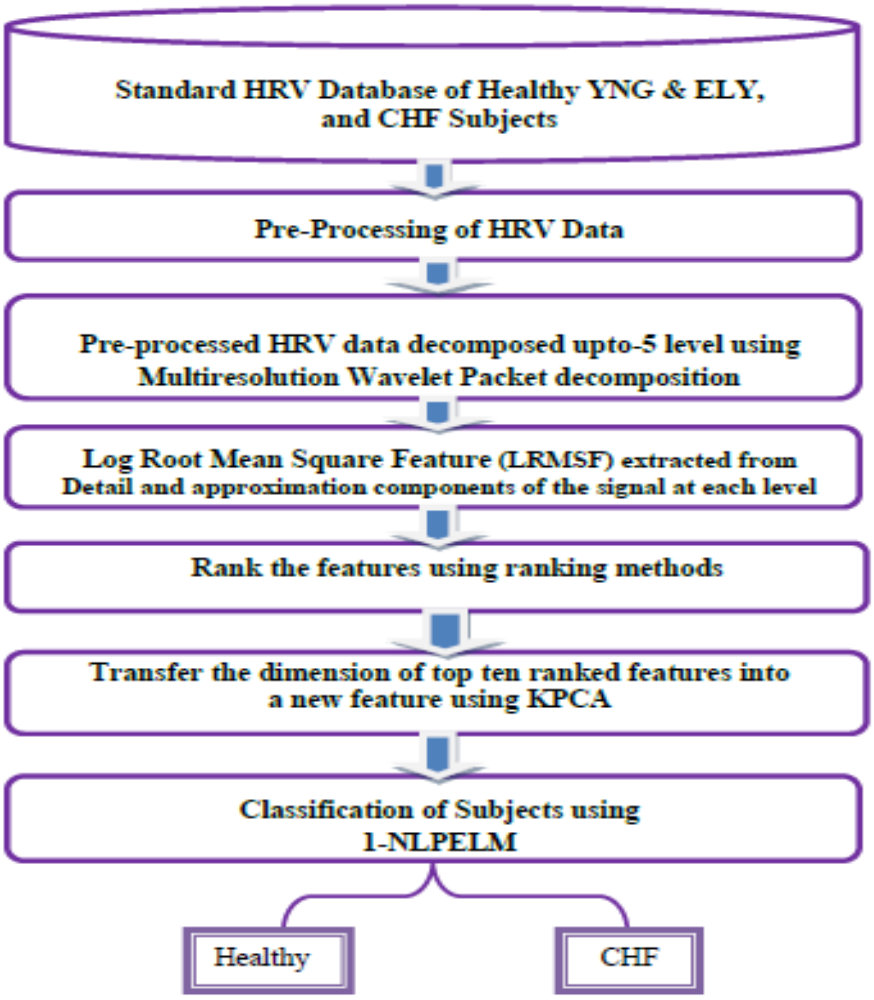


Fig. 6.1: Represents a block diagram of proposed model for CHF detection

6.3 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 [206], [207]. 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 equation (1) and (2) by the following recursive associations [208].

$$W^{2k}(t) = \sqrt{2}\ \sum_n S(n)\ W^k(2t-n) \tag{6.1}$$

$$W^{2k+1}(t) = \sqrt{2}\ \sum_n g(n)\ W^k(2t-n) \tag{6.2}$$

Here assume that there are the first two components of wavelet packets $W^0(t) = \emptyset(t)$ and $W^1(t) = \emptyset(t)$ 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 equation (3) and (4) may be disintegrated recursively

$$X_{j+1,2k}(t) = \sum_m s(m-2n)x_{j,k}(t) \tag{6.3}$$

$$X_{j+1,2k+1}(t) = \sum_m g(m-2n)x_{j,k}(t) \tag{6.4}$$

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 equation (5).

$$X(t) = \sum_{k=0}^{2^j-1} x_{j,k}(t) \tag{6.5}$$

The 3^{rd} l level decompositions of HRV signal $X(t)$ using the MRWP is shown in **Fig. 5.2**. 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 comportment of HRV signal [209]. The $LRMSF$ attribute is expressed in equation (6).

$$LRMSF = \log \sqrt{\frac{X_1^2 + X_2^2 + ----+ X_N^2}{N}} \quad (6.6)$$

Where X_1, X_2 …… 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.

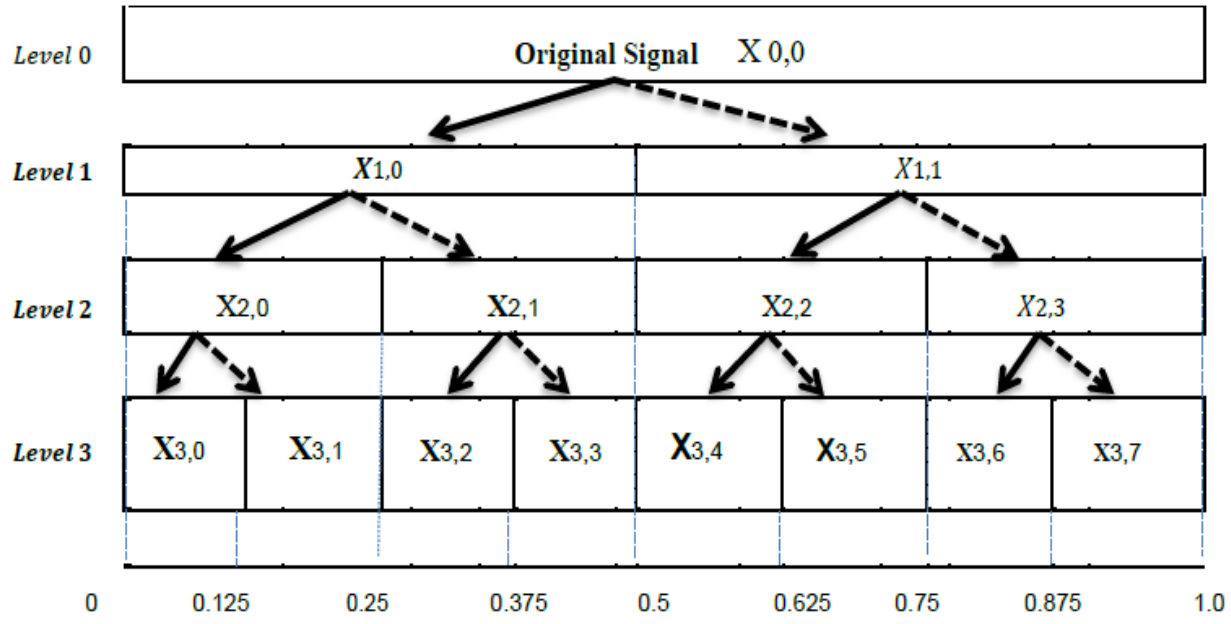


Figure.6.2: Demonstrates 3 levels decomposition of HRV signal using MRWP method, horizontal axis denotes frequency variation as a fraction of the sample frequency. The $X1,0$; $X1,1$; $X2,0$…… represents LF and HF components of HRV.

6.4 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 [210]. 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 [211]. The *Fisher* score is attained by applying equation (7). The *Fisher score* of i^{th} attribute in j^{th} level matrix is define as

$$F(i) = \frac{\sum_{j=1}^{n} N_J(\mu_{i,j} - \mu_i)^2}{\sum_{j=1}^{n} N_J(\sigma_{i,j})^2} \tag{6.7}$$

Where μ_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 σ denotes standard deviation.

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

$$En(X^J) = \frac{\left(\frac{\mu_1}{\mu_2} + \frac{\mu_2}{\mu_1} - 2\right)}{2} + (\mu_1 - \mu_2)^2 \times \frac{\left(\frac{1}{V_1} + \frac{1}{V_2}\right)}{2} \tag{6.8}$$

$En(X^J)$ indicates Entropy score of j^{th} attributes. μ_1 and μ_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 equation (9)

$$Bh(X^J) = \frac{4\times(\mu_1 - \mu_2)^2}{(d1+d2)} + 2\ \log\left(\frac{(d1+d2)}{\sqrt{\frac{d1\times d2}{2}}}\right) \tag{6.9}$$

Where $d1 = \sqrt{V_1}$, $d2 = \sqrt{V_2}$ and $Bh(X^J)$ denotes Bhattacharya score of j^{th} attributes. μ_1 and μ_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 equation (10)

$$Wx(X^J) = Absolut\left(\frac{N_2 \times sum(Ranks(group,:))}{N_1} - 1\right) \tag{6.10}$$

$Ranks = \text{tiedrank}(X)$,$N_1 = number\ of\ first\ group\ attributs,$

$N2 = number\ of\ second\ group\ attributs.$

6.5 Kernel Principle 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 [140]. Different strategies have been proposed to reduce the dimension of attributes to classification and detection of cardiovascular diseases [141], [195], [196]. In this paper, KPCA has been applied for reduction of attributes. Its advantages are nonlinearity of eigenvectors and greater number of eigenvectors [212]. 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 [198]. 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.

6.6 Result and Discussion

6.6.1 *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. 6.3. To comprehend the separate completely 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.6.3 and after reduction of top ten LRMSF by $KPCA$ are illustrated in Fig. 6.4(a); and by GDA are shown in Fig. 6.4 (b). The box-plot patterns (median value indicated by red line in box) of ten LRMSF associated with $\mathrm{NSR} - \mathrm{CHF}$ 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. 6.4(a) and Fig. 6.4(b) that KPCA produces more separate linearly pattern in NSR-CHF group compared to GDA after attributes size reduction. Hence, KPCA reduction scheme can be used for dimension reduction of top ten LRMSF extracted from NSR-CHF dataset.

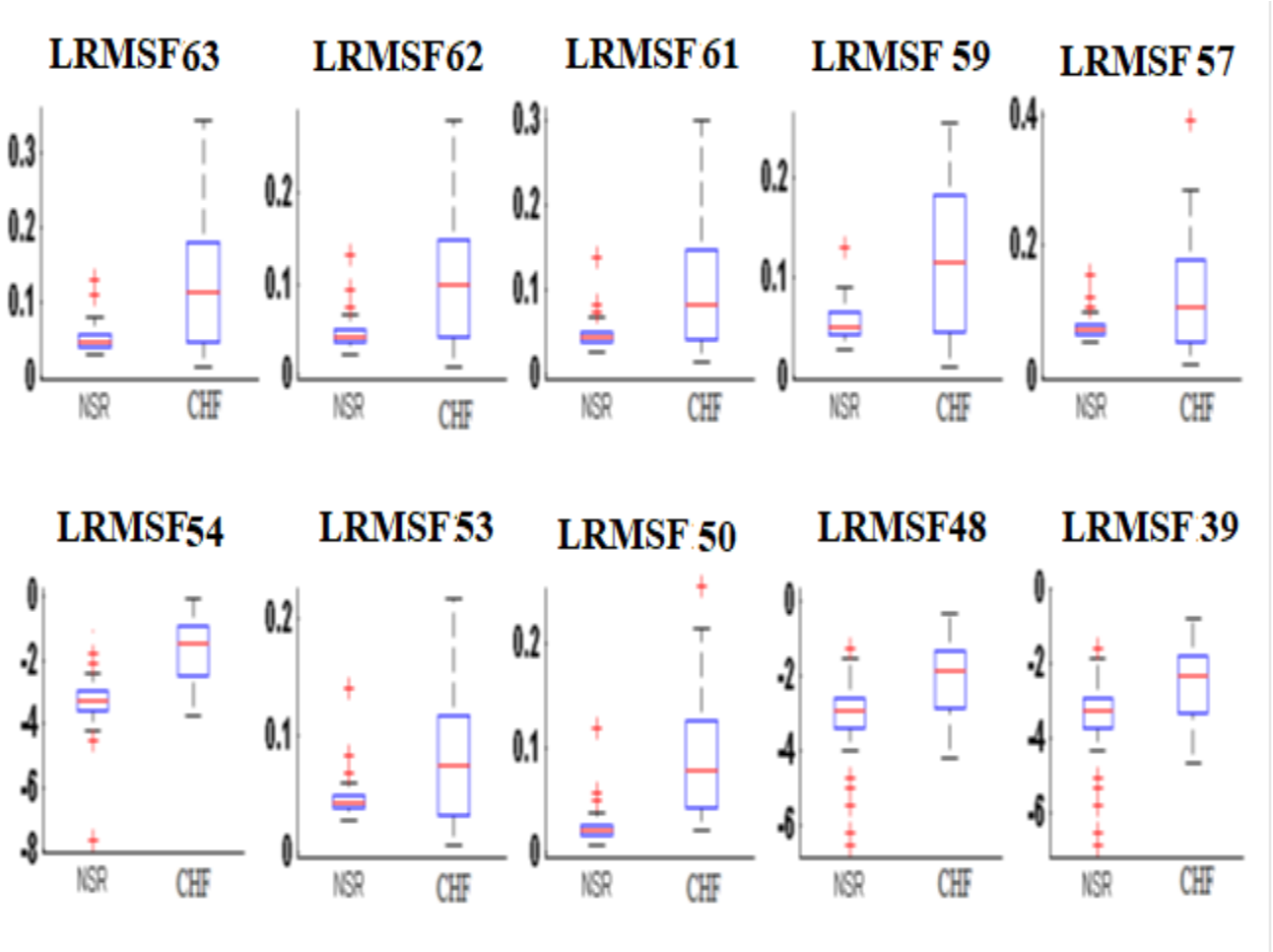


Fig. 6.3: Box plot of top ten ranked LRMSF attributes of NSR-CHF dataset before dimension mapping by GDA and KPCA.

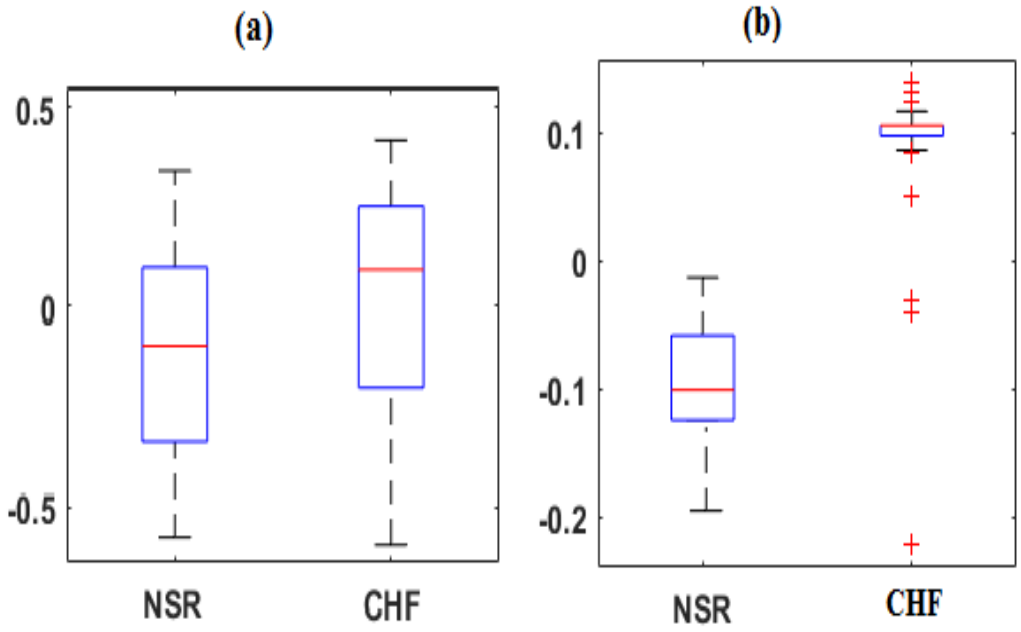


Fig. 6.4. Box plot, after dimension mapping of top ten ranked LRMSF to a new attribute for $NSR-CHF$ dataset using (a) GDA with radial basis function. (b) KPCA with Radial Basis Function.

6.6.2 *Classifier Performance Using AUC and Ranked Features*

In this section, the AUC value achieved by machine learning (ML) such as $KPCA+1-NELM$, $KPCA+SVM$, $KPCA+PNN$, $1-NELM$, 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. 6.5 for $NSR-CHF$ (Fig. 6.5(a)), for $NSR-ELY$ (Fig. 6.5(b)) and for $ELY-CHF$ (Fig.6.5(c)). 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. Fig.6.5 (a) 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. Fig.6.5 (b) depicts that

KPCA + 1 − NLPELM also accomplished highest AUC (0.72) among all considered ML after eight attributes for NSR − ELY

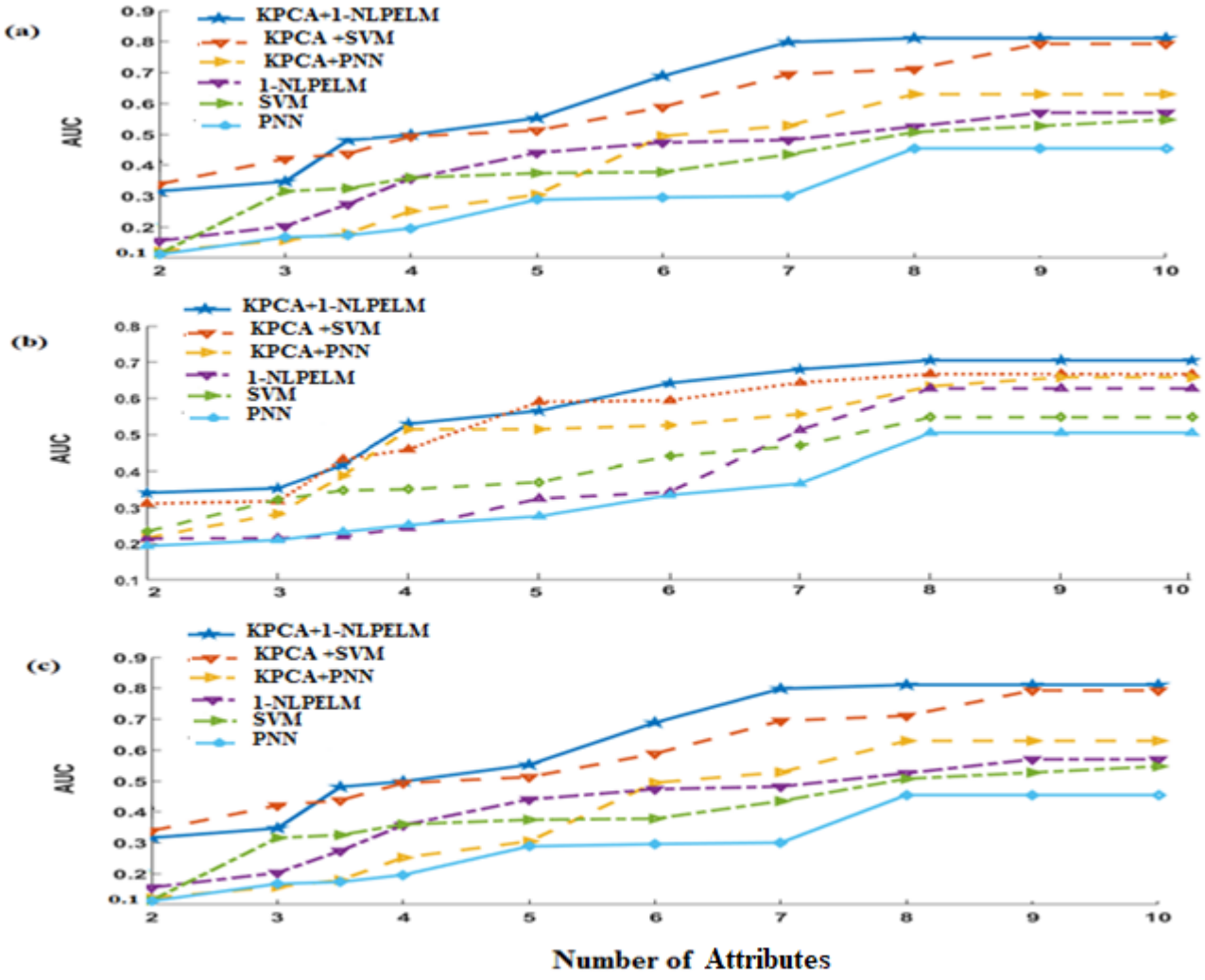


Fig. 6.5: Demonstrates the AUC value achieved by MLs at number of ranked attributes for (a) NSR-CHF (b) NSR-ELY (c) ELY-CHF datasets.

dataset. Fig.6.5 (c) 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. Fig. 6.5 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.

6.6.3 *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 γ 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 γ varies between $\{2^{-1}, 2^{40}\}$ and N varies from $\{50, 60 \ldots 100, 300, 400, 800 \; and \; 1000\}$.

Fig.6.6 depicts the value of accuracy for variation in every pair of γ and N for $1 - NLPELM$ with $Sigmoid$ additive node and $Multiquadric$ radial basis function node. Fig.6.6.(a) demonstrates that as γ 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.6.6. (b) that as γ 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.6.6 that the validation AC is much more sensitive for every pair of γ 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.

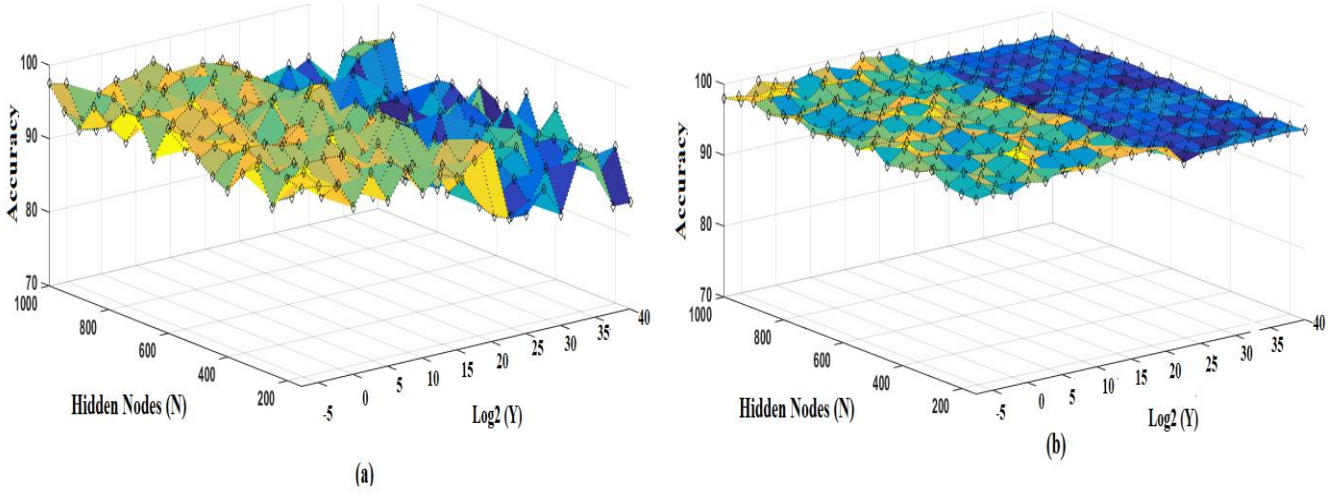


Fig. 6.6. Generalization performance of 1-NELM on the γ and N for (a) 1-NELM with Sigmoid additive node for NSR-CHF dataset; (b) 1-NELM with Multiquadric radial basis function node for NSR-CHF dataset.

Fig. 6.7 illustrates the value of accuracy for variation in every pair of γ and N for $KPCA + 1 - NLPELM$ with $Sigmoid$ additive node and $Multiquadric$ radial basis function. Fig. 6.7. (a) depicts that as γ 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

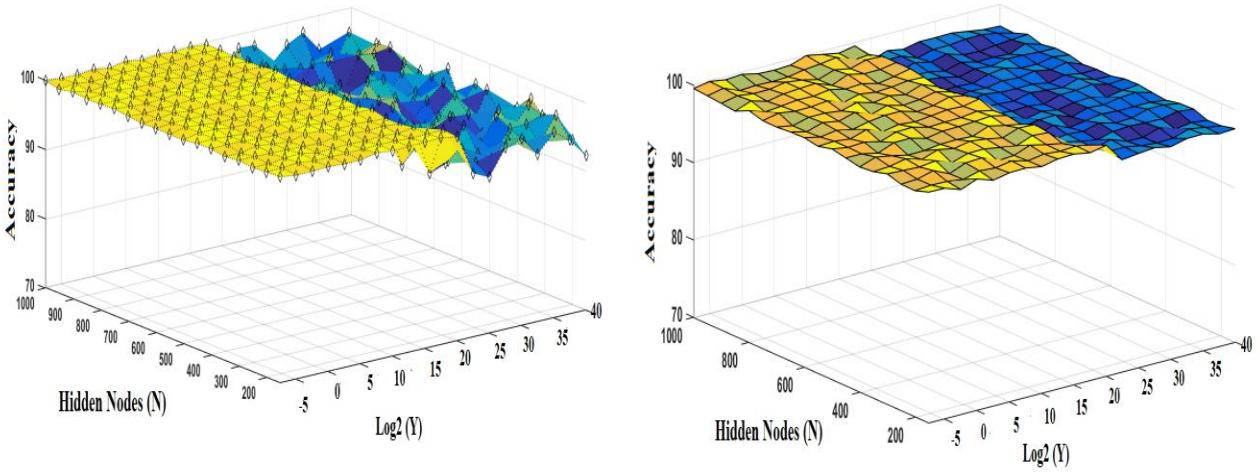


Fig. 6.7. Generalization performance of proposed model using (a) KPCA with RBF +1-NLPELM with Sigmoid additive node for NSR-CHF dataset (b) KPCA with RBF +1-NLPELM with Multiquadric Radial Basis Function node for NSR-CHF dataset.

when $Sigmoid$ additive function node is used with $KPCA + 1 - NLPELM$. It can also be observed that if the value of γ varies from $2^{25} < \gamma < 2^{40}$, the validation AC falls below 99%. It can be observed from Fig.6.7. (b) that with the variation in value γ and N ($2^{-2} < \gamma < 2^{25}, 50 < N < 400$), the validation AC remains constant at 99.96% when Multiquadratic RBF node is used with KPCA+ 1-NLPELM. It can be concluded from Fig.8 that for every pair of γ 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.

6.6.4 *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 ≤ 0.05 at 95% confidence limit is said to have considerable significance*, if p-value ≤ 0.001 then it said to be very significant ** and if p-value > 0.05 then it said to be statistically insignificant.

Table 6. 2: Repersents 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
$NSR-CHF$ **(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
	$7.19E-6$ **	$1.16E-6$ **	$1.99E-7$ **	$4.5E-7$ **	$1.60E-7$ **	$1.5\ E-8$ **	$3.12E-13$**	$2.51E-12$ **	$2.4E-14$ **	$1.7\ E-12$ **
Features	**LRMSF1**	**LRMSF 2**	**LRMSF 3**	**LRMSF4**	**LRMSF 5**	LRMSF 6	LRMSF 7	LRMSF 8	LRMSF 9	LRMSF 10
$NSR-ELY$ **(p-Value)**	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
	0.213	0.0004 **	0.00046 **	0.143	0.0041 **	$1.2\ E-4$ **	$1.1E-5$ **	0.00021 **	$1.4E-6$	$1.7E-7$ **
Features	**LRMSF1**	**LRMSF 2**	**LRMSF 3**	**LRMSF4**	**LRMSF 8**	LRMSF 6	LRMSF 7	LRMSF 13	LRMSF 9	LRMSF 10
$ELY-CHF$ **(p-Value)**	3.16	2.15	0.01021 *	1.10	0.1494	0.0067 **	0.0308 **	1.20	0.629	0.0178 *
	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 **

The results listed in Table 6.2 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. [140] 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.

6.6.5 *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 10-fold cross validation method. In each trial, 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 purposes which are listed in Table 6.3. The AC error rate was measured in the form of standard deviation (±S.d).

Table 6.3: Demonstrates the comparative performance of 1-NLPELM with ranking methods and KPCA+1-NLPELM with ranking methods in terms of AC (%) ±S.D for considered datasets. Bolded text in table shows the best results.

Database (Train size, Test size)	***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 (Train size, Test size)	**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	**96.06±3.17**	**97.91±1.43**
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 (Train size, Test size)	**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	**98.44±1.4**
NSR − ELY (100 × 10,60 × 10)	82.07 ± 3.49	89.03 ± 7.49	96.27±4.37	**97.4±1.45**
ELY − CHF (80 × 10,60 × 10)	71.58 ± 6.87	79.2 ± 4.19	97.06 ± 5.39	**99.13 ± 1.85**
Database (Train size, Test size)	**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 (Train size, Test size)	**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**

Table 6.3 shows the results achieved by various ranking methods in association of $1-\text{NLPELM}$ and $\text{KPCA}+1-\text{NLPELM}$. It can be seen from the Table 6.3 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 ± 1.32 for $\text{NSR}-\text{CHF}$ dataset respectively whereas an AC of 72.54±5.04 and 76.77±4.7 has been achieved with $\text{Fisher}+1-\text{NLPELM}$. With the same proposed method an AC of 95.93±1.12% and 98.44±1.4% has been achieved when the Bhattacharya ranking method is used whereas $ROC + KPCA + 1 - NLPELM$ achieved an accuracy of 97.19±0.61 and 98.24±0.96 %. The proposed method combined with Fisher, Wilcoxon, Bhattacharya, Entropy and ROC achieved an accuracy of 96.76±3.94, 95.32±1.79, 98.12±1.85, 97.11±1.51 and 98.82±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 6.4 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 6.4 reveals that proposed method does not require too much computational time to diagnose CHF disease and is easy to design too.

Table 6.4: Validation times (in seconds) compared for all six methodologies for classification of data sets namely NSR-ELY, NSR- CHF and ELY-CHF. The validation time calculated after ranking by Bhattacharya.

Methods	**NSR-ELY**	$NSR-CHF$	$ELY-CHF$
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

6.7 Summary

In this chapter, 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^{th} and 5^{th} level of HRV decomposition by MRWP have lowest $p-value\ (<0.001)$. Since $p<0.001$ attributes have greatest discernment capability so 4^{th} and 5^{th} level of MRWP attributes are considered much suitable than 1^{st}, 2nd and 3^{rd} 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.

CHAPTER: 4

Conclusion and Future Work

4.1 Conclusions

This proposal work has presented the development of new computational methods for investigation of cardiac diseases like CAD and CHF. The major conclusions from this work are summarized under.

- For the detection of cardiac diseases, three approaches have been discussed. In first approach (3rd chapter, section 4), the detection of CAD and CHF patients based on nine nonlinear features which have been extracted from HRV signal. In order to enhance the detection capability of proposed model, the dimension of nine features were reduced using GDA technique. The results of the comparative study have indicated that the GDA with OSELM binary classifier outperforms as compared to other existing techniques. The GDA with Gaussian Kernel + OS-ELM with Sine activation function achieved classification performances of 100% for NSR-CHF datasets. The GDA with RBF Kernel + OS-ELM with Sine Activation Function attained an AC of 99.34%, PPV of 99% and SP of 99.32 % for NSR-CAD. In second approach (3rd chapter, section 5), an expert system approach for the diagnosis of CAD subject has been demonstrated. The proposed ML was based on assortment of ranking approaches, GDA and 1-NELM in which five different ranking approaches and one features space reduction technique has been used to enhance the classification performance. The excellent classification accuracy, very good generalization performance and least execution time have been achieved through proposed scheme. The ROC with GDA and 1- NELM approach achieved an accuracy of 99.76±0.14, 99.87±0.12 and 100±0 for YNG-CAD, YNG-ELY and ELY-CAD groups respectively. The Fisher with GDA and 1-NELM; and Bhattacharya with GDA and 1-NELM approach achieved an accuracy of 100±0 for all considered datasets. The GDA+ 1-NELM (Sigmoid) method achieved lowest validation time 4.56, 3.962, 3.962 seconds among all other considered ML for YNG-ELD, YNG-CAD, ELD-CAD datasets. In third approach (3rd chapter, section 6), a novel algorithm has been presented for detection of CHF. This algorithm various ranking methods were combined with KPCA and 1-NLPELM classifier. The analysis results have shown 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 4th and 5th 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.

- Excellent classification accuracy achieved by proposed method makes us assured that the suggested scheme can be used for most suitable methods for the clinical setting, treatment, monitoring, diagnosis of cardiac diseases and distinguish the CAD, CHF affected patients, healthy YNG and ELY subjects.

4.2 Future Work

For training and validation of machine learning required large data size. Due to limited number of standard database available on PhysioBank and self recorded Databases, the 2000 samples of R-R interval are divided into segments of 500 R-R intervals to increase training and validating data size. Finally segments of R-R intervals are used for the feature extraction by entropy and nonlinear, MRWP methods. Less than 500 NN intervals samples, the CD, DFA and Poincare plot as SD1/SD2 ratio are less suitable for HRV analysis. Hence, for short HRV time series, some features are less suitable for HRV analysis and also segmentation of database may be affected on detection performance. Therefore, increases of number of samples have been recommended for future work. All above proposed methods can be employed for detection and classification of cardiac diseases like acute infection, arrhythmia disease detection and myocardial infarction.

References

[1] World Health Organization, "Noncommunicable Diseases. Country Perfil 2011," *World Health Organization*, 2011. http://www.who.int/nmh/publications/ncd_profiles2011/en/.

[2] B. L. Dake and C. L. Oltman, "Cardiovascular, metabolic, and coronary dysfunction in high-Fat-Fed obesity-resistant/prone rats," *Obesity*, vol. 23, no. 3, pp. 623–629, 2015, doi: 10.1002/oby.21009.

[3] S. Yusuf *et al.*, "Cardiovascular Risk and Events in 17 Low-, Middle-, and High-Income Countries," *N. Engl. J. Med.*, vol. 371, no. 9, pp. 818–827, 2014, doi: 10.1056/NEJMoa1311890.

[4] S. Mendis *et al.*, "World Health Organization definition of myocardial infarction: 2008-09 revision," *Int. J. Epidemiol.*, vol. 40, no. 1, pp. 139–146, 2011, doi: 10.1093/ije/dyq165.

[5] D. M. Salerno *et al.*, "Exercise seismocardiography for detection of coronary artery disease," *Am. J. Noninvasive Cardiol.*, vol. 6, no. 5, pp. 321–330, 1992, doi: 10.1159/000470383.

[6] A. Cassar, D. R. Holmes, C. S. Rihal, and B. J. Gersh, "Chronic coronary artery disease: Diagnosis and management," in *Mayo Clinic Proceedings*, 2009, vol. 84, no. 12, pp. 1130–1146, doi: 10.4065/mcp.2009.0391.

[7] M. R. Patel *et al.*, "Low diagnostic yield of elective coronary angiography," *N. Engl. J. Med.*, vol. 362, no. 10, pp. 886–895, 2010, doi: 10.1056/NEJMoa0907272.

[8] C. Martin-Isla *et al.*, "Image-Based Cardiac Diagnosis With Machine Learning: A Review," *Frontiers in Cardiovascular Medicine*, vol. 7. pp. 1–10, 2020, doi: 10.3389/fcvm.2020.00001.

[9] A. Moreno, J. Rodriguez, and F. Martínez, "Regional Multiscale Motion Representation for Cardiac Disease Prediction," in *2019 22nd Symposium on Image, Signal Processing and Artificial Vision, STSIVA 2019 - Conference Proceedings*, 2019, pp. 1–15, doi: 10.1109/STSIVA.2019.8730231.

[10] H. Bagher-Ebadian, H. Soltanian-Zadeh, S. Setayeshi, and S. T. Smith, "Neural network and fuzzy clustering approach for automatic diagnosis of coronary artery disease in nuclear medicine," *IEEE Trans. Nucl. Sci.*, vol. 51, no. 1 I, pp. 184–192, 2004, doi: 10.1109/TNS.2003.823047.

[11] G. Bin Huang, Q. Y. Zhu, and C. K. Siew, "Extreme learning machine: A new learning scheme of feedforward neural networks," in *IEEE International Conference on Neural Networks - Conference Proceedings*, 2004, vol. 2, pp. 985–990, doi: 10.1109/IJCNN.2004.1380068.

[12] G.-B. Huang *et al.*, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, no. 1–3, pp. 489–501, 2006, doi: 10.1016/j.neucom.2005.12.126.

[13] S. Ding, H. Zhao, Y. Zhang, X. Xu, and R. Nie, "Extreme learning machine: algorithm, theory and applications," *Artif. Intell. Rev.*, vol. 44, no. 1, pp. 103–115, 2015, doi: 10.1007/s10462-013-9405-z.

[14] Q. Leng, H. Qi, J. Miao, W. Zhu, and G. Su, “One-Class Classification with Extreme Learning Machine,” *Math. Probl. Eng.*, vol. 6, no. 2, pp. 447–461, 2015, doi: 10.1155/2015/412957.

[15] C. Gautam, A. Tiwari, and Q. Leng, “On the construction of extreme learning machine for online and offline one-class classification—An expanded toolbox,” *Neurocomputing*, vol. 261, pp. 126–143, 2017, doi: 10.1016/j.neucom.2016.04.070.

[16] Mariano Llamedo Soria, “Signal Processing for Automatic Heartbeat Classification and Patient Adaptation in the Electrocardiogram,” Universidad de Zaragoza, 2012.

[17] A. L. Goldberger *et al.*, “PhysioBank, PhysioToolkit, and PhysioNet,” *Circulation*, vol. 101, no. 23, pp. E215-20, 2000, doi: 10.1161/01.CIR.101.23.e215.

[18] G. B. Moody and R. G. Mark, “The impact of the MIT-BIH arrhythmia database,” *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3. pp. 45–50, 2001, doi: 10.1109/51.932724.

[19] A. L. Goldberger *et al.*, “PhysioBank, PhysioToolkit, and PhysioNet : Components of a New Research Resource for Complex Physiologic Signals,” *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000, doi: 10.1161/01.CIR.101.23.e215.

[20] S. Shah, G. Gnanasegaran, J. Sundberg-Cohon, and J. R. Buscombe, “The heart: Anatomy, physiology and exercise physiology,” in *Integrating Cardiology for Nuclear Medicine Physicians: A Guide to Nuclear Medicine Physicians*, 2009, pp. 3–22.

[21] M. E. Silverman and J. Willis Hurst, “Willem einthoven—the father of electrocardiography,” *Clin. Cardiol.*, vol. 15, no. 10, pp. 785–787, 1992, doi: 10.1002/clc.4960151020.

[22] **R**. S. Singh, D. J. Gelmecha and D. K. Sinha, "Expert System Based Detection and Classification of Coronary Artery Disease Using Ranking Methods and Nonlinear Attributes," *Multimedia Tools and Applications*, vol. 81(2), pp. 19723–19750, 2022, doi: 10.1007/s11042-021-11528-1.

[23] U. Rajendra Acharya, J. S. Suri, J. A. E. Spaan, and S. M. Krishnan, *Advances in cardiac signal processing*. 2007.

[24] M. Beheshti, Karthikeyan Umapathy, and Sridhar Krishnan, “Electrophysiological Cardiac Modeling: A Review,” *Crit. Rev. Biomed. Eng.*, vol. 44, no. 1–2, pp. 99–122, 2016, doi: 10.1615/CritRevBiomedEng.2016016454.

[25] Demissie J. Gelmecha , R. S. Singh, D. K. Sinha and Dereje Tekilu, " Automated Health Detection of Congestive Heart Failure Subject Using Rank Multiresolution Wavelet Packet Attributes and 1-Norm Linear Programming ELM," Multimedia Tools and Applications, vol. 81(6), pp. 19723–19750, 2022, doi: 10.1007/s11042-021-11562-z .

[26] L. Sörnmo and P. Laguna, *Bioelectrical Signal Processing in Cardiac and Neurological Applications*. 2005.

[27] P. E. Dilaveris, A. Pantazis, G. Zervopoulos, J. Kallikazaros, C. Stefanadis, and P. K. Toutouzas, "Differences in the morphology and duration between premature P waves and the preceding sinus complexes in patients with a history of paroxysmal atrial fibrillation," *Clin Cardiol*, vol. 26, no. 7, pp. 341–347, 2003, [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/12862301.

[28] D. Durrer, "Electrical aspects of human cardiac activity: A clinical-physiological approach to excitation and stimulation," *Cardiovascular Research*, vol. 2, no. 1. pp. 1–18, 1968, doi: 10.1093/cvrese/2.1.1.

[29] J. W. Hurst, "Abnormalities of the S-T segment--Part I.," *Clin. Cardiol.*, vol. 20, no. 6, pp. 511–520, 1997, doi: 10.1002/clc.4960200703.

[30] C. Merritt and S. Y. Tan, "Willem Einthoven (1860-1927): Father of electrocardiography," *Singapore Med. J.*, vol. 53, no. 1, pp. 17–18, 2012, doi: 10.1016/S0167-5273(95)90055-1.

[31] A. J. Moss, "Measurement of the QT interval and the risk associated with QTc interval prolongation: A review," *The American Journal of Cardiology*, vol. 72, no. 6. 1993, doi: 10.1016/0002-9149(93)90036-C.

[32] R. Sinha and C. Shahnaz, "An Approach for Classifying ECG Arrhythmia Based on Features Extracted from EMD and Wavelet Packet Domains," Florida International University, 2012.

[33] J. W. Hurst, "The value of using the entire New York Heart Association's classification of heart and vascular disease," *Clin. Cardiol.*, vol. 29, no. 9, pp. 415–417, 2006, doi: 10.1002/clc.4960290909.

[34] A. U. Rajendra, N. Kannathal, and S. M. Krishnan, "Comprehensive analysis of cardiac health using heart rate signals.," *Physiol. Meas.*, vol. 25, no. 5, pp. 1139–51, 2004.

[35] U. Rajendra Acharya, O. Faust, N. Adib Kadri, J. S. Suri, and W. Yu, "Automated identification of normal and diabetes heart rate signals using nonlinear measures," *Comput. Biol. Med.*, vol. 43, no. 10, pp. 1523–1529, 2013, doi: 10.1016/j.compbiomed.2013.05.024.

[36] K. Howorka *et al.*, "Effects of guided breathing on blood pressure and heart rate variability in hypertensive diabetic patients," *Auton. Neurosci. Basic Clin.*, vol. 179, no. 1–2, pp. 131–137, 2013, doi: 10.1016/j.autneu.2013.08.065.

[37] **R**. S. Singh, B. S. Saini and R. K. Sunkaria, "Time-varying spectral coherence investigation of cardiovascular signals based on energy concentration in healthy young and elderly subjects by the adaptive continuous Morlet wavelet transform," *Innovation and Research in BioMedical engineering*, vol.31 (1), pp. 1071–1081, 2018, doi: 10.1016/j.irbm.2017.12.004.

[38] T. M. Salo, J. S. Viikari, K. J. Antila, L. M. Voipio-Pulkki, J. O. Jalonen, and I. A. Välimäki, "Antihypertensive treatment and heart rate variability in diabetic patients: role of cardiac autonomic neuropathy.," *J. Auton. Nerv. Syst.*, vol. 60, no. 1–2, pp. 61–70, 1996, doi:

10.1016/0165-1838(96)00036-7.

[39] V. Magagnin *et al.*, “Heart Rate Variability and Respiratory Sinus Arrhythmia Assessment of Affective States by Bivariate Autoregressive Spectral Analysis.,” *Comput. Cardiol. (2010).*, vol. 37, no. 5737930, pp. 145–148, 2010, doi: 10.1097/MPG.0b013e3181a15ae8.Screening.

[40] M. J. Reed, C. E. Robertson, and P. S. Addison, “Heart rate variability measurements and the prediction of ventricular arrhythmias,” *QJM - Monthly Journal of the Association of Physicians*, vol. 98, no. 2. pp. 87–95, 2005, doi: 10.1093/qjmed/hci018.

[41] S. Akselrod, D. Gordon, F. Ubel, D. Shannon, A. Berger, and R. Cohen, “Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control,” *Science (80-.).*, vol. 213, no. 4504, pp. 220–222, 1981, doi: 10.1126/science.6166045.

[42] B. Pomeranz *et al.*, “Assessment of autonomic function in humans by heart rate spectral analysis.,” *Am. J. Physiol.*, vol. 248, no. 1 Pt 2, pp. H151-3, 1985, doi: 10.1152/ajpheart.1985.248.1.H151.

[43] M. Malik, J. Bigger, A. Camm, and R. Kleiger, “Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Heart rate variability: Standards of measurement, physiological interpretation, and clinical use,” *Eur. Heart J.*, vol. 17, pp. 354–381, 1996, doi: 10.1161/01.CIR.93.5.1043.

[44] S. C. Malpas and T. J. Maling, “Heart-rate variability and cardiac autonomic function in diabetes.,” *Diabetes*, vol. 39, no. 10, pp. 1177–81, 1990, [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/2210071.

[45] S. Scalvini *et al.*, “Is heart rate variability a reliable method to assess autonomic modulation in left ventricular dysfunction and heart failure? Assessment of autonomic modulation with heart rate variability.,” *Int. J. Cardiol.*, vol. 67, no. 1, pp. 9–17, 1998, [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/9880196.

[46] R. K. Sunkaria, S. C. Saxena, K. Vinod, and S. Telles, “HRV dynamics in four yogic-based meditation states using optimised AR model,” *Int. J. Biomed. Eng. Technol.*, vol. 9, no. 1, p. 45, 2012, doi: 10.1504/IJBET.2012.047370.

[47] M. V Kamath, E. L. Fallen, and R. McKelvie, “Effects of steady state exercise on the power spectrum of heart rate variability.,” *Med. Sci. Sports Exerc.*, vol. 23, no. 4, pp. 428–34, 1991, [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/2056900.

[48] R. Barbieri, J. K. Triedman, and J. P. Saul, “Heart rate control and mechanical cardiopulmonary coupling to assess central volume: a systems analysis,” *Am. J. Physiol. - Regul. Integr. Comp. Physiol.*, vol. 283, no. 5, pp. R1210–R1220, 2002, doi: 10.1152/ajpregu.00127.2002.

[49] N. Rentero, A. Cividjian, D. Trevaks, J. M. Pequignot, L. Quintin, and R. M. McAllen, “Activity patterns of cardiac vagal motoneurons in rat nucleus ambiguus,” *Am. J. Physiol. - Regul. Integr.*

Comp. Physiol., vol. 283, no. 6, pp. R1327–R1334, 2002, doi: 10.1152/ajpregu.00271.2002.

[50] Penaz Jan, *Mayer waves: History and methodology*. Automedica, 1978.

[51] G. Mulder and L. J. M. Mulder, "Information Processing and Cardiovascular Control," *Psychophysiology*, vol. 18, no. 4. pp. 392–402, 1981, doi: 10.1111/j.1469-8986.1981.tb02470.x.

[52] S. Akselrod, D. Gordon, J. B. B. Madwed, N. C. C. Snidman, D. C. C. Shannon, and R. J. J. Cohen, "Hemodynamic regulation: investigation by spectral analysis.," *Am. J. Physiol.*, vol. 249, no. 4 Pt 2, pp. H867--75, 1985,

[53] A. N. G. Braga, M. da Silva Lemos, J. R. da Silva, W. R. P. Fontes, and R. A. S. dos Santos, "Effects of angiotensins on day-night fluctuations and stress-induced changes in blood pressure.," *Am. J. Physiol. Regul. Integr. Comp. Physiol.*, vol. 282, no. 6, pp. R1663-71, 2002, doi: 10.1152/ajpregu.00583.2001.

[54] R I KITNE, "An analysis of the thermoregulatory influences on heart-rate variability," *Study Hear. Var.*, vol. 6, no. 1, pp. 81–106, 1980.

[55] J. Koh, T. E. Brown, L. A. Beightol, C. Y. Ha, and D. L. Eckberg, "Human autonomic rhythms: vagal cardiac mechanisms in tetraplegic subjects.," *J. Physiol.*, vol. 474, no. 3, pp. 483–495, 1994, doi: 10.1113/jphysiol.1994.sp020039.

[56] D. L. Eckberg, "Human Respiratory-Cardiovascular Interactions in Health and Disease," in *Cardiorespiratory and Motor Coordination*, 1985, pp. 253–258.

[57] K. Aoki, D. P. Stephens, and J. M. Johnson, "Diurnal variation in cutaneous vasodilator and vasoconstrictor systems during heat stress.," *Am. J. Physiol. Regul. Integr. Comp. Physiol.*, vol. 281, no. 2, pp. R591–R595, 2001, doi: 10.1152/ajpregu.2001.281.2.R591.

[58] M. Vornanen, A. Ryökkynen, and A. Nurmi, "Temperature-dependent expression of sarcolemmal K^+ currents in rainbow trout atrial and ventricular myocytes," *Am. J. Physiol. - Regul. Integr. Comp. Physiol.*, vol. 282, no. 4, pp. R1191–R1199, 2002, doi: 10.1152/ajpregu.00349.2001.

[59] G. A. Porter and S. A. Rivkees, "Ontogeny of humoral heart rate regulation in the embryonic mouse.," *Am. J. Physiol. Regul. Integr. Comp. Physiol.*, vol. 281, no. 2, pp. R401-7, 2001, [Online]. Available: http://ajpregu.physiology.org/content/281/2/R401.abstract.

[60] L. J. M. Mulder, "Measurement and analysis methods of heart rate and respiration for use in applied environments," *Biol. Psychol.*, vol. 34, no. 2–3, pp. 205–236, 1992, doi: 10.1016/0301-0511(92)90016-N.

[61] R. Acharya U, N. Kannathal, and S. M. Krishnan, "Comprehensive analysis of cardiac health using heart rate signals," *Physiol. Meas.*, vol. 25, no. 5, pp. 1139–1151, 2004, doi: 10.1088/0967-3334/25/5/005.

[62] D. L. Eckberg, "Human sinus arrhythmia as an index of vagal cardiac outflow.," *J. Appl. Physiol.*,

vol. 54, no. 4, pp. 961–966, 1983.

[63] MV Kamath, “Power spectral analysis of HRV,” *Biomed. Eng.*, vol. 21, no. 3, pp. 245–311, 1993.

[64] M. V Kamath and E. L. Fallen, “Power spectral analysis of heart rate variability: a noninvasive signature of cardiac autonomic function.,” *Crit. Rev. Biomed. Eng.*, vol. 21, no. 3, pp. 245–311, 1993, doi: 8243093.

[65] S. M. Kay and S. L. Marple, “Spectrum Analysis—A Modern Perspective,” *Proc. IEEE*, vol. 69, no. 11, pp. 1380–1419, 1981, doi: 10.1109/PROC.1981.12184.

[66] M. Pagani *et al.*, “Power spectral analysis of heart rate and arterial pressure variabilities as a marker of sympatho-vagal interaction in man and conscious dog,” *Circ. Res.*, vol. 59, no. 2, pp. 178–193, 1986, doi: 10.1161/01.RES.59.2.178.

[67] G. Parati, J. P. Saul, M. Di Rienzo, and G. Mancia, “Spectral Analysis of Blood Pressure and Heart Rate Variability in Evaluating Cardiovascular Regulation : A Critical Appraisal,” *Hypertension*, vol. 25, no. 6, pp. 1276–1286, 1995, doi: 10.1161/01.HYP.25.6.1276.

[68] World Health Organization, “Cardiovascular diseases (CVDs),” *Fact sheet N°317, September 2012*, 2012. http://www.who.int/mediacentre/factsheets/fs317/en/%5Cnhttp://www.who.int/mediacentre/factsheets/fs317/en/index.html.

[69] L. Stankovic, “An analysis of some time-frequency and time-scale distributions,” *Ann. Des Télécommunications*, vol. 49, no. 9–10, pp. 505–517, 1994, doi: 10.1007/BF02999442.

[70] I. Orović, M. Orlandić, S. Stanković, and Z. Uskoković, “A virtual instrument for time-frequency analysis of signals with highly nonstationary instantaneous frequency,” *IEEE Trans. Instrum. Meas.*, vol. 60, no. 3, pp. 791–803, 2011, doi: 10.1109/TIM.2010.2060227.

[71] L. Stankovic, “Highly concentrated time-frequency distributions: Pseudo quantum signal representation,” *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 543–551, 1997, doi: 10.1109/78.558467.

[72] D. Gabor, “Theory of communication,” *J. Institute of Electr. Eng.*, vol. 93. pp. 429–457, 1946, doi: 10.1049/ji-3-2.1946.0074.

[73] M. Wacker and H. Witte, “Time-frequency techniques in biomedical signal analysis: A tutorial review of similarities and differences,” *Methods Inf. Med.*, vol. 52, no. 4, pp. 279–296, 2013, doi: 10.3414/ME12-01-0083.

[74] I. Daubechies, “The wavelet transform, time-frequency localization and signal analysis,” *IEEE Trans. Inf. Theory*, vol. 36, no. 5, pp. 961–1005, 1990, doi: 10.1109/18.57199.

[75] R. G. Stockwell, L. Mansinha, and R. P. Lowe, “Localization of the complex spectrum: The S transform,” *IEEE Trans. Signal Process.*, vol. 44, no. 4, pp. 998–1001, 1996, doi:

10.1109/78.492555.

[76] R. G. Stockwell, “A basis for efficient representation of the S-transform,” *Digit. Signal Process. A Rev. J.*, vol. 17, no. 1, pp. 371–393, 2007, doi: 10.1016/j.dsp.2006.04.006.

[77] S. Ventosa, C. Simon, M. Schimmel, J. J. Danobeitia, and A. Manuel, “The S -Transform From a Wavelet Point of View,” *Signal Process. IEEE Trans.*, vol. 56, no. 7, pp. 2771–2780, 2008, doi: 10.1109/TSP.2008.917029.

[78] W. Lin and M. Xiaofeng, “An adaptive Generalized S-transform for instantaneous frequency estimation,” *Signal Processing*, vol. 91, no. 8, pp. 1876–1886, 2011, doi: 10.1016/j.sigpro.2011.02.010.

[79] R. H. Clayton and A. Murray, “Estimation of the ECG signal spectrum during ventricular fibrillation using the fast Fourier transform and maximum entropy methods,” in *Computers in Cardiology*, 1993, pp. 2–10, doi: 10.1109/CIC.1993.378299.

[80] A. Govindan, G. Deng, and John Power, “Electrogram analysis during atrial fibrillation using wavelet and neural network techniques,” in *Wavelet Applications in Signal and Image Processing*, 1997, pp. 1–10, doi: 10.1117/12.279706.

[81] L. Khadra, a S. al-Fahoum, and H. al-Nashash, “Detection of life-threatening cardiac arrhythmias using the wavelet transformation.,” *Med. Biol. Eng. Comput.*, vol. 35, no. 6, pp. 626–632, 1997, doi: 10.1007/BF02510970.

[82] X. S. Zhang, Y. S. Zhu, N. V. Thakor, Z. M. Wang, and Z. Z. Wang, “Modeling the relationship between concurrent epicardial action potentials and bipolar electrograms,” *IEEE Trans. Biomed. Eng.*, vol. 46, no. 4, pp. 365–376, 1999, doi: 10.1109/10.752933.

[83] A. S. Al-Fahoum and I. Howitt, “Combined wavelet transformation and radial basis neural networks for classifying life-threatening cardiac arrhythmias,” *Med. Biol. Eng. Comput.*, vol. 37, no. 5, pp. 566–573, 1999, doi: 10.1007/BF02513350.

[84] D. Morlet, F. Peyrin, P. Desseigne, P. Touboul, and P. Rubel, “Wavelet analysis of high-resolution signal-averaged ECGs in postinfarction patients,” *J. Electrocardiol.*, vol. 26, no. 4, pp. 311–320, 1993, doi: 10.1016/0022-0736(93)90052-F.

[85] A. Englund, K. Hnatkova, P. Kulakowski, P. M. Elliot, W. J. McKenna, and M. Malik, “Wavelet decomposition analysis of the signal averaged electrocardiogram used for risk stratification of patients with hypertrophic cardiomyopathy.,” *Eur. Heart J.*, vol. 19, no. 9, pp. 1383–90, 1998, doi: 10.1053/euhj.1998.1082.

[86] P. S. Addison, “Wavelet transforms and the ECG: A review,” *Physiological Measurement*, vol. 26, no. 5. 2005, doi: 10.1088/0967-3334/26/5/R01.

[87] H. Dickhaus and H. Heinrich, “Classifying Biosignals with wavelet networks : A method for

nonivasive diagnosis," *IEEE Engineering in Medicine and Biology Magazine*, vol. 15, no. 5. pp. 103–111, 1996, doi: 10.1109/51.537066.

[88] C. H. Lin and C. H. Wang, "Adaptive wavelet networks for power-quality detection and discrimination in a power system," *IEEE Trans. Power Deliv.*, vol. 21, no. 3, pp. 1106–1113, 2006, doi: 10.1109/TPWRD.2006.874105.

[89] C. H. Lin, Y. C. Du, and T. Chen, "Adaptive wavelet network for multiple cardiac arrhythmias recognition," *Expert Syst. Appl.*, vol. 34, no. 4, pp. 2601–2611, 2008, doi: 10.1016/j.eswa.2007.05.008.

[90] J. P. Mart??nez, R. Almeida, S. Olmos, A. P. Rocha, and P. Laguna, "A Wavelet-Based ECG Delineator Evaluation on Standard Databases," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 4, pp. 570–581, 2004, doi: 10.1109/TBME.2003.821031.

[91] N. V. Thakor, Y. S. Zhu, and K. Y. Pan, "Ventricular Tachycardia and Fibrillation Detection by a Sequential Hypothesis Testing Algorithm," *IEEE Trans. Biomed. Eng.*, vol. 37, no. 9, pp. 837–843, 1990, doi: 10.1109/10.58594.

[92] N. V. Thakor, A. Natarajan, and G. F. Tomaselli, "Multiway Sequential Hypothesis Testing for Tachyarrhythmia Discrimination," *IEEE Trans. Biomed. Eng.*, vol. 41, no. 5, pp. 480–487, 1994, doi: 10.1109/10.293223.

[93] R. H. Clayton, A. Murray, and R. W. F. Campbell, "Comparison of four techniques for recognition of ventricular fibrillation from the surface ECG," *Med. Biol. Eng. Comput.*, vol. 31, no. 2, pp. 111–117, 1993, doi: 10.1007/BF02446668.

[94] R. Clayton, A. Murray, and R. Campbell, "Recognition of ventricular fibrillation using neural networks.," *Med Biol Eng Comput.*, vol. 32, no. 217–220, 1994.

[95] S. Barro, R. Ruiz, D. Cabello, and J. Mira, "Algorithmic sequential decision-making in the frequency domain for life threatening ventricular arrhythmias and imitative artefacts: a diagnostic system," *J. Biomed. Eng.*, vol. 11, no. 4, pp. 320–328, 1989, doi: 10.1016/0141-5425(89)90067-8.

[96] X. S. Zhang, Y. S. Zhu, N. V. Thakor, and Z. Z. Wang, "Detecting ventricular tachycardia and fibrillation by complexity measure," *IEEE Trans. Biomed. Eng.*, vol. 46, no. 5, pp. 548–555, 1999, doi: 10.1109/10.759055.

[97] Y. Wang, Y. S. Zhu, N. V. Thakor, and Y. H. Xu, "A short-time multifractal approach for arrhythmia detection based on fuzzy neural network," *IEEE Trans. Biomed. Eng.*, vol. 48, no. 9, pp. 989–995, 2001, doi: 10.1109/10.942588.

[98] Y. Zhan, D. Halliday, P. Jiang, X. Liu, and J. Feng, "Detecting time-dependent coherence between non-stationary electrophysiological signals-A combined statistical and time-frequency approach," *J. Neurosci. Methods*, vol. 156, no. 1–2, pp. 322–332, 2006, doi: 10.1016/j.jneumeth.2006.02.013.

[99] J. Benesty, J. Chen, and Y. Huang, “A generalized MVDR spectrum,” *IEEE Signal Process. Lett.*, vol. 12, no. 12, pp. 827–830, 2005, doi: 10.1109/LSP.2005.859517.

[100] H. K. Lackner, I. Papousek, J. J. Batzel, A. Roessler, H. Scharfetter, and H. Hinghofer-Szalkay, “Phase synchronization of hemodynamic variables and respiration during mental challenge,” *Int. J. Psychophysiol.*, vol. 79, no. 3, pp. 401–409, 2011, doi: 10.1016/j.ijpsycho.2011.01.001.

[101] D. S. Fonseca, A. Beda, A. M. F. L. Miranda de Sá, and D. M. Simpson, “Gain and coherence estimates between respiration and heart-rate: Differences between inspiration and expiration,” *Auton. Neurosci. Basic Clin.*, vol. 178, no. 1–2, pp. 89–95, 2013, doi: 10.1016/j.autneu.2013.03.015.

[102] G. Valenza, A. Lanata, and E. P. Scilingo, “Oscillations of heart rate and respiration synchronize during affective visual stimulation,” *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 4, pp. 683–690, 2012, doi: 10.1109/TITB.2012.2197632.

[103] C. Gallet and C. Julien, “The significance threshold for coherence when using the Welch’s periodogram method: Effect of overlapping segments,” *Biomed. Signal Process. Control*, vol. 6, no. 4, pp. 405–409, 2011, doi: 10.1016/j.bspc.2010.11.004.

[104] K. Yoshino and K. Matsuoka, “Causal coherence analysis of heart rate variability and systolic blood pressure variability under mental arithmetic task load,” *Biol. Psychol.*, vol. 69, no. 2, pp. 217–227, 2005, doi: 10.1016/j.biopsycho.2004.07.001.

[105] F. Mormann, K. Lehnertz, P. David, and C. E. Elger, “Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients,” *Phys. D Nonlinear Phenom.*, vol. 144, no. 3, pp. 358–369, 2000, doi: 10.1016/S0167-2789(00)00087-7.

[106] K. Keissar, L. R. Davrath, and S. Akselrod, “Coherence analysis between respiration and heart rate variability using continuous wavelet transform,” *Phil Trans R Soc A*, vol. 367, no. 1892, pp. 1393–1406, 2009, doi: 10.1098/rsta.2008.0273.

[107] M. Hassan, J. Terrien, B. Karlsson, and C. Marque, “Application of wavelet coherence to the detection of uterine electrical activity synchronization in labor,” *IRBM*, vol. 31, no. 3, pp. 182–187, 2010, doi: 10.1016/j.irbm.2009.12.004.

[108] A. Grinsted, J. C. Moore, and S. Jevrejeva, “Application of the cross wavelet transform and wavelet coherence to geophysical time series,” *Nonlinear Process. Geophys.*, vol. 11, no. 5/6, pp. 561–566, 2004, doi: 10.5194/npg-11-561-2004.

[109] V. Di Virgilio, R. Barbieri, L. Mainardi, S. Strano, and S. Cerutti, “A multivariate time-variant AR method for the analysis of heart rate and arterial blood pressure.,” *Med. Eng. Phys.*, vol. 19, no. 2, pp. 109–24, 1997, [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/9203145.

[110] S. Pola, A. Macerata, M. Emdin, and C. Marchesi, “Estimation of the power spectral density in

nonstationary cardiovascular time series: Assessing the role of the time-frequency representations (TFR)," *IEEE Trans. Biomed. Eng.*, vol. 43, no. 1, pp. 46–59, 1996, doi: 10.1109/10.477700.

[111] G. Matz, H. Bolcskei, and F. Hlawatsch, "Time-frequency foundations of communications: Concepts and tools," *IEEE Signal Process. Mag.*, vol. 30, no. 6, pp. 87–96, 2013, doi: 10.1109/MSP.2013.2269702.

[112] K. Keissar and L. Davrath, "Time-frequency wavelet transform coherence of cardio-respiratory signals during exercise," *Comput. Cardiol.*, vol. 33, pp. 733–736, 2008, [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4511956%5Cnpapers2://publication/uuid/88632949-171C-40EA-A0FC-F0C8B6F052B8.

[113] N. Östlund, O. B. Suhr, and U. Wiklund, "Wavelet coherence detects non-autonomic heart rate fluctuations in familial amyloidotic polyneuropathy," in *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, 2007, pp. 4660–4662, doi: 10.1109/IEMBS.2007.4353379.

[114] B. Boashash, G. Azemi, and N. Ali Khan, "Principles of time-frequency feature extraction for change detection in non-stationary signals: Applications to newborn EEG abnormality detection," *Pattern Recognit.*, vol. 48, no. 3, pp. 616–627, 2015, doi: 10.1016/j.patcog.2014.08.016.

[115] B. Boashash, N. A. Khan, and T. Ben-Jabeur, "Time-frequency features for pattern recognition using high-resolution TFDs: A tutorial review," *Digit. Signal Process. A Rev. J.*, vol. 40, no. 1, pp. 1–30, 2015, doi: 10.1016/j.dsp.2014.12.015.

[116] D. Hoyer *et al.*, "Nonlinear analysis of heart rate and respiratory dynamics," *IEEE Engineering in Medicine and Biology Magazine*, vol. 16, no. 1. pp. 31–39, 1997, doi: 10.1109/51.566150.

[117] H. V. Huikuri, T. H. Mäkikallio, C. K. Peng, a. L. Goldberger, U. Hintze, and M. Møller, "Fractal correlation properties of R-R interval dynamics and mortality in patients with depressed left ventricular function after an acute myocardial infarction.," *Circulation*, vol. 101, no. 1, pp. 47–53, 2000, doi: 10.1161/01.CIR.101.1.47.

[118] A. Porta, S. Guzzetti, R. Furlan, T. Gnecchi-Ruscone, N. Montano, and A. Malliani, "Complexity and nonlinearity in short-term heart period variability: Comparison of methods based on local nonlinear prediction," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 1, pp. 94–106, 2007, doi: 10.1109/TBME.2006.883789.

[119] L. Faes, K. H. Chon, and G. Nollo, "A method for the time-varying nonlinear prediction of complex nonstationary biomedical signals," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 2, pp. 205–209, 2009, doi: 10.1109/TBME.2008.2008726.

[120] A. Bianchi, B. Bontempi, S. Cerutti, P. Gianoglio, G. Comi, and M. G. Natali Sora, "Spectral analysis of heart rate variability signal and respiration in diabetic subjects," *Med. Biol. Eng.*

Comput., vol. 28, no. 3, pp. 205–211, 1990, doi: 10.1007/BF02442668.

[121] R. K. Sunkaria, "Recent Trends in Nonlinear Methods of HRV Analysis : A Review," *Eng. Technol.*, vol. 75, no. 3, pp. 566–571, 2011.

[122] S. M. Pincus, "Approximate entropy as a measure of system complexity," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 88, no. 6, pp. 2297–2301, 1991, doi: 10.1073/pnas.88.6.2297.

[123] H. Azami, A. Fernández, and J. Escudero, "Refined multiscale fuzzy entropy based on standard deviation for biomedical signal analysis," *Med. Biol. Eng. Comput.*, vol. 55, no. 11, pp. 2037–2052, 2017, doi: 10.1007/s11517-017-1647-5.

[124] A. Kraskov, H. Stögbauer, and P. Grassberger, "Estimating mutual information," *Phys. Rev. E - Stat. Physics, Plasmas, Fluids, Relat. Interdiscip. Top.*, vol. 69, no. 6, p. 16, 2004, doi: 10.1103/PhysRevE.69.066138.

[125] M. Zanin, L. Zunino, O. A. Rosso, and D. Papo, "Permutation entropy and its main biomedical and econophysics applications: A review," *Entropy*, vol. 14, no. 8. pp. 1553–1577, 2012, doi: 10.3390/e14081553.

[126] H. Azami and J. Escudero, "Improved multiscale permutation entropy for biomedical signal analysis: Interpretation and application to electroencephalogram recordings," *Biomed. Signal Process. Control*, vol. 23, pp. 28–41, 2016, doi: 10.1016/j.bspc.2015.08.004.

[127] R. R. Coifman, Y. Meyer, and M. V Wickerhauser, "Wavelet Analysis and Signal Processing," in *Wavelets and Their Applications*, 1992, pp. 153–178.

[128] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy.," *Am. J. Physiol. Heart Circ. Physiol.*, vol. 278, no. 6, pp. H2039–H2049, 2000, doi: 10.1103/physreva.29.975.

[129] P. W. Kamen, H. Krum, and A. M. Tonkin, "Poincaré plot of heart rate variability allows quantitative display of parasympathetic nervous activity in humans.," *Clin. Sci. (Lond).*, vol. 91, no. 2, pp. 201–8, 1996, doi: 10.1042/cs0910201.

[130] H. Shaobo, S. Kehui, and W. Huihai, "Modified multiscale permutation entropy algorithm and its application for multiscroll chaotic systems," *Complexity*, vol. 21, no. 5, pp. 52–58, 2016, doi: 10.1002/cplx.21629.

[131] M. Costa, A. L. Goldberger, and C.-K. Peng, "Multiscale Entropy Analysis of Complex Physiologic Time Series," *Phys. Rev. Lett.*, vol. 89, no. 6, p. 068102, 2002, doi: 10.1103/PhysRevLett.89.068102.

[132] D. Kugiumtzis and A. Tsimpiris, "Measures of Analysis of Time Series (MATS):," *J. Stat. Softw.*, vol. 33, no. 5, 2010.

[133] C. A. Boneau, "The effects of violations of assumptions underlying the t test," *Psychol. Bull.*, vol.

57, no. 1, pp. 49–64, 1960, doi: 10.1037/h0041412.

[134] M. Pfannkuch, “Comparing Box Plot Distributions: A Teacher’s Reasoning,” *Stat. Educ. Res. J.*, vol. 5, no. 2, pp. 27–45, 2006, doi: 10.11120/msor.2004.04030058.

[135] Y. H. Hu, S. Palreddy, and W. J. Tompkins, “A patient-adaptable ECG beat classifier using a mixture of experts approach,” *IEEE Trans. Biomed. Eng.*, vol. 44, no. 9, pp. 891–900, 1997, doi: 10.1109/10.623058.

[136] M. Llamedo, J. P. Martínez, A. Inst, and I. I. S. Aragón, “Analysis of a Semiautomatic Algorithm for ECG Heartbeat Classification,” *Comput. Cardiol. (CinC), 2011*, pp. 137–140, 2011.

[137] W. Jiang and S. G. Kong, “Block-based neural networks for personalized ECG signal classification,” *IEEE Trans. Neural Networks*, vol. 18, no. 6, pp. 1750–1761, 2007, doi: 10.1109/TNN.2007.900239.

[138] P. deChazal, M. O’Dwyer, and R. B. Reilly, “Automatic Classification of Heartbeats Using ECG Morphology and Heartbeat Interval Features,” *IEEE Trans. Biomed. Eng.*, vol. 51, no. 7, pp. 1196–1206, 2004, doi: 10.1109/TBME.2004.827359.

[139] M. Lagerholm and G. Peterson, “Clustering ECG complexes using hermite functions and self-organizing maps,” *IEEE Trans. Biomed. Eng.*, vol. 47, no. 7, pp. 838–848, 2000, doi: 10.1109/10.846677.

[140] A. Kampouraki, G. Manis, and C. Nikou, “Heartbeat time series classification with support vector machines,” in *IEEE Transactions on Information Technology in Biomedicine*, 2009, vol. 13, no. 4, pp. 512–518, doi: 10.1109/TITB.2008.2003323.

[141] S. Dua, X. Du, S. Vinitha Sree, and V. I. Thajudin Ahamed, “Novel classification of coronary artery disease using heart rate variability analysis,” *J. Mech. Med. Biol.*, vol. 12, no. 4, p. 1240017, 2012, doi: 10.1142/S0219519412400179.

[142] B. M. Asl, S. K. Setarehdan, and M. Mohebbi, “Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal,” *Artif. Intell. Med.*, vol. 44, no. 1, pp. 51–64, 2008, doi: 10.1016/j.artmed.2008.04.007.

[143] I. Christov, G. Gómez-Herrero, V. Krasteva, I. Jekova, A. Gotchev, and K. Egiazarian, “Comparative study of morphological and time-frequency ECG descriptors for heartbeat classification,” *Med. Eng. Phys.*, vol. 28, no. 9, pp. 876–887, 2006, doi: 10.1016/j.medengphy.2005.12.010.

[144] T. Ince, S. Kiranyaz, and M. Gabbou, “A generic and robust system for automated patient-specific classification of ECG signals,” *IEEE Trans. Biomed. Eng.*, vol. 56, no. 5, pp. 1415–1426, 2009, doi: 10.1109/TBME.2009.2013934.

[145] G. Bin Huang, D. H. Wang, and Y. Lan, “Extreme learning machines: A survey,” *Int. J. Mach.*

Learn. Cybern., vol. 2, no. 2, pp. 107–122, 2011, doi: 10.1007/s13042-011-0019-y.

[146] Z. X. Chen, H. Y. Zhu, and Y. G. Wang, "A modified extreme learning machine with sigmoidal activation functions," *Neural Comput. Appl.*, vol. 22, no. 3–4, pp. 541–550, 2013, doi: 10.1007/s00521-012-0860-2.

[147] B. Frénay and M. Verleysen, "Using SVMs with randomised feature spaces : an extreme learning approach," *ESANN 2010 proceedings, Eur. Symp. Artif. Neural Networks - Comput. Intell. Mach. Learn.*, no. April, pp. 28–30, 2010.

[148] Guang-Bin Huang, Hongming Zhou, Xiaojian Ding, and Rui Zhang, "Extreme Learning Machine for Regression and Multiclass Classification," *IEEE Trans. Syst. Man, Cybern. Part B*, vol. 42, no. 2, pp. 513–529, 2011, doi: 10.1109/tsmcb.2011.2168604.

[149] L. Zhang and W. Zhou, "On the sparseness of 1-norm support vector machines," *Neural Networks*, vol. 23, no. 3, pp. 373–385, 2010, doi: 10.1016/j.neunet.2009.11.012.

[150] O. L. Mangasarian, "Exact 1-Norm Support Vector Machines via Unconstrained Convex Differentiable Minimization," *J. Mach. Learn. Res.*, vol. 7, pp. 1517–1530, 2006, [Online]. Available: http://dl.acm.org/citation.cfm?id=1248603.

[151] G. Fung and O. L. Mangasarian, "Finite Newton method for Lagrangian support vector machine classification," *Neurocomputing*, vol. 55, no. 1–2, pp. 39–55, 2003, doi: 10.1016/S0925-2312(03)00379-5.

[152] P. Jerčić, C. Sennersten, and C. Lindley, "Modeling cognitive load and physiological arousal through pupil diameter and heart rate," *Multimed. Tools Appl.*, vol. 79, no. 5–6, pp. 3145–3159, 2020, doi: 10.1007/s11042-018-6518-z.

[153] S. Balasundaram, D. Gupta, and Kapil, "1-Norm extreme learning machine for regression and multiclass classification using Newton method," *Neurocomputing*, vol. 128, pp. 4–14, 2014, doi: 10.1016/j.neucom.2013.03.051.

[154] R. S. Singh, B. S. Saini, and R. K. Sunkaria, "Classification of Cardiac Heart Disease Using Reduced Chaos Features and 1-Norm Linear Programming Extreme Learning Machine," *Int. J. Multiscale Comput. Eng.*, vol. 16, no. 5, pp. 465–486, 2018, doi: 10.1615/intjmultcompeng.2018026587.

[155] F. Marmolejo-Ramos and T. Siva Tian, "The shifting boxplot. A boxplot based on essential summary statistics around the mean.," *Int. J. Psychol. Res.*, vol. 3, no. 1, pp. 37–45, 2010, doi: 10.21500/20112084.823.

[156] P. Grassberger and I. Procaccia, "Measuring the strangeness of strange attractors," *Phys. D Nonlinear Phenom.*, vol. 9, no. 1–2, pp. 189–208, 1983, doi: 10.1016/0167-2789(83)90298-1.

[157] M. Vollmer, "A robust, simple and reliable measure of heart rate variability using relative RR

intervals," *Comput. Cardiol. (2010).*, vol. 42, no. 6, pp. 609–612, 2016, doi: 10.1109/CIC.2015.7410984.

[158] R. Yan, Y. Liu, and R. X. Gao, "Permutation entropy: A nonlinear statistical measure for status characterization of rotary machines," in *Mechanical Systems and Signal Processing*, 2012, vol. 29, pp. 474–484, doi: 10.1016/j.ymssp.2011.11.022.

[159] G. Manis, M. Aktaruzzaman, and R. Sassi, "Bubble entropy: An entropy almost free of parameters," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 11, pp. 2711–2718, 2017, doi: 10.1109/TBME.2017.2664105.

[160] R. Yan, Q. Zheng, and W. Peng, "Multi-scale entropy and Renyi cross entropy based traffic anomaly detection," in *2008 11th IEEE Singapore International Conference on Communication Systems, ICCS 2008*, 2008, pp. 554–558, doi: 10.1109/ICCS.2008.4737245.

[161] R. Kalpana, M. Chitra, and G. Ratna-Sagari, "A Case Study Analysis of EEG Signals under Conditions of Cognition," *Asian J. Med. Sci.*, vol. 7, no. 4, pp. 41–49, 2015, [Online]. Available: http://www.airitilibrary.com/Publication/alDetailedMesh?docid=20408773-201510-201512080004-201512080004-41-49.

[162] O. W. Samuel, G. M. Asogbon, A. K. Sangaiah, P. Fang, and G. Li, "An integrated decision support system based on ANN and Fuzzy_AHP for heart failure risk prediction," *Expert Syst. Appl.*, vol. 68, no. 2, pp. 163–172, 2017, doi: 10.1016/j.eswa.2016.10.020.

[163] R. Mahajan, T. Viangteeravat, and O. Akbilgic, "Improved detection of congestive heart failure via probabilistic symbolic pattern recognition and heart rate variability metrics," *Int. J. Med. Inform.*, vol. 108, no. 1, pp. 55–63, 2017, doi: 10.1016/j.ijmedinf.2017.09.006.

[164] S. Bashir, U. Qamar, and F. H. Khan, "IntelliHealth: A medical decision support application using a novel weighted multi-layer classifier ensemble framework," *J. Biomed. Inform.*, vol. 59, no. 1, pp. 185–200, 2016, doi: 10.1016/j.jbi.2015.12.001.

[165] S. N. Yu and M. Y. Lee, "Bispectral analysis and genetic algorithm for congestive heart failure recognition based on heart rate variability," *Comput. Biol. Med.*, vol. 42, no. 8, pp. 816–825, 2012, doi: 10.1016/j.compbiomed.2012.06.005.

[166] D. Tay, C. L. Poh, and R. I. Kitney, "A novel neural-inspired learning algorithm with application to clinical risk prediction," *J. Biomed. Inform.*, vol. 54, pp. 305–314, 2015, doi: 10.1016/j.jbi.2014.12.014.

[167] P. Pławiak, "Novel methodology of cardiac health recognition based on ECG signals and evolutionary-neural system," *Expert Syst. Appl.*, vol. 92, no. 2, pp. 334–349, 2018, doi: 10.1016/j.eswa.2017.09.022.

[168] A. Davari Dolatabadi, S. E. Z. Khadem, and B. M. Asl, "Automated diagnosis of coronary artery

disease (CAD) patients using optimized SVM," *Comput. Methods Programs Biomed.*, vol. 138, no. 2, pp. 117–126, 2017, doi: 10.1016/j.cmpb.2016.10.011.

[169] A. Mustaqeem, S. M. Anwar, A. R. Khan, and M. Majid, "A statistical analysis based recommender model for heart disease patients," *Int. J. Med. Inform.*, vol. 108, no. 2, pp. 134–145, 2017, doi: 10.1016/j.ijmedinf.2017.10.008.

[170] Z. Arabasadi, R. Alizadehsani, M. Roshanzamir, H. Moosaei, and A. A. Yarifard, "Computer aided decision making for heart disease detection using hybrid neural network-Genetic algorithm," *Comput. Methods Programs Biomed.*, vol. 141, no. 2, pp. 19–26, 2017, doi: 10.1016/j.cmpb.2017.01.004.

[171] R. J. Martis, U. R. Acharya, K. M. Mandana, A. K. Ray, and C. Chakraborty, "Application of principal component analysis to ECG signals for automated diagnosis of cardiac health," *Expert Syst. Appl.*, vol. 39, no. 14, pp. 11792–11800, 2012, doi: 10.1016/j.eswa.2012.04.072.

[172] N. C. Long, P. Meesad, and H. Unger, "A highly accurate firefly based algorithm for heart disease prediction," *Expert Syst. Appl.*, vol. 42, no. 21, pp. 8221–8231, 2015, doi: 10.1016/j.eswa.2015.06.024.

[173] V. K. Sudarshan *et al.*, "Automated diagnosis of congestive heart failure using dual tree complex wavelet transform and statistical features extracted from 2 s of ECG signals," *Comput. Biol. Med.*, vol. 83, no. 1, pp. 48–58, 2017, doi: 10.1016/j.compbiomed.2017.01.019.

[174] M. Tayefi *et al.*, "hs-CRP is strongly associated with coronary heart disease (CHD): A data mining approach using decision tree algorithm," *Comput. Methods Programs Biomed.*, vol. 141, pp. 105–109, 2017, doi: 10.1016/j.cmpb.2017.02.001.

[175] K. H. Boon, M. Khalil-Hani, and M. B. Malarvili, "Paroxysmal atrial fibrillation prediction based on HRV analysis and non-dominated sorting genetic algorithm III," *Comput. Methods Programs Biomed.*, vol. 153, pp. 171–184, 2018, doi: 10.1016/j.cmpb.2017.10.012.

[176] Purushottam, K. Saxena, and R. Sharma, "Efficient Heart Disease Prediction System," in *Procedia Computer Science*, 2016, vol. 85, pp. 962–969, doi: 10.1016/j.procs.2016.05.288.

[177] D. Pal, K. M. Mandana, S. Pal, D. Sarkar, and C. Chakraborty, "Fuzzy expert system approach for coronary artery disease screening using clinical parameters," *Knowledge-Based Syst.*, vol. 36, no. 2, pp. 162–174, 2012, doi: 10.1016/j.knosys.2012.06.013.

[178] H. Fujita *et al.*, "Sudden cardiac death (SCD) prediction based on nonlinear heart rate variability features and SCD index," *Appl. Soft Comput. J.*, vol. 43, pp. 510–519, 2016, doi: 10.1016/j.asoc.2016.02.049.

[179] G. Altan, Y. Kutlu, and N. Allahverdi, "A new approach to early diagnosis of congestive heart failure disease by using Hilbert–Huang transform," *Comput. Methods Programs Biomed.*, vol.

137, pp. 23–34, 2016, doi: 10.1016/j.cmpb.2016.09.003.

[180] Y. Zheng, X. Guo, J. Qin, and S. Xiao, “Computer-assisted diagnosis for chronic heart failure by the analysis of their cardiac reserve and heart sound characteristics,” *Comput. Methods Programs Biomed.*, vol. 122, no. 3, pp. 372–383, 2015, doi: 10.1016/j.cmpb.2015.09.001.

[181] F. Miao, Y. P. Cai, Y. X. Zhang, X. M. Fan, and Y. Li, “Predictive modeling of hospital mortality for patients with heart failure by using an improved random survival forest,” *IEEE Access*, vol. 6, no. 1, pp. 7244–7253, 2018, doi: 10.1109/ACCESS.2018.2789898.

[182] B. Jin, C. Che, Z. Liu, S. Zhang, X. Yin, and X. Wei, “Predicting the Risk of Heart Failure with EHR Sequential Data Modeling,” *IEEE Access*, vol. 6, no. 2, pp. 9256–9261, 2018, doi: 10.1109/ACCESS.2017.2789324.

[183] S. A. Mokeddem, “A fuzzy classification model for myocardial infarction risk assessment,” *Appl. Intell.*, vol. 48, no. 5, pp. 1233–1250, 2018, doi: 10.1007/s10489-017-1102-1.

[184] J. K. Kim, J. S. Lee, D. K. Park, Y. S. Lim, Y. H. Lee, and E. Y. Jung, “Adaptive mining prediction model for content recommendation to coronary heart disease patients,” *Cluster Comput.*, vol. 17, no. 3, pp. 881–891, 2014, doi: 10.1007/s10586-013-0308-1.

[185] J. W. Kantelhardt, S. A. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde, and H. E. Stanley, “Multifractal detrended fluctuation analysis of nonstationary time series,” *Phys. A Stat. Mech. its Appl.*, vol. 316, no. 1–4, pp. 87–114, 2002, doi: 10.1016/S0378-4371(02)01383-3.

[186] P. Castiglioni, D. Lazzeroni, P. Coruzzi, and A. Faini, “Multifractal-Multiscale Analysis of Cardiovascular Signals: A DFA-Based Characterization of Blood Pressure and Heart-Rate Complexity by Gender,” *Complexity*, vol. 2018, no. 1, pp. 1–14, 2018, doi: 10.1155/2018/4801924.

[187] R. Nagarajan, “Quantifying physiological data with Lempel-Ziv complexity - Certain issues,” *IEEE Trans. Biomed. Eng.*, vol. 49, no. 11, pp. 1371–1373, 2002, doi: 10.1109/TBME.2002.804582.

[188] M. Aboy, R. Hornero, D. Abasolo, and D. Álvarez, “Interpretation of the Lempel-Ziv complexity measure in the context of biomedical signal analysis,” *IEEE Trans. Biomed. Eng.*, vol. 53, no. 11, pp. 2282–2288, 2006, doi: 10.1109/TBME.2006.883696.

[189] A. Lempel and J. Ziv, “On the Complexity of Finite Sequences,” *IEEE Trans. Inf. Theory*, vol. 22, no. 1, pp. 75–81, 1976, doi: 10.1109/TIT.1976.1055501.

[190] H. Azami, M. Rostaghi, D. Abasolo, and J. Escudero, “Refined Composite Multiscale Dispersion Entropy and its Application to Biomedical Signals,” *IEEE Trans. Biomed. Eng.*, vol. 64, no. 12, pp. 2872–2879, 2017, doi: 10.1109/TBME.2017.2679136.

[191] F. S. Fazan, F. Brognara, R. F. Junior, and L. O. Junior, “Changes in the Complexity of Heart Rate

Variability with Exercise Training Measured by Multiscale Entropy-Based Measurements," *Entropy*, vol. 20, no. 8, pp. 1–10, 2018.

[192] T. Costa, D. Galati, and E. Rognoni, "The Hurst exponent of cardiac response to positive and negative emotional film stimuli using wavelet," *Auton. Neurosci. Basic Clin.*, vol. 151, no. 2, pp. 183–185, 2009, doi: 10.1016/j.autneu.2009.08.011.

[193] K. Liu, H. Wang, and J. Xiao, "The multivariate largest lyapunov exponent as an age-related metric of quiet standing balance," *Comput. Math. Methods Med.*, vol. 2015, no. 1, pp. 1–12, 2015, doi: 10.1155/2015/309756.

[194] Z. K. Gao, Y. X. Yang, P. C. Fang, Y. Zou, C. Y. Xia, and M. Du, "Multiscale complex network for analyzing experimental multivariate time series," *Epl*, vol. 109, no. 3, pp. 1–8, 2015, doi: 10.1209/0295-5075/109/30005.

[195] I. Babaoglu, O. Findik, and E. Ulker, "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.*, vol. 37, no. 4, pp. 3177–3183, 2010, doi: 10.1016/j.eswa.2009.09.064.

[196] D. Giri *et al.*, "Automated diagnosis of Coronary Artery Disease affected patients using LDA, PCA, ICA and Discrete Wavelet Transform," *Knowledge-Based Syst.*, vol. 37, pp. 274–282, 2013, doi: 10.1016/j.knosys.2012.08.011.

[197] R. Saeidi, R. F. Astudillo, and D. Kolossa, "Uncertain LDA: Including Observation Uncertainties in Discriminative Transforms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 7, pp. 1479–1488, 2016, doi: 10.1109/TPAMI.2015.2481420.

[198] A. M. Martinez and A. C. Kak, "PCA versus LDA," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 2, pp. 228–233, 2001, doi: 10.1109/34.908974.

[199] C. Chang and C. Lin, "LIBSVM:Alibraryforsupportvectormachine." 2001.

[200] P. P. Raghu and B. Yegnanarayana, "Supervised texture classification using a probabilistic neural network and constraint satisfaction model," *IEEE Trans. Neural Networks*, vol. 9, no. 3, pp. 516–522, 1998, doi: 10.1109/72.668893.

[201] I. Babaoglu, O. Findik, and E. Ülker, "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.*, vol. 37, no. 4, pp. 3177–3183, 2010, doi: 10.1016/j.eswa.2009.09.064.

[202] M. Kumar, R. B. Pachori, and U. Rajendra Acharya, "An efficient automated technique for CAD diagnosis using flexible analytic wavelet transform and entropy features extracted from HRV signals," *Expert Syst. Appl.*, vol. 63, pp. 165–172, 2016, doi: 10.1016/j.eswa.2016.06.038.

[203] M. Karimi, “Noninvasive detection and classification of coronary artery occlusions using wavelet analysis of heart sounds with neural networks,” *IET*, pp. 117–120, 2006, doi: 10.1049/ic:20050342.

[204] I. Babaoğlu, O. Fındık, and M. Bayrak, “Effects of principle component analysis on assessment of coronary artery diseases using support vector machine,” *Expert Syst. Appl.*, vol. 37, no. 3, pp. 2182–2185, 2010, doi: 10.1016/j.eswa.2009.07.055.

[205] M. G. Poddar, V. Kumar, and Y. P. Sharma, “Automated diagnosis of coronary artery diseased patients by heart rate variability analysis using linear and non-linear methods,” *J. Med. Eng. Technol.*, vol. 39, no. 6, pp. 331–341, 2015, doi: 10.3109/03091902.2015.1063721.

[206] G. G. Yen, “Wavelet packet feature extraction for vibration monitoring,” *IEEE Trans. Ind. Electron.*, vol. 47, no. 3, pp. 650–667, 2000, doi: 10.1109/41.847906.

[207] Y. Gao, Z. Xie, and X. Yu, “A hybrid algorithm for integrated scheduling problem of complex products with tree structure,” *Multimed. Tools Appl.*, vol. 79, no. 43–44, pp. 32285–32304, 2020, doi: 10.1007/s11042-020-09477-2.

[208] B. Hemamalini and V. Nagarajan, “Wavelet transform and pixel strength-based robust watermarking using dragonflyoptimization,” *Multimed. Tools Appl.*, vol. 79, no. 13–14, pp. 8727–8746, 2020, doi: 10.1007/s11042-018-6096-0.

[209] R. N. Khushaba, S. Kodagoda, S. Lal, and G. Dissanayake, “Driver drowsiness classification using fuzzy wavelet-packet-based feature-extraction algorithm,” *IEEE Trans. Biomed. Eng.*, vol. 58, no. 1, pp. 121–131, 2011, doi: 10.1109/TBME.2010.2077291.

[210] J. Tang, S. Alelyani, and H. Liu, “Feature Selection for Classification: A Review,” *Data Classif. Algorithms Appl.*, 2014, doi: 10.1.1.409.5195.

[211] O. Aran and L. Akarun, “A multi-class classification strategy for Fisher scores: Application to signer independent sign language recognition,” *Pattern Recognit.*, vol. 43, no. 5, pp. 1776–1788, 2010, doi: 10.1016/j.patcog.2009.12.002.

[212] G. Baudat and F. Anouar, “Generalized Discriminant Analysis Using a Kernel Approach,” *Neural Comput.*, vol. 12, no. 10, pp. 2385–2404, 2000, doi: 10.1162/089976600300014980.