GENETIC EVOLUTION BASED CARDIAC ARRHYTHMIA CLASSIFIER

PROJECT REPORT

Submitted by

IKRAM.N
Reg. no. 13MAE04

in partial fulfillment for the requirement of award of the degree

of

MASTER OF ENGINEERING

in

APPLIED ELECTRONICS

Department of Electronics and Communication Engineering

KUMARAGURU COLLEGE OF TECHNOLOGY
(An Autonomous Institution affiliated to Anna University, Chennai)
COIMBATORE - 641049

ANNA UNIVERSITY: CHENNAI 600 025

April 2015

BONAFIDE CERTIFICATE

Certified that this project report titled **"GENETIC EVOLUTION BASED CARDIAC ARRHYTHMIA CLASSIFIER"** is the bonafide work of **N.IKRAM [Reg. No. 13MAE04]** who carried out the project under my supervision. Certified further that, to the best of my knowledge the work reported herein does not form part of any other project or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

SIGNATURE	**SIGNATURE**
Ms. S.N.SHIVAPPRIYA	**Dr. RAJESWARI MARIAPPAN**
PROJECT SUPERVISOR	**HEAD OF THE DEPARTMENT**
Department of ECE	Department of ECE
Kumaraguru College of Technology	Kumaraguru College of Technology
Coimbatore-641 049	Coimbatore-641 049

The Candidate with **Register No. 13MAE04** was examined by us in the project viva –voice examination held on...........................

INTERNAL EXAMINER **EXTERNAL EXAMINER**

ACKNOWLEDGEMENT

First I would like to express my praise and gratitude to the Lord, who has showered his grace and blessing enabling me to complete this project in an excellent manner. He has made all things in beautiful in his time.

I express my sincere thanks to our beloved Joint Correspondent, **Shri. Shankar Vanavarayar** for his kind support and for providing necessary facilities to carry out the project work.

I would like to express my sincere thanks to our beloved Principal **Dr.R.S.Kumar M.E., Ph.D.,** who encouraged us with his valuable thoughts.

I would like to express my sincere thanks and deep sense of gratitude to our HOD, **Dr. Rajeswari Mariappan M.E., Ph.D.,** for her valuable suggestions and encouragement which paved way for the successful completion of the project.

I am greatly privileged to express my deep sense of gratitude to the Project Coordinator **Ms.S.Sasikala M.Tech**, Associate Professor, for her continuous support throughout the course.

I am indeed grateful to my project guide, **Ms.Shivappriya M.E, (PhD),** Assistant Professor-II, Department of Electronics and Communication Engineering, Kumaraguru College of Technology, Coimbatore, for her immense contribution, guidance, support and constructive criticism not only during this project but also during this two years of my master program.

Finally I thank my parents and my family members for giving me the moral support and abundant blessings in all of my activities and my dear friends who helped me to endure my difficult times with their unfailing support and warm wishes.

. ABSTRACT

Some of the major diseases that have a high impact on the society are Cardio Vascular Diseases (CVDs). An important category of CVDs are the Cardiac Arrhythmias. In this project the performance of ECG signal can be analyzed using Genetic evolutions (Genetic Algorithm and Genetic Programming) and Artificial Neural Network (ANN). The ECG signal must be clearly represented and filtered to remove all noise and artifacts from the signal. DWT is a powerful tool for ECG signal enhancement, different wavelet transforms used to enhance the ECG signal and pick out the best wavelet from analyzing performance measures. This enhanced signal helpful for classification task. Various parameters are extracted from denoised signal, one cardiac cycle in an ECG signal consists of the P-QRS-T waves. Irrelevant or noisy features unnecessarily increase the complexity of the problem and can degrade classification performance. GP and GA is used only for best feature selection from a large features data set. The algorithm is used to effectively select a smaller subset of features that together form a genetically fit family for classification tasks using ANN. The databases are extracted from MIT-BIH Database (Massachusetts Institute of technology/ Beth Israel Hospital).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACRONYMS	ABBREVIATION
ANN	Artificial Neural Network
BP	Back Propagation
BPN	Back Propagation Network
DWT	Discrete Wavelet Transform
ECG	Electro Cardio Gram
GA	Genetic Algorithm
GP	Genetic Programming
MLP	Multi Layered Perceptron
MSE	Mean Squared Error
PRS	Pattern Recognition System

CHAPTER 1

INTRODUCTION

1.1 ELECTROCARDIOGRAM

The electrocardiogram (ECG) indicates the electrical activity of the human heart. It offers cardiologists with helpful information regarding the rhythm and functioning of the heart. Medical attention is required for patients with abnormal morphology and heart rate in ECG signals since these abnormal cardiac rhythms may lead to life threatening situations. ECG signals are acquired by placing electrodes across the thorax or chest of human body for a limited time and the electrical recordings are visualized using an external device. The electrodes pick the electrical signals generated by the polarization and depolarization of heart muscle. The flat line (isoelectric line) indicates the absence of electrical activity while a deviation from the isoelectric line indicates the electrical action of heart muscles.

1.2 THE HUMAN HEART

The heart cells have small difference in the concentration of ions across the cell membrane at resting stage. The positive ion concentration is high in the outside membrane when compared to the inner membrane of heart cells resulting in resting potential of 90mV. The activation of heart cells causes high permeability of Na+ ion, resulting in a change of polarity across the cell membrane (depolarization). The change in cell potential from negative to positive and back causes a voltage pulse called action potential. The action potential leads to contraction of heart muscles. The depolarization and repolarisation activity of heart cells are measured using ECG. Sino atrial (SA) node is considered as the pacemaker of the human heart because the bunch of cells around SA node causes fast depolarization of heart cells. The human heart contains four chambers as shown in Fig. 1.1.

The top left and right chambers corresponds to left and right atria whereas the bottom left chamber corresponds to left ventricle and bottom right chambers corresponds to right ventricles. The SA node is located inside the right atrium and the depolarization of heart cells in the SA node causes the left and right atria to contract almost simultaneously.

The atria and ventricles are interconnected with each other by atria ventricular valve. A group of cells in the right atrium called Atria Ventricular (AV) node conducts the depolarization of atria by bundle of conducting fibers (Bundle of His) to the ventricles. All part of ventricles is depolarized by purkinje fibers which are located in the muscle walls of ventricles. The polarization and depolarization activity of heart muscles causes the electrical current that moves across the body. The change in the electric current is maximum when one portion of heart is completely polarized while the other is completely depolarized. The summation of action potential from the heart is represented by the electrocardiogram signal. The contraction of two atria forces the blood to flow into ventricles. The two ventricles start contracting once the signal is conducted from the atria and the blood is pumped from ventricles through pulmonary and aortic arteries.

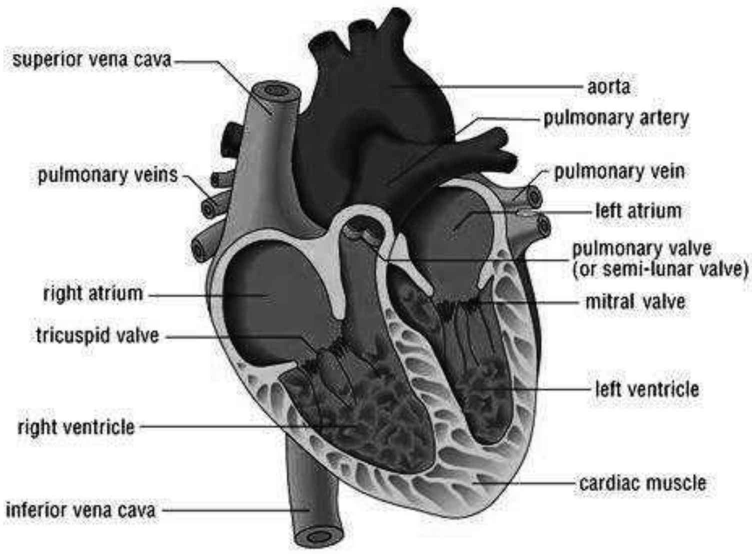

Figure 1.1 Anatomy of Human Heart

1.3 ECG RECORDING

The recording of ECG signals is performed by placing electrodes across the chest of the human body. The term lead refers to the electrical voltage difference between two electrodes. The most commonly used lead system for recording the ECG signals is 12-lead ECG system which includes three classes of leads.

Bipolar leads – It refers to the lead I, II and III. The electrodes are connected to the right arm (RA), left arm (LA) and left leg (LL).

(i) Lead I correspond to the voltage difference between the LA and RA.

(ii) Lead II corresponds to the voltage difference between the LL and RA.

(iii) Lead III corresponds to the voltage difference between the LL and LA.

Augmented unipolar leads – It refers to the lead IV, V and VI as shown in Fig. 1.2.

The positive electrode is placed at right arm for lead IV, left arm for lead V and left leg for lead VI while the negative electrode is placed at the centre of heart electric field which acts as the reference point. The leads under this category are:

(i) Augmented Vector Right (AVR) is the difference between the electrical potential of right arm and centre of heart electric field.

(ii) Augmented Vector Left (AVL) is the difference between the electrical potential of left arm and centre of heart electric field.

(iii)Augmented Vector Foot (AVF) is the difference between the electrical potential of left foot and centre of heart electric field.

Unipolar precordinal leads–It refers to the lead V1-V6. The leads are obtained by taking the difference between potential of electrodes placed on chest and centre of the heart electric field.

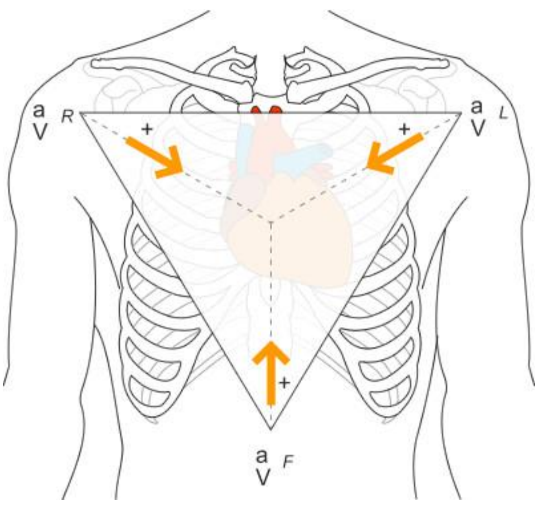

Figure 1.2 Augmented Unipolar Leads

1.3 ECG WAVEFORMS

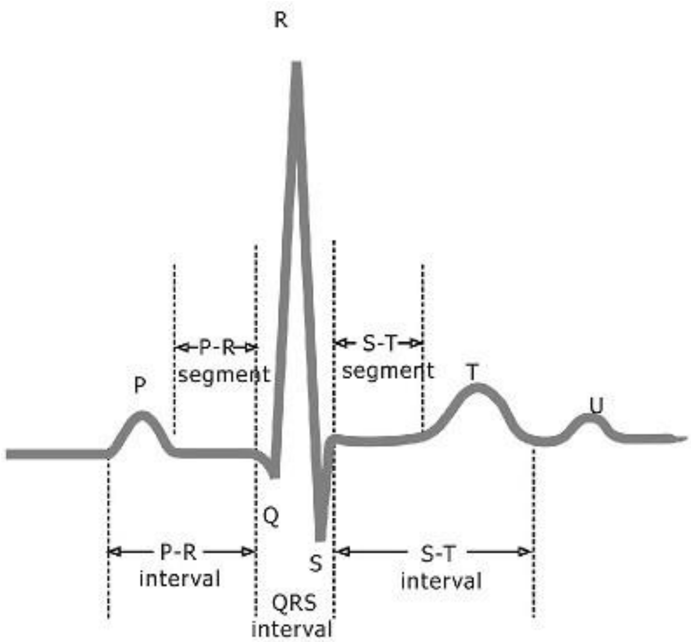

Figure 1.3 ECG signal representing one cardiac cycle

Each cardiac cycle of ECG signal consist of a P, QRS, T and U wave as shown in Fig. 1.3.

P wave: The first upward pulse from the isoelectric line followed by return to the isoelectric line corresponds to the P wave and it exists for duration of 0.04 sec. The depolarization activity of atria's causes the contraction of left and right atria resulting in P wave.

QRS complex wave: This signal results due to the depolarization activity of ventricles. The Q wave is a downward pulse which is followed by R wave with sharp positive peak. The R wave is followed by the S wave with a negative swing. The Q, R and S wave together constitute the QRS complex signal which indicates the time for the contraction of ventricles and it exist for duration of 80-120 msec.

T wave: The QRS complex signal is followed by the T wave. It is an upward pulse which indicates the repolarisation of ventricles.

U wave: The repolarisation activity of purkinji fibers is indicated by a small deflection following T wave called the U wave.

PR segment: The P and QRS waves are connected by PR segment. In this duration of time the electrical impulse from AV node travels from atria to the ventricle. The PR segment represents a flat signal since there is no contraction of heart muscles occurring during that interval.

ST segment: This segment interconnects the QRS and T wave. The ST segment represents the depolarization activity of ventricles. The duration of ST segment is approximately 80-120 msec.

1.5 NOISES EMBEDDED IN ECG SIGNALS

It is difficult for medical experts to analyse ECG signals embedded with noise. The noise removal technique is a challenging task due to spectral overlap between ECG signal and noise signal. The different types of noises that affect the ECG signals are:

1.5.1 Baseline wander

Baseline wander results in the movement of isometric line in upward and downward direction. It is caused due to the movement of electrodes connected across the chest during breathing or due to the movement of arm or leg. The variation in temperature and bias of the instrumentation amplifier circuit can also be a cause of baseline drift. It is a low frequency noise with a frequency range of 0 - 0.5Hz.

1.5.2 Power line interference

It occurs due to the poor grounding of ECG machines connected to the power supply. The ECG machine picks up the ac signal of 50/60 Hz frequency and displays a thick looking ECG signal.

1.5.3 Electromyogram (EMG) interference

The electrical activity of muscles causes the contraction of muscles. The resulting signals are band limited Gaussian noise with a zero mean distribution. The EMG interference causes fast fluctuations which are faster than the ECG signals. Its frequency range is 0-10 KHz and occurs for duration of 50msec.

1.6 ECG ARRHYTHMIAS

A Normal Sinus Rhythm (NSR) represents an ECG signal with no cardiac disorder and has a heart rate of 60-100 Beats Per Minute (BPM). A heart rate beyond 100 BPM indicates sinus tachycardia while a heart rate below 60 BPM indicates sinus bradycardia which affects the vital organs.

1.6.1 Sinus node arrhythmias

The SA node of the heart is responsible for this type of arrhythmias. The distinguishing feature of this type of arrhythmias is that the morphology of P wave of ECG signals remains normal. Sinus arrhythmia, Sinus bradycardia, and Sinus arrest are different type of arrhythmias which comes under the category of sinus node arrhythmias.

1.6.2 Atrial arrhythmias

These types of arrhythmias originate inside the atria but outside the SA node. The different types of atrial arrhythmias are:

Premature Atrial Contractions (PAC)

The P wave has abnormal morphology while the QRS and T wave have normal morphology. This problem arises due to the early firing of ectopic pacemaker before the SA node. Atrial tachycardia is characterized by the occurrence of PAC as triplet or more.

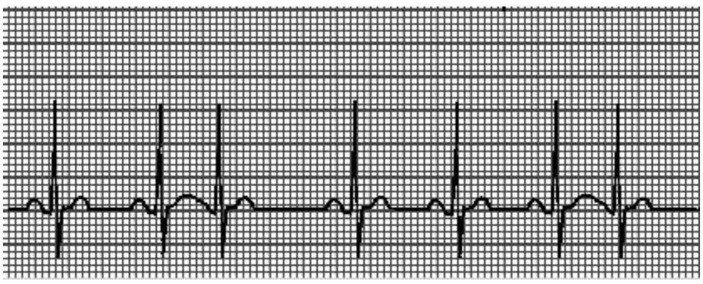

Figure 1.4 Premature atrial contraction

Atrial Tachycardia

This arrhythmia shows a heart rate of 160 to 240 BPM. The symptoms of atrial tachycardia are feeling of palpitations, nervousness and anxiety.

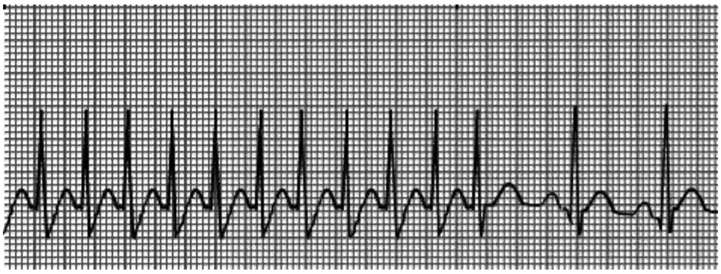

Figure 1.5 Atrial Tachycardia

1.6.3 Ventricular Arrhythmia

This type of arrhythmias is characterized by wide QRS complex signal with bizarre shape. The impulse signal originates on ventricles and move to the remaining portion of heart.

Premature Ventricular contraction

The heart beat is triggered by the purkinje fibers in the ventricles resulting in the contraction of ventricles before the complete depolarization of atria. Lack of oxygen in cardiac muscles may cause PVC beats and they can occur anywhere in the ECG beat cycle.

1.6.4 Bundle Branch Blocks

The bundle branch blocks results in myocardial infarction since there is a restriction in the movement of impulse signal from AV node to the entire conduction system. There are two types of Bundle Branch Blocks (BBB). The right BBB prevents the impulse signal from AV node to depolarize the right ventricle while the left BBB prevents the depolarization of left ventricular muscles. These blocks results in myocardial infarction.

1.7 ECG DATABASE

All the experiments in this work are carried out by considering three different public databases namely MIT-BIH Arrhythmia, Noise Stress Data Base (NSTDB) and INCARTDB change which is available freely on Physionet..For all databases the AAMI recommendations for class-labelling are adopted..The database includes different types of ECG recordings. In this work, data with a broad range of normal and pathological ECG recordings are considered to evaluate the algorithm performance**.**

1.7.1 MIT-BIH Arrhythmia Database (MITBIHAR)

The MIT institute in collaboration with Boston's Beth Israel Hospital (BIH) developed a standard database in 1980 for arrhythmia detection. This is the first standard database used by researchers to validate the performance of their proposed technique and compare their result with pre-existing algorithms. A total of 25 male patients with an age of 32 to 89 years and 22 female patients with an age of 23 to 89 years were chosen in order to include normal beat and common type of life threatening arrhythmias. Each ECG record consists of two channel recordings. The first channel recording uses modified lead limb II (MLII) while the second channel recording commonly uses lead V1(V2, V4 or V5 for some patients). The database comprises of 48 recordings each containing 30 minutes segment of ECG selected from 24 hour recording of 47 different patients (200 and 201 ECG records are acquired from same patient). The first 23 (100-124) recordings correspond to the routine clinical recordings while the remaining recordings (200 – 234) contain the complex arrhythmias. The analog signals were sampled at a frequency of 360 Hz in order to use a notch filter with a notch frequency of 60 Hz for eliminating power line interference and band pass filtered at 0.1-100 Hz in order to avoid anti-aliasing and saturation in analog to digital conversion.

1.7.2 Noise Stress Data Base (NSTDB)

This database includes 12 half-hour ECG recordings and 3 half-hour recordings of noise typical in ambulatory ECG recordings. The noise recordings were made using physically active volunteers and standard ECG recorders, leads, and electrodes; the electrodes were placed on the limbs in positions in which the subjects ECGs were not visible. The three noise records were assembled from the recordings by selecting intervals that contained predominantly baseline wander, muscle and electrode motion artifact. Electrode motion artifact is generally considered the most troublesome, since it can mimic the appearance of ectopic beats and cannot be removed easily by simple filters

1.7.3 1nstitute of Cardiological Technics Data Base (INCARTDB)

This database consists of 75 annotated recordings extracted from 32 holter records. Each record is 30 minutes long and contains 12 standard leads, each sampled at 257 Hz, with gains varying from 250 to 1100 analog-to-digital converter units per mill volt. Gains for each record are specified in its .hea file. The reference annotation files contain over 175,000 beat annotations in all.

The original records were collected from patients undergoing tests for coronary artery disease. None of the patients had pacemakers; most had ventricular ectopic beats. In selecting records to be included in the database, preference was given to subjects with ECGs consistent with ischemia, coronary artery disease, conduction abnormalities, and arrhythmias.

1.8 LITERATURE SURVEY

1. Feature Selection by Genetic Programming, Artificial Neural Network-based Machine Condition Monitoring (2012) – In this paper the performance of bearing fault diagnosis using Genetic Programming and Artificial Neural Networks (ANNs). The experimental data is collected for four bearings conditions namely: Healthy, defective Outer race, defective Inner race and defective ball fault condition. Artificial neural network have been widely used for health diagnosis of rotating machinery using features extracted from vibration emission signals. One of the most important considerations in applying neural networks to condition monitoring of electrical machine is the proper selection of training features. Irrelevant or noisy features unnecessarily increase the complexity of the problem and can degrade modelling performance. A Genetic programming for feature selection is developed, based on the concept of dominance. GP is used for two purpose Feature extractor and feature selector but in this work GP is used only for best feature selection from a large features data set. The algorithm is used to effectively select a smaller subset of features that together form a genetically fit family for fault identification and classification tasks.

2. Classification of cardiac arrhythmia with respect to ECG and HRV signal by genetic Programming (2012) - Consistent or periodical heart rhythm disorders may result cardiac arrhythmias. In this paper, Heart Rate Variability (HRV) signals are analyzed and various features including time domain, frequency domain and nonlinear parameters are extracted. Moreover, additional nonlinear features are extracted from electrocardiogram signals. These features are helpful in classifying cardiac arrhythmias. It describes that, genetic programming is applied to classify heart arrhythmias using both HRV and ECG features. Genetic programming selects effective features, and then finds the most suitable trees to distinguish between different types of arrhythmia. By considering the variety of extracted parameters from ECG and HRV signals, genetic programming can precisely differentiate various arrhythmias. The performance of this algorithm is evaluated on MIT–BIH Database. The results show that seven different types of arrhythmia classes including normal beat, left bundle branch block beat, right bundle branch beat, premature ventricular contraction, fusion of ventricular and normal beat, atrial premature contraction and paced beat are classified.

3. Genetic Programming for Object Detection: A Two-Phase Approach with an Improved Fitness Function (2008) - In this paper describes two innovations that improve the efficiency and effectiveness of a genetic programming approach to object detection problems. The approach uses genetic programming to construct object detection programs that are applied, in a moving window fashion, to the large images to locate the objects of interest. The first innovation is to break the GP search into two phases with the first phase applied to a selected subset of the training data, and a simplified fitness function. The second phase is initialised with 1the programs from the first phase, and uses the full set of training data with a complete fitness function to construct the final detection programs. The second innovation is to add a program size component to the fitness function. This approach is examined and compared with a neural network approach on three object detection problems of increasing difficulty. The results suggest that the innovations increase both the effectiveness and the efficiency of the genetic programming search, and also that the genetic programming approach outperforms a neural network approach for the most difficult data set in terms of the object detection accuracy.

4. ECG Feature Extraction Techniques - A Survey approach (2010) - ECG Feature Extraction plays a significant role in diagnosing most of the cardiac diseases. One cardiac cycle in an ECG signal consists of the P-QRS-T waves. This feature extraction scheme determines the amplitudes and intervals in the ECG signal for subsequent analysis. The amplitudes and intervals value of P-QRS-T segment determines the functioning of heart of every human. Recently, numerous research and techniques have been developed for analyzing the ECG signal. In this schemes were mostly based on Fuzzy Logic Methods, Artificial Neural Networks (ANN), Genetic Algorithm (GA), Support Vector Machines (SVM), and other Signal Analysis techniques. All these techniques and algorithms have their advantages and limitations. This work discusses various techniques and transformations proposed earlier in literature for extracting feature from an ECG signal.

5. Cardiac Arrhythmia Classification using Cartesian Genetic Programming Evolved Artificial Neural Network (2013) - Some of the major diseases that have a high impact on the society are Cardiovascular Diseases (CVDs). An important category of CVDs are the Cardiac Arrhythmias. Conventional methods of diagnosis for the disease are prone to errors and need experience on part of the diagnosing physician. For automatic detection of Cardiac Arrhythmia, author developed an algorithm that first applies digital signal processing and logical operations to the time domain ECG signal and hence detects the fiducial points of an ECG complex. From the fiducial points, the lengths and slopes of a number of segments and amplitudes of peaks are determined.These parameter values are applied to the fast learning evolutionary algorithm of Cartesian Genetic Programming Evolved Artificial Neural Network (CGPANN) to classify the beats.A number of setups were experimented with.In these experiments, the CGPANN was first trained with known data from the popular MIT-BIH database and then tested with another part of the known data and the results found highly satisfactory.

6. QRS Complex Detection Using Optimal Discrete Wavelet transform (2009) – Wavelet transform has emerged as a powerful tool for time frequency analysis of complex non stationary signals such as the electrocardiogram signal. In this paper the design of good wavelets for cardiac signal is discussed from the perspective of orthogonal filter banks. Optimum wavelet for ECG signal is designed and evaluated based on perfect reconstruction conditions and QRS complex detection. The performance is evaluated by using the ECG records from the MIT-BIH arrhythmia database. In the first step, the filter coefficients (optimum wavelet) are designed by re parameterization of filter coefficients. In the second step, ECG signal is decomposed to three levels using the optimum wavelet and reconstructed. From the reconstructed signal, the range of error signal is calculated and it is compared with the performance of other suitable wavelets already available in the literature. The optimum wavelet gives the maximum error range as 10-14–10-11 which is better than that of other wavelets existing in the literature. In the third step, the baseline wandering is removed from the ECG signal for better detection of QRS complex. The optimum wavelet detects all R peaks of all records. That is using optimum wavelet 100% sensitivity and positive predictions are achieved. Based on the performance, it is confirmed that optimum wavelet is more suitable for ECG signal.

CHAPTER 2

METHODOLOGY

2.1 INTRODUCTION

In literature, numerous feature extraction techniques are applied to analyze and classify ECG beats such as the particle swarm optimization technique, principle component analysis, ECG morphological features in conjunction with timing information. Though these techniques have provided good classification results but the wavelet based techniques have outperformed when compared to the aforementioned techniques. The Fourier transform technique fails in analyzing non stationary signals due to its poor time frequency localization of the signal. The Short Time Fourier Transform (STFT) analyses every spectral component with fixed window thereby providing fixed time and frequency resolution. The uncertainty principle explains the inability to achieve high frequency and time resolution at same time instant. The Wavelet Transform (WT) technique helps in analyzing high frequency signal with high time resolution and low frequency signal with high frequency resolution by decomposing the input signal over scaled and shifted versions of a prototype wavelet. The use of varying window size in wavelet transform helps in analyzing non stationary signal at multiple 9 resolutions. In practical situations, the signals with high frequency component exist for short duration and low frequency component exist for long duration. The WT provides varying time-frequency window thereby analyzing high frequency component with narrow window and low frequency component with wide window. The wavelet transform can be represented in continuous and discrete form.

2.2 WAVELET TRANSFORM

Although the discretized continuous wavelet transform enables the computation of the continuous wavelet transform by computers, it is not a true discrete transform. As a matter of fact, the wavelet series is simply a sampled version of the CWT, and the information it provides is highly redundant as far as the reconstruction of the signal is concerned. This redundancy, on the other hand, requires a significant amount of computation time and resources. The Discrete Wavelet Transform (DWT), on the other hand, provides sufficient information both for analysis and synthesis of the original signal, with a significant reduction in the computation time.The DWT is considerably easier to implement when compared to the

CWT. The basic concepts of the DWT with its properties and the algorithms used to compute it.

Wavelet transform is an efficient tool for analyzing non stationary ECG signals. The wavelet transform is used to extract morphological information from the ECG data. Wavelet transformation maps the ECG signal into a time-scale plane. The fourth scale has the good projection between the frequency ranges 12.25 to 22.5Hz.It is worth to use the WT which can be efficiently implemented as a filter bank. The WT function of the signal is given as

$$\int_{-\alpha}^{\alpha} f(x)\varphi *_{(a,b)(x)dx} \tag{2.1}$$

*-denotes complex conjugate

2.2.1 Wavelet types

Symlet, Haar, Biorthogonal wavelets are used for denoising the signal.Among these wavelets pick out the best wavelet for pre processing and feature extraction it depends on the performance measures.

Biorthogonal Wavelet

A biorthogonal wavelet is a wavelet where the associated wavelet transform is invertible but not necessarily orthogonal. Designing biorthogonal wavelets allows more degrees of freedom than orthogonal wavelets.One additional degree of freedom is the possibility to construct symmetric wavelet functions.

In the biorthogonal case, there are two scaling functions $\emptyset, \theta$ which may generate different multiresolution analyses, and accordingly two different wavelet functions .So the numbers *M* and *N* of coefficients in the scaling sequences a, a* may differ.The scaling sequences must satisfy the following biorthogonality condition

$$\sum_{n\in z}^{\mathrm{n}} \mathrm{a_n}\, \mathrm{\breve{a}_n} + 2\mathrm{m} = 2\, .\, \delta_{\mathrm{m},0} \tag{2.2}$$

Scaling function of Biorthogonal wavelet can be defined as

$$b_n = (-1)^n{}_{a*_{M-1-n}}(n = 0, \ldots \ldots, N-1) \tag{2.3}$$

$$b *_n = (-1)^n{}_{a_{M-1-n}}(n = 0, \ldots \ldots, N-1) \tag{2.4}$$

Haar Wavelet

Haar wavelet is also the simplest possible wavelet. The technical disadvantage of the haar wavelet is that it is not continuous, and therefore not differentiable.

The Haar mother wavelet function φ(t) can be described as

$$\varphi(t) = \begin{cases} 1 & 0 \le t < 1/2 \\ -1, & 1/2t \le t < 1 \end{cases} \tag{2.5}$$

The scaling function of Haar can be described as

$$\emptyset(t) = \begin{cases} 1, & 0 \le t < 1, \\ 0, & otherwise \end{cases} \tag{2.6}$$

Daubechies Wavelet

The Daubechies wavelets, based on the work of Ingrid Daubechies, are a family of orthogonal wavelets defining a discrete wavelet transform and characterized by a maximal number of vanishing moments for some given support. With each wavelet type of this class, there is a scaling function (called the *father wavelet*) which generates an orthogonal multiresolution analysis.

The scaling function of Daubechies wavelet can be defined as

$$b_k = -1^k{}_{a_{N-1-k}} \tag{2.7}$$

Compare the different wavelet transform in terms of performance for the three database, pick out the best wavelet for pre processing and feature extraction to detect R-peaks, RR interval, QRS points and ST points in the ECG waveforms since those are going to help the classification.

2.3 PRE-PROCESSING

Considerable attention has been paid to the design of filters for the purpose of removing baseline wander and power-line interference; both types of disturbance imply the design of a narrowband filter. Removal of noise because of muscle activity represents another important filtering problem being much more difficult to handle because of the substantial spectral overlap between the ECG and muscle noise. Muscle noise present in the ECG can, however, be reduced whenever it is appropriate to employ techniques that benefit from the fact that the ECG is a recurrent signal. For example, ensemble averaging techniques can be successfully applied to time-aligned heartbeats for reduction of muscle noise. The filtering techniques are primarily used for pre-processing of the signal and have as such been implemented in a wide variety of systems for ECG analysis. It should be remembered that filtering of the ECG is contextual and should be performed only when the desired information remains undistorted. This important insight may be ex-emplified by filtering for the removal of power-line interference.

Removal of baseline wander is required in order to minimize changes in beat morphology that do not have cardiac origin, which is especially important when subtle changes in the‘‘low-frequency’’ ST segment are analyzed for the diagnosis of ischemia, which may be observed, for example, during the course of a stress test. The frequency content of baseline wander is usually in the range below 0.5Hz; however, increased movement of the body during the latter stages of a stress test further increases the frequency content of baseline wander Patients unable to perform a traditional treadmill or ergometer stress test may still be able to perform a stress test by sitting, running an ergometer by hand, or using a special rowing device. In such cases, baseline wander related to motion of the arms severely distorts the ECG signal. The design of a linear, time-invariant, high pass filter for removal of baseline wander involves several considerations, of which the most crucial are the choice offilter cut- off frequency and phase response characteristic. The cut- off frequency should obviously be chosen so that the clinical information in the ECG signal remains undistorted while as much as possible of the baseline wander is removed. Hence, it is essential to find the lowest frequency component of the ECG spectrum. In general, the slowest heart rate is considered to define this particular frequency component; the PQRST waveform is attributed to higher frequencies. During bradycardia, the heart rate may drop to approximately 40beats/minute, implying that the lowest frequency contained in the ECG is approximately 0.67Hz . As the heart rate is not perfectly regular but always fluctuates from one beat to the next, it is necessary to choose a slightly lower cut-off frequency such as 0.5Hz. If too high a cut-off frequency is employed, the output of the highpass filter contains an unwanted, oscillatory component that is strongly correlated to the heart rate. In certain situations, baseline wander becomes particularly pronounced at higher heart rates such as during the latter stages of a stress test when the workload increases. Then, it may be advantageous to couple the cut-off frequency to the prevailing heart rate, rather than to the lowest possible heart rate, to further improve base- line removal. Linear filtering with time-variable cut-off frequency was initially suggested for offline processing of ECG signals and later extended for online use (3,4). The other crucial design consideration is related to the properties of the phase response and, consequently, the choice of filter structure. Linear phase filtering is highly desirable in order to prevent phase distortion from altering various wave properties of the cardiac cycle such as the duration of the QRS complex, the ST–T segment level, or the endpoint of the T wave. It is well-known that FIR filters can have an exact linear phase response, provided that the impulse response is either symmetric or anti symmetric; however, FIR designs result in high filter orders.

ECG signal is applied as input to the pre processing stage. These signals consist of baseline wander, power line interference, and high frequency noises. The pre-processing stage utilizes a filtering unit to remove artifact signals from the ECG signals. The following steps explain the function of pre-processing stage

Step 1: First read the annotation file, header file and Data file

Step 2: Apply smoothening function to that signal to remove the noise interference.

Step 3: To remove the base line interference, the signal obtained from the step2 is subtracted from original signal.

Step 4: Determine Wavelet coefficients at scale 5 and using that denoise the output of step 3 signals.

Step 5: Finally a Clean ECG signal is obtained and it is considered as input to the next stage.

2.4 ECG SEGMENTATION

Electrocardiogram (ECG) introduced into clinical practice more than 100 years ago constitutes a graphical recording of the heart’s electrical activity that occurs successively over time. The ECG results determine whether the heart is performing normally or suffering from abnormalities. The recorded ECG is the representation of the depolarization and re-polarization of the heart and can diagnose a patient by looking at the characteristics of the traced ECG readings. There are 3 main deflections in an ECG: the P-wave, the QRS complex, and the T-wave. The first upright wave is called the P wave and is normally round in shape and its duration is usually not more than 0.1 second. The QRS complex corresponds to ventricular depolarization. It is normally 0.04 - 0.12 second in duration. Between the P wave and the QRS complex is the PR interval. It represents the time taken by the SA node electrical impulse to travel from its exit out of the SA node to the beginning of ventricular excitation. It is normally 0.1 - 0.2 second in duration. The T wave is another rounded upright wave corresponding to repolarization of the ventricles. Finally, the ST segment is the electric line between the end of the QRS complex and the beginning of the T wave. Additionally, it is useful for epidemiologic studies and screening.The ECG waveforms may differ for the same patient to such extend that they are unlike to each other and at the same time alike for different types of beats. The Wavelets are a powerful tool for the representation and analysis of such physiologic waveforms because a wavelet has finite duration as contrast to Fourier methods based on sinusoids of infinite duration. The wavelet transform or wavelet analysis is probably the most recent solution to overcome the shortcomings of the fourier transform. In wavelet analysis the use of a fully scalable modulated window solves the signal-cutting problem. The window is shifted along the signal and for every position the spectrum is calculated. Then this process is repeated many times with a slightly shorter (or longer) window for every new cycle. In the end the result will be a collection of time-frequency representations of the signal, all with different resolutions.

2.4.1 Pan Tompkins's Algorithm

ECG is essentially responsible for patient monitoring and diagnosis.The extracted feature from the ECG signal plays a vital in diagnosing the cardiac disease.The development of accurate and quick methods for automatic ECG feature extraction is of major importance. Therefore it is necessary that the feature extraction system performs accurately.The purpose of feature extraction is to find as few properties as possible within ECG signal that would allow successful abnormality detection and efficient prognosis.

In every cardiac cycle of heart, R peak occurred due to the contraction of the ventricles.The main reason for considering the R peak for this work is that it has higher amplitude value when compared to other points in the ECG signal. Therefore it can be easily identifiable. The normal duration of between two R peaks (i.e) RR interval varies from 0.6 to 1.2 seconds. Abnormal beats shows variation in their range.That abnormality is due to delay in the contraction of ventricles which leads to a possibility of getting more than one R peak within one cycle.

To find QRS points Pan-Tompkins's algorithm is used.The following steps explain how to obtain the QRS points in beat.

Step 1: Apply WT to clean ECG signal which is the output of pre processing stage.

Step 2: Reconstruct the signal only using approximate coefficients.

Step 3: As per Pan-Tompkins algorithm apply band pass filter and moving window integrator. The threshold value is obtained by subtracting the mean and maximum value of the previous step output signal.

Step 4: Compare the peaks of the beats which are greater than the threshold are considered as R peaks.

Step 5: For some records there is a need to change the threshold for finding the missing beats.

Step 6: After finding R peak, set the window 50ms before R peak and 150ms after R peak.

Step 7: To the left of R peak, within that window find the minimum point that is considered as Q point.

Step 8: To the Right of R peak, within that window find the minimum point that is considered as S point.

Step 9: Using that points find the QRS duration.

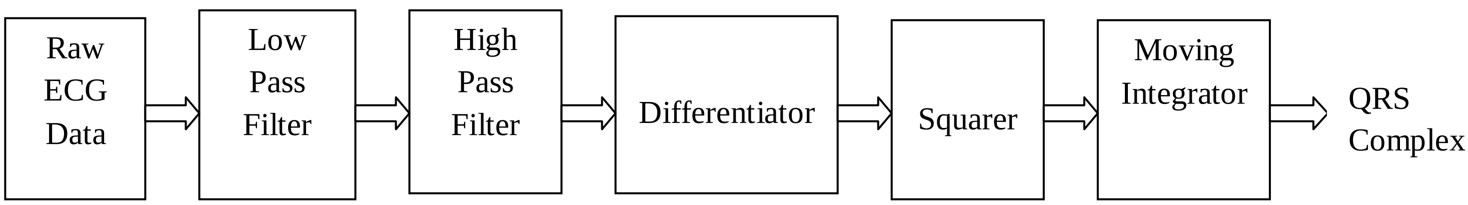

Figure 2.1 Workflow of Pan Tompkins's Algorithm

2.4.2 Features extracted

Table 2.1 Feature Extracted from De noised ECG Signal

S.NO	FEATURES	DESCRIPTION
1	QRS Complex	Due to contraction of ventricles occurs duration between 80-120 ms
2	R-amp	Depolarization activity of atria's
3	S-T segment	Depolarization activity of ventricles occurs between 80-120ms (approx)
4	S amp	Downward deflection of QRS complex occurs after R-wave
5	RR interval	Interval between two consecutive R-amp
6	Q-amp	Depolarization of the interventricular septum

CHAPTER 3

CLASSIFICATION USING ARTIFICIAL NEURAL NETWORK

3.1 INTRODUCTION

In recent years neural network has been used as a tool for classification problems in many fields. The feed forward network architecture such as MLP, RBF network is commonly used in pattern recognition problems. The patterns are given as input to the input layer which propagates in the forward direction from one layer to another layer. The MLP networks are trained by back propagation algorithm which employs supervised learning technique.A well-known error correction algorithm used in MLP network is the Least Mean Square (LMS) algorithm. The MLP network consists of two passes in the network. In the forward pass, the input patterns are applied to the input neurons and the signal propagates in the forward direction resulting in a group of outputs (actual response). In the backward pass, the error correction rule will determine the synaptic weights and the error signal propagates in the direction opposite to the synaptic weight connection. The weights are modified so that the actual output approaches the desired output.

3.1.1 Overview

Neural networks are composed of simple elements operating in parallel. These elements are inspired by biological nervous systems. As in nature, the connections between elements largely determine the network function. You can train a neural network to perform a particular function by adjusting the values of the connections (weights) between elements.

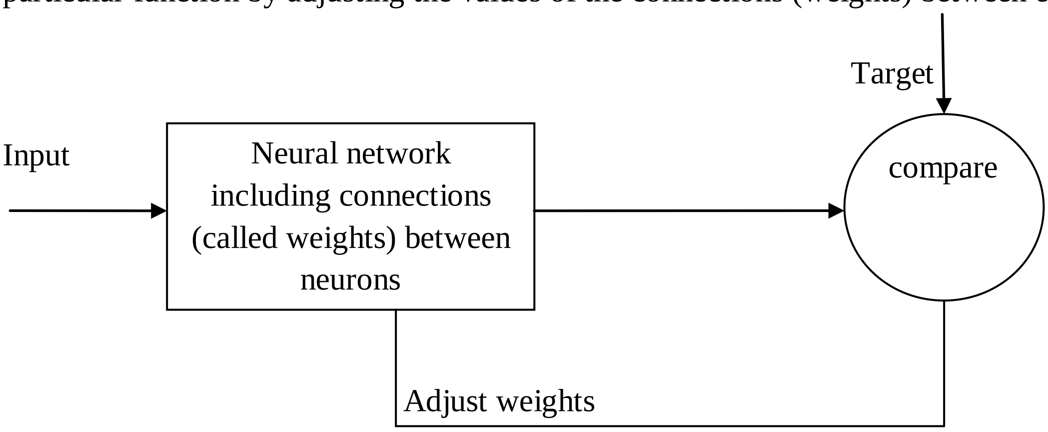

Figure 3.1 Typical neural network flow diagram

Typically, neural networks are adjusted, or trained, so that a particular input leads to a specific target output. The figure 3.1 illustrates such a situation. Here, the network is adjusted, based on a comparison of the output and the target, until the network output matches the target. Typically, many such input/target pairs are needed to train a network.This have been trained to perform complex functions in various fields, including pattern recognition, identification, classification, speech, vision, and control systems.It can also be trained to solve problems that are difficult for conventional computers or human beings

3.2 MLP NETWORK ARCHITECTURE

The three important characteristics of a MLP network are:

1) An MLP network uses a nonlinear smooth activation function (differentiable at any point) which is superior to the hard limiter activation function

2) The network consists of more than one hidden layers. The hidden neurons help to learn difficult problems by extracting significant information from input data.

3) The network has high connectivity.

The Fig. 3.2 represents a 3 layered MLP network. The MLP network contains two types of signals (Functional signal and Error signal).The functional signal is the input signal which propagates in the forward direction and emerges at the output of the network. The output signal is considered as the function of the input signal and corresponding weights. The error signal generated at the output propagates in the backward direction.The hidden or output neuron performs two important tasks. They determine the output of a neuron which is the combination of nonlinear function of input and synaptic weight of that neuron. They also determine the gradient vector required for the synaptic weight updation during backward pass

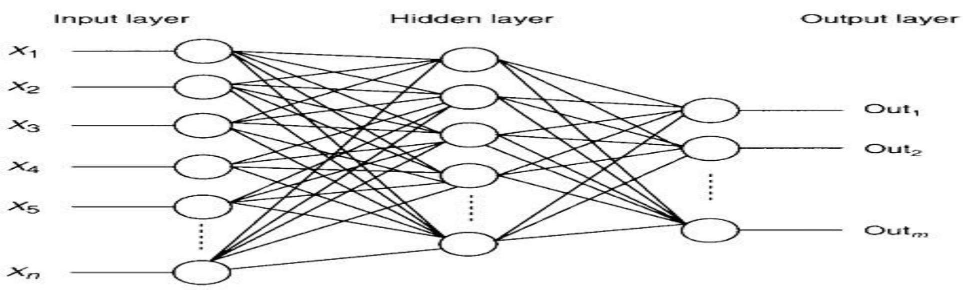

Figure 3.2 Three layered MLP network

If $y_j(n)$ indicates the output response of j^{th} neuron and $d_j(n)$ represents the desired response of j^{th} neuron then the error signal $e_j(n)$ is given by

$$e_j(n)=d_j(n)-y_j(n) \quad (3.1)$$

The weight updating equation is given by

$$w_{ji}(n+1)=w_{ji}(n)+\acute{\eta}\$_j(n)y_i(n) \quad (3.2)$$

Where $\acute{\eta}$ corresponds to learning rate parameter, $\$_j$ refers to local gradient function and $y_i(n)$ refers to input of j^{th} neuron for n^{th} iteration.If j is consider as output neuron then

$$\$_j(n)=e_j(n)\varphi(v_j(n)) \quad (3.3)$$

Where φ is the activation function and

$$v_j(n)= v(\sum_{i=0}^{m} \quad w_{ji\,(n+1)\;y_{i(n)}}n) \quad (3.4)$$

If neuron J is considered as the hidden neuron and neuron k as the output neuron then

$$\delta_j(n) = \varphi(v_j(n)) \sum_k \delta_k(n)\, w_{kj}(n) \quad (3.5)$$

Though back propagation algorithm is considered as an efficient computation technique but it doesn’t ensures optimal solution for all complex problems

3.3 FEED FORWARD MODEL

Generalized feed forward neural networks are a generalization of the Multilayer Perceptron (MLP) such that connections can jump over one or more layers. In theory, a MLP can solve any problem that a generalized feed forward network can solve.In practice however, generalized feed forward networks often solve the problem much more efficiently. A classic example of this is the two spiral problem. Without describing the problem, it suffices to say that a standard MLP requires hundreds of times more training epochs than the generalized feed forward network containing the same number of processing elements.

3.4 BACK PROPAGATION LEARNING ALGORITHM

The back propagation learning algorithm uses the delta-rule. What this does is that it computes the deltas, (local gradients) of each neuron starting from the output neurons and going backwards until it reaches the input layer. To compute the deltas of the output neurons, first step have to get the error of each output neuron. Thus back propagation rule propagates the errors through the network and allows adaptation of the hidden processing elements (PEs). Two important characteristics of the multilayer perceptron are: its nonlinear processing elements (PEs) which have a nonlinearity that must be smooth (the logistic function and the hyperbolic tangent are the most widely used); and their massive interconnectivity (i.e. any element of a given layer feeds all the elements of the next layer). The multilayer perceptron is trained with error correction learning, which means that the desired response for the system must be known. In pattern recognition this is normally the case, since input data is labeled, i.e. it justified that which data belongs to which experiment.

Error correction learning works in the following way: From the system response at PE_i at iteration n, $y_i(n)$ and the desired response $d_i(n)$ for a given input pattern an instantaneous error, $e_i(n)$ is defined by

$$e_i(n) = a_i(n) - y_i(n) \tag{3.6}$$

Using the theory of gradient descent learning, each weight in the network can be adapted by correcting the present value of the weight with a term that is proportional to the present input and error at the weight, i.e.

$$w_{ii}(n + 1) = w_{ii}(n) + \eta\delta_i(n)\, x_j(n) \tag{3.7}$$

The local error $\delta_i(n)$ can be directly computed from $e_i(n)$ at the output PE or can be computed as a weighted sum of errors at the internal PEs. The constant η is called the step size. This procedure is called the back propagation algorithm. Back propagation computes the sensitivity of a cost functional with respect to each weight in the network, and updates each weight proportional to the sensitivity. The beauty of the procedure is that it can be implemented with local information and requires just a few multiplications per weight, which is very efficient. Because this is a gradient descent procedure, it only uses the local information so can be caught in local minima. Moreover, the procedure is inherently noisy since we are using a poor estimate of the gradient, causing slow convergence.

Momentum learning is an improvement to the straight gradient descent in the sense that a memory term (the past increment to the weight) is used to speed up and stabilize convergence. In momentum learning the equation to update the weights becomes: Where α, is the momentum and normally it should be set between 0.1 and 0.9.

$$w_{ij}(n+1) = w_{ij}(n) + \eta\delta_i(n)\, x_j(n) \qquad (3.8)$$

$$+ \alpha\, w_{ij}(n) - w_{ij}(n-1)$$

3.5 LEVENBERG-MARQUARDT (LM) ALGORITHM

The Levenberg-Marquardt (LM) algorithm is basically an iterative method that locates the minimum of a multivariate function that is expressed as the sum of squares of non-linear real-valued functions. LM can be thought of as a combination of steepest descent and the Gauss-Newton (GN) method. LM algorithm is more robust than GN algorithm which essentially means that it finds a solution even if it starts far off the final minimum. During the iterations, the new configuration of weights in step *k+1* is calculated as follows

$$w(k+1) = w(k) - (J^t J + \alpha I)^{-1} J^t \dot{\varepsilon}(k) \qquad (3.9)$$

where *J* - Jacobian matrix, α - adjustable parameter, $\dot{\varepsilon}$ - error vector. The parameter is modified based on the development of error function E. If the step causes a reduction of *E*, we accept it. Otherwise, is changed; reset the original value and recalculate *w (k+1)*.

3.6 NEURAL NETWORK DESIGN STEPS

Standard steps for designing neural networks to solve problems in four application areas: function fitting, pattern recognition, clustering, and time series analysis. The work flow for any of these problems has seven primary steps. (Data collection in step 1, while important, generally occurs outside the MATLAB environment.)

1. Collect data

2. Create the network

3. Configure the network

4. Initialize the weights and biases

5. Train the network

6. Validate the network

7. Use the network

3.7 PERFORMANCE MEASURES

3.7.1 Mean Squared Error (MSE)

The mean squared error is simply two times the average cost. The formula for the mean squared error is

$$MSE = \sum_{j=0}^{p} \sum_{i=0}^{N} \left(d_{ij-} y_{ij}\right)^{2} / NP \qquad (3.10)$$

P = number of output processing elements (PEs)
N = number of exemplars (instances) in the data set
y_{ij} = Network output for exemplar i at processing element *j*
d_{ij} = desired output for exemplar i at processing element *j*

3.7.2 Accuracy, Sensitivity, Specificity, Positive predictivity

The performances of the classification are expressed in terms of Sensitivity (Sen), Specificity (Spe), and Positive Predictivity (Ppr), Accuracy (Acc). Respective definitions using True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN), all of which can be obtained from the classification results, are as follows:

- True positive (TP) are beats which have been correctly assigned to a certain class
- True negative(TN) are beats which have been wrongly assigned to a certain class
- False positive (FP) are beats which have been incorrectly assigned to that same class.

- False negative (FN) occurs when a beat should have been assigned to that class but missed and assigned to another class.

- **Sensitivity** is the rate of correctly classified events among all events, *Sen = TP / (TP + FN)*
- **Specificity** is the rate of correctly classified non events among all non events, *Spe = TN / (TN + FP)*

- **Positive Predictivity** is the rate of correctly classified events in all detected events,

  Ppr = TP / (TP + FP)
- **Accuracy (Acc)** measures the overall system performance over all classes of beats above metrics are used to quantify the performance of the proposed system with respect to detection of each class of beat in ECG samples from the databases. Accuracy is usually the most crucial metric for determining overall system performance, however due to large variation in the number of beats from different classes in the long-term ECG dataset, sensitivity, specificity, and positive predictivity can too be critical and relevant performance criteria for medical diagnosis.

CHAPTER 4

GENETIC EVALUATION FOR FEATURE SELECTION

4.1 GOAL OF FEATURE SELECTION

Based on the literature review there are multiple reasons for doing feature selection. In some cases the amount of features can make construction of an induction model hard, either because the model cannot fit in memory or construction would take too long. Creating a limited subset can help construct a model within these constraints.One of the reasons for building classifiers, and in particular decision trees, is to create an understandable decision model. Now an unpruned decision tree made from instances with 100 features is most likely going to be very messy. If the number of features is reduced in a manner that does not significantly reduce accuracy this may get a smaller model. This would make the model more readable and it could also be less over-fitted.

Removing redundant and irrelevant features can help improve the performance of classifiers. This is done by reducing the potential for over- fitting and, in the case of redundant features, selecting those that work best.

4.2 OVERVIEW OF EVOLUTIONARY ALGORITHM

A typical Pattern Recognition System (PRS) consists of three tasks namely, data acquisition and/or preprocessing, feature analysis, classification . In the first step, data are collected and then these raw data may be preprocessed. Preprocessing may involve noise reduction, normalization and conversion of raw data into suitable form for pattern recognition. After obtaining the data, good features are extracted by mapping data to other domain or a subset of good features is selected from the available features. This process finds useful features to obtain an efficient and improved solution to a given problem. Success of pattern recognition depends on the features used. Finally, in the classification, the actual task of PRS is performed. Classification involves assigning a class label to a given pattern while clustering finds homogeneous subgroups in data.If data are available for pattern recognition, then it may require schemes for feature extraction/selection and for classification.

Computational intelligence includes mainly three tools: Artificial Neural Networks (ANNs), Fuzzy logic and Evolutionary algorithms (EAs) are most popular among computational intelligence methods. They have good learning and generalization abilities but sometimes lack interpretability and may work as a black box.

Evolutionary Algorithms (EAs) evolve desired solution to a given problem using biologically inspired operations like crossover, mutation and the Darwinian Principle of the Survival of the Fittest. At first, typically a population of representation of possible solutions is randomly generated. Each representation is called a chromosome. Then, genetic operations and selection operations are implemented on the current population to create the next population. This is motivated by the hope that new generation will be better than previous generation. The process of evolution is continued till the desired solution is obtained or till the termination criteria are satisfied. The computation using Evolutionary algorithms is called Evolutionary Computation. Figure 4.1 represents a general evolutionary algorithm for feature selection.

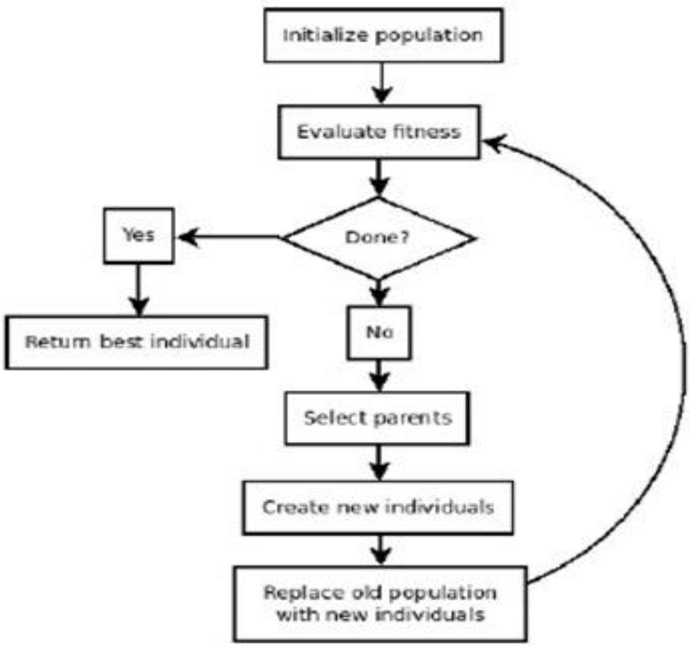

Figure 4.1 A general evolutionary algorithm

4.3 GENETIC ALGORITHM

Genetic Algorithms (GA) are direct, parallel, stochastic method for global search and optimization, which imitates the evolution of the living beings, described by Charles Darwin. GA are part of the group of Evolutionary Algorithms (EA). The evolutionary algorithms use the three main principles of the natural evolution: reproduction, natural selection and diversity of the species, maintained by the differences of each generation with the previous.

Genetic Algorithms works with a set of individuals, representing possible solutions of the task. The selection principle is applied by using a criterion, giving an evaluation for the individual with respect to the desired solution. The best-suited individuals create the next generation. The large variety of problems in the engineering sphere, as well as in other fields, requires the usage of algorithms from different type, with different characteristics and settings

4.3.1 Introduction

High dimensional feature set can negatively affect the performance of pattern or image recognition systems. In other words, too many features sometimes reduce the classification accuracy of the recognition system since some of the features may be redundant and non informative. Different combinatorial set of features should be obtained in order to keep the best combination to achieve optimal accuracy. In machine learning and statistics, feature selection, which is also called variable selection, attribute selection or variable subset selection, is the process of obtaining a subset of relevant features (Probably optimal) for use in machine model construction. Genetic Algorithm has been known to be a very adaptive and efficient method of feature selection.

Genetic Algorithms (GA) is an optimization technique, population-based and algorithmic search heuristic methods that mimic natural evolution process of man. The operations in a GA are iterative procedures manipulating one population of chromosomes (solution candidates) to produce a new population through genetic functionals such as crossover and mutation (in a similar way to Charles Darwin evolution principle of reproduction, genetic recombination, and the survival of the fittest). The terminology between human genetic and GA can be summarised as shown in Table 4.1

Table 4.1 Comparative Terminology between human genetic and GA

S.NO	Human Genetic	GA Terminology
1	Chromosomes	Bit Strings
2	Genes	Features
3	Allele	Feature Value
4	Locus	Bit Position
5	Encoded	String
6	Phenotype	Decoded Genotype

4.4 GA-BASED FEATURE SELECTION

The five important issues in the GA are chromosome encoding, fitness evaluation, selection mechanisms, genetic operators and criteria to stop the GA The GA operates on binary search space as the chromosomes are bit strings. The GA manipulates the finite binary population in similitude of human natural evolution. To begin with, an initial population is created (mostly randomly) and evaluated using a fitness function. For binary chromosome employed in this work, a gene value ’1’ depicts that the particular feature indexed by the position of the ’1’is selected. Otherwise, (i.e if it is ’0’), the feature is not selected for chromosomal evaluation. Using the positional index of features indexed by the ’1s’, the chromosomes are then ranked and based on the rankings, the top n fittest kids (Elitism of size n) are selected to survive to the next generation. The fitness evaluation is done through Algorithm. After the elite kids are pushed automatically to the next generation, the remaining kids (individuals) in the current population are allowed to genetically pass through the functional crossover and mutation to form crossover and mutation kids respectively. The three kids viz elite, crossover and mutation then form the new one population (new generation). Crossover (a genetic functional) is a combination of two individuals (chromosomes) to form a crossover kids. Mutation operator on the other hand, is used for genetic perturbation of the genes in each chromosomes through bits flipping depending on the mutation probability.

4.4.1 Fitness evaluation

For GA to select a subset of features, a fitness function (a driver for the GA) must be defined to evaluate the discriminative capability of each subset of features. The fitness of each chromosome in the population are evaluated using kNN-based fitness function.

kNN algorithm solves classification problem by looking for the shortest distance between the test data and training sets. The kNN algorithm computes Euclidean distance between test data x_{test} and the training sets and then find the nearest point (shortest distance) from the train set from the training set to the test data.This distance is expressed in equ 4.1.

$$D(x_{test}, x_i) = \sqrt{\sum_{m=1}^{M} (x_{test} - x_i)^2} \tag{4.1}$$

The kNN count each category min the class information (accumulated as (*count* x_m) using 3 Nearest Neighbors and then report classification results and errors based on the

$$\text{argmax}(\text{count}(x_m)) \tag{4.2}$$

Subjected to

$$\sum_{i=1}^{M} count(x_m) = class \tag{4.3}$$

The iterations involved in running the GA ensures that the GA reduce the error rate and picks the individual with the least (best) fitness value since error rate is reported for each chromosomes involved and the smallest of error rate is finally picked up by the GA.

$$fit = \frac{\alpha}{N_f} + \exp\left(\frac{-1}{N_f}\right) \tag{4.4}$$

α = kNN-Based classification error.

N_f = Cardinality of the selected features.

The algebraic structure of this equation ensures the learning of the GA, error minimization and reduced number of features selected.

Algorithm

1: procedure fit ()

2: Feat Index Indices of ones from Binary Chromosome

3: New Data Set Data Set indexed by Feat Index

4: Num Feat Number of elements in Feat Index

5: 3 ← Num Neighbors kNN

6: kNN Error ← Classifier KNN (DataSet, Class Information, Num Neighbors kNN)

7: Return kNN Error

8: **e**nd procedure

The carefully chosen fitness function enabled the GA to minimize classification error from kNN the best fitness and mean fitness should be close in value as the GA reaches the termination condition. The stall generation is number of generations produced by the GA since the last upgrade of the fitness value. More features are extracted from the signals used. With the application of GA for dimensionality reduction, more discriminating features were obtained

4.5 GENETIC PROGRAMMING

Genetic programming based feature selection was used to improve the classification results and reduce the dimensionality of the data. Genetic programming as a form of evolutionary algorithm and an extension of genetic algorithms, is used as the feature selection. The major difference between the GP and GA approaches lies in the way that each algorithm solves the problem under consideration. With a GA-based solution, the basic form of the solution is predefined; the GA is able to optimize parameters of the solution, however not the actual structure of the solution. GP by comparison has control over both the structure and the parameters of the solution to the problem.

Genetic Programming (GP) technique provides a framework for automatically creating a working computer program from a high-level problem statement of the problem. Genetic programming achieves this goal of automatic programming by genetically breeding a population of computer programs using the principles of Darwinian natural selection and biologically inspired operations. The operations include most of the techniques discussed in the previous sections.

The main difference between genetic programming and genetic algorithms is the representation of the solution. Genetic programming creates computer programs in the LISP or scheme computer languages as the solution. LISP is an acronym for LISP Processor and was developed by John McCarthy in the late 1950s.Unlike most languages, LISP is usually used as an interpreted language. This means that, unlike compiled languages, an interpreter can process and respond directly to programs written in LISP. The main reason for choosing LISP to implement GP is due to the advantage of having the programs and data have the same structure, which could provide easy means for manipulation and evaluation.

Genetic programming is the extension of evolutionary learning into the space of computer programs. In GP the individual population members are not fixed length character strings that encode possible solutions to the problem at hand, they are programs that, when executed, are the candidate solutions to the problem. These programs are expressed in genetic programming as parse trees, rather than as lines of code.

4.5.1 Process of Genetic programming

GP based feature selector is used to select the most significant input features for the classifier from a large set of possible features obtain from the extracted features. The purpose of GP is to try to maximize the extra information content in the sample of ECG signals, and it implicitly maximizes the separation between different conditions within the data. First, an initial population with a chosen number of individuals is generated on a random basis meaning that there is no human influence, or bias, in the selection of original features. Calculated features data set is fed as the inputs to the initial population. Each individual represents a transformation network, which tries to transform raw data into information for classification. In terms of the usefulness of each individual for classification, a fitness value is assigned to each individual by fitness function. Therefore, the members with the best fitness values survive from the current generation and will be chosen as the origins of the next generation, only the elite will survive the natural selection. This mechanism allows the feature to evolve in a direction toward the best classification performance, thus achieving the best selection of features. At the beginning of the next generation, three operations reproduction, crossover, and mutation are conducted to produce new members based on the surviving member. If the termination criterion is met, the best solution is preserved.

4.5.2 Primitive Operations

GP evolves tree individuals representing possible solutions to the problem. A population of such individuals is randomly created and then evolved by probability of genetic operations.

Crossover - GP carries out a crossover operation to create new individuals with a probability which controls the occurrence of the crossover throughout generations. Two new individuals are generated by selecting compatible nodes randomly from each parent and swapping them. Figure 4.2 represents some example of crossover operation.

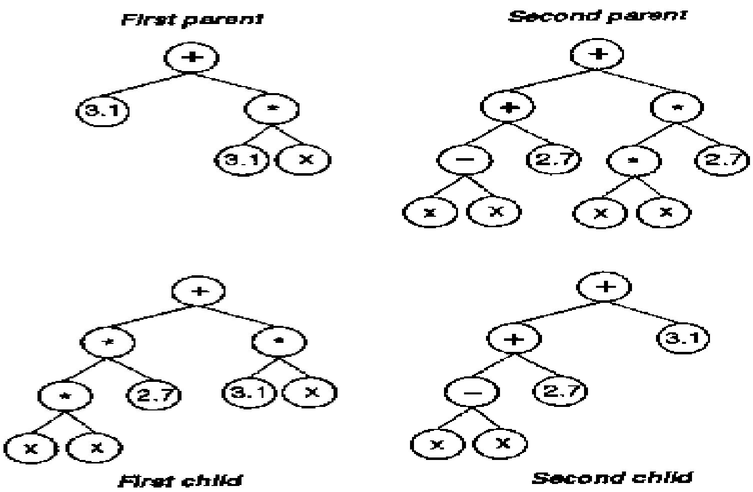

Figure 4.2 Crossover operation

Mutation - The mutation operation is performed by the creation of a sub-tree at a randomly selected node with the probability. First, for a given parent, there is an index assigned to each node for identification. A random index number is generated to indicate the place where mutation will happen. The node is located, then the tree downstream from this node is deleted and a new sub tree is generated from this node, exactly in the same way as growing initial population.Figure 4.3 represents general form of mutation operation.

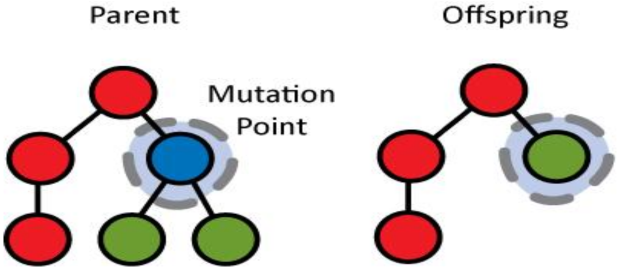

Figure 4.3 A general mutation operation

Reproduction: The reproduction operation is performed by copying individuals to the next population without any change in terms of a certain probability. All these three operation happen within one generation based on three probabilities.

4.5.3 Parameters of GP

Table 4.2 Parameters of GP

Population size	**50**
Maximum number of evaluated individuals	**1920**
Type of selection	**Tournament selection**
Type of mutation	**Point mutation**
Type of crossover	**One point crossover**
Max tree depth	**5**
Probability of crossover	**0.5**
Probability of mutation	**0.5**

Steps of GP:

A typical implementation of GP involves the following steps.

Initialization: GP begins with a randomly generated population of solutions of size P

Termination: GP is terminated when termination criteria are satisfied. Unlike GA , GP will not converge. So, GP is terminated when a desired solution (may be with fitness value 1) is achieved. Otherwise, it is terminated after a predefined number of generations.

Evaluation: A fitness value is assigned to each solution of the population.

Next Generation: The current population is replaced by a new population by means of applying genetic operations probabilistically.

This completes one generation. Go to termination step and repeat if termination criteria are not satisfied.

CHAPTER 5

DESIGN METHODOLOGY AND REQUIREMENTS

5.1 WORK FLOW

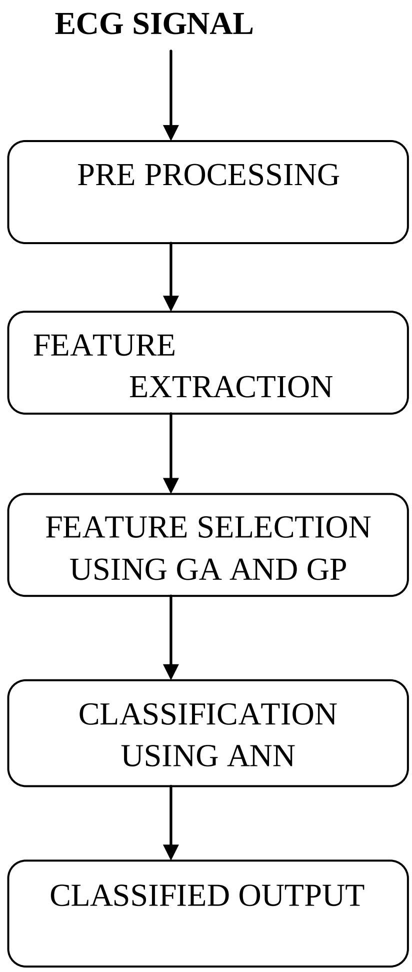

Figure 5.1 Workflow of proposed method

5.2 HARDWARE REQUIREMENTS

Processor : CORE^{Tm} i5

Hard Disk: 800 GB

RAM : 2 GB

5.2.1 SOFTWARE REQUIREMENTS

Operating system: Windows 7

Programming language: MATLAB2014a

5.3 SOFTWARE DESCRIPTIONS – MATLAB2014A

MATLAB (Matrix Laboratory) is a numerical computing environment and fourth-generation programming language. Developed by Math Works, MATLAB allows matrix manipulations, plotting of functions and data, implementation of algorithms creation of user interfaces, and interfacing with programs written in other languages including C, C++, Java, and FORTRAN.

Although MATLAB is intended primarily for numerical computing, an optional toolbox uses the MuPAD symbolic engine, allowing access to symbolic computing capabilities. An additional package, Stimulant, adds graphical multi-domain simulation and Model-Based Design for dynamic and embedded systems.

In 2004, MATLAB had around one million users across industry and academia.MATLAB users come from various backgrounds of engineering, science, and economics. MATLAB is widely used in academic and research institutions as well as industrial enterprises.

Features

- High-level language for technical computing
- Development environment for managing code, files, and data
- Interactive tools for iterative exploration, design, and problem solving
- Mathematical functions for linear algebra, statistics, Fourier analysis, filtering, optimization, and numerical integration
- 2-D and 3-D graphics functions for visualizing data
- Functions for integrating MATLAB based algorithms with external applications and languages, such as C, C++, Fortran, Java, COM, and Microsoft® Excel

CHAPTER 6

RESULTS AND DISCUSSION

6.1 DENOISING PERFORMANCE MEASURES

Original ECG is denoised and obtain clean ECG signal by using different wavelet transforms.The performance measures are calculated for Symlet, Haar, Biorthogonal and Daubechies wavelets by using three data base namely, MIT-BIH, NSTDB, INCARTDB.

Table 6.1 Denoising performance measures of INCARTDB

INCARTDB	HARR	BIORTHO	SYMLET	DB
variance	1.398	3.5208	1.7029	1.7953
Mean square error	$4.416e^{0.6}$	$7.1716e^{0.6}$	$5.0745e^{0.6}$	$5.0751e^{.6}$
Mean absolute error	1.8771	2.2387	1.9767	1.9715

Table 6.1 Shows the performance measures (MSE, MAE and Variance) of Haar, Biorthogonal, Symlet and Daubechies wavelets. Haar wavelet gives higher performance than other wavelet for INCARTDB.

Table 6.2 Denoising performance measures of MIT-BIH

MIT-BIH	HARR	BIORTHO	SYMLET	DB
Variance	0.55483	0.77618	0.16646	0.22365
Mean square error	3.4341	3.9491	3.3811	3.3424
Mean absolute error	1.8029	1.8106	1.9933	1.9964

Table 6.2 Shows the performance measures (MSE, MAE and Variance) of Haar, Biorthogonal, Symlet and Daubechies wavelets. Haar wavelet gives higher performance than other wavelet for MIT-BIH.

Table 6.3 Denoising performance measures of NSTDB

NSTDB	HARR	BIORTHO	SYMLET	DB
Std-deviation	4.0098	8.2811	4.4684	5.8511
Mean square error	16.731	70.455	21.294	36.244
Mean absolute error	3.2571	6.3031	3.255	3.9364

Table 6.3 Shows the performance measures (MSE, MAE and Standard deviation) of Haar, Biorthogonal, Symlet and Daubechies wavelets. Haar wavelet gives higher performance than other wavelet for NSTDB

6.2 CLASSES FOR CLASSIFICATION

- Normal
- Premature ventricular contraction
- Bundle branch block(left and right)
- Atria premature contraction

6.3 SIMULATION RESULTS FOR ANALYSIS OF WAVELET

- Total number of sample points taken=10,000
- Samples (cycles) taken from each record =20
- No of features taken from each cycle =6
- Sampling frequency=250 Hz

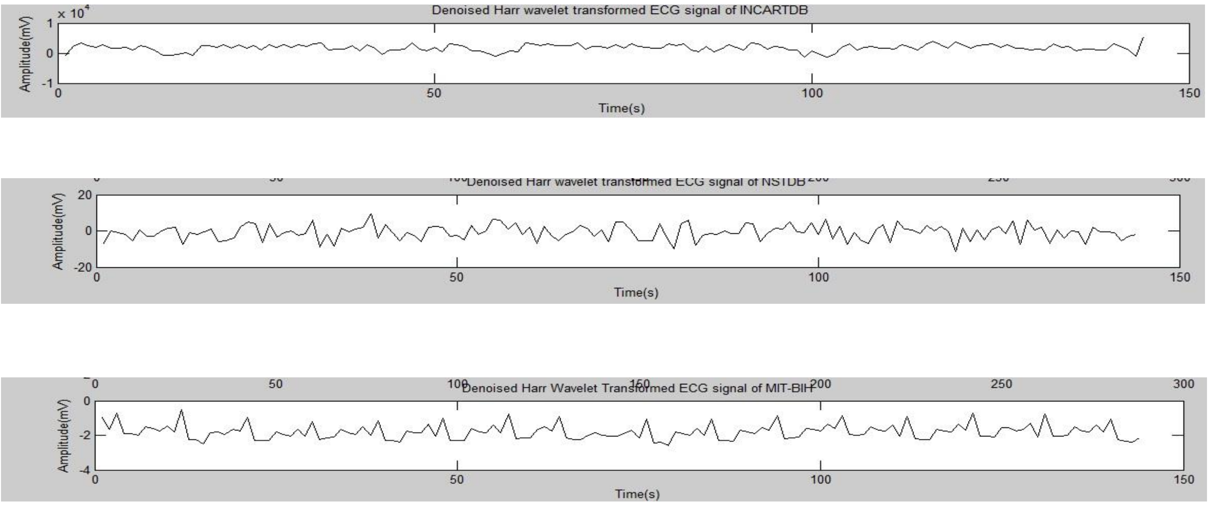

Figure 6.1 Denoised ECG signal using haar wavelet for different database

Figure 6.1 Shows Denoised ECG signal using haar wavelet for MIT-BIH, NSTDB and INCARTDB database.

6.4 FEATURES EXTRACTED FOR MIT-BIH

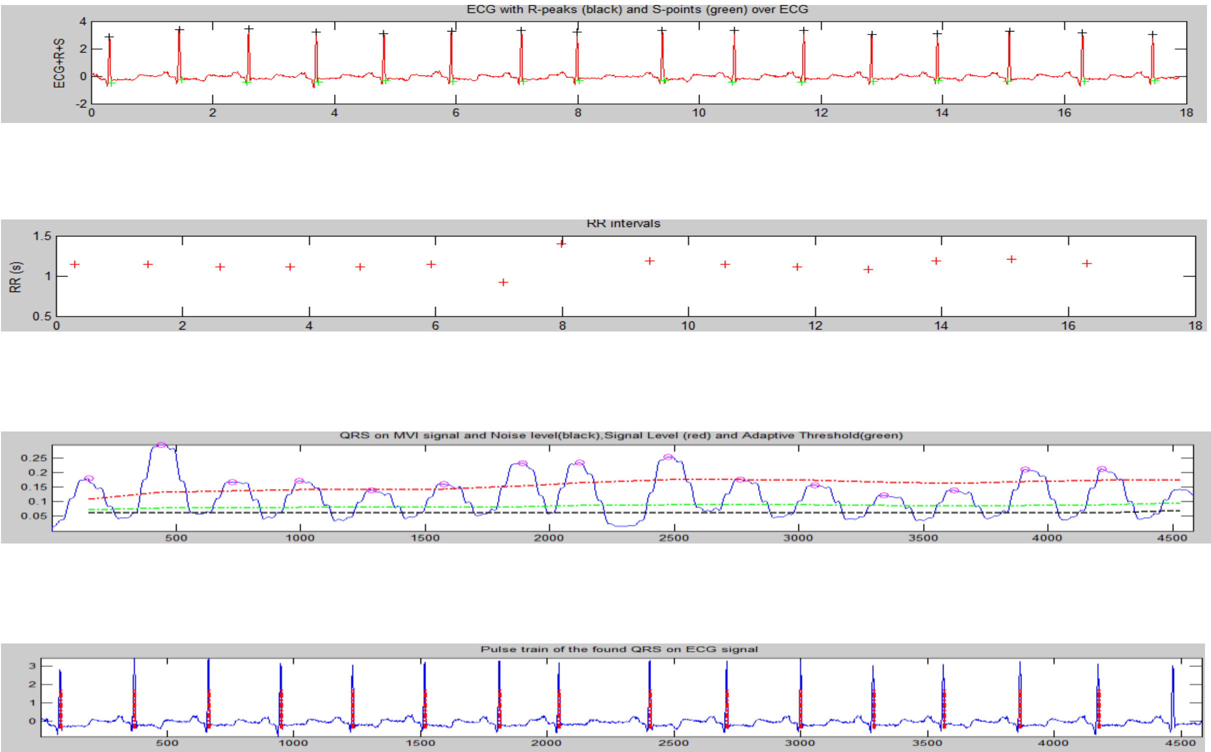

Figure 6.2 Extracted parameters R-amp, RR interval, QRS complex using pantompkins's

6.5 DESCRIPTION ABOUT NEURAL NETWORK TOOLBOX

Neural network in Matlab can be dealt very easily due to the existing new neural network toolbox. This provides a complete set of functions to design an implement various computational intelligence processes. The major neural network operation includes function fitting, classification and pattern recognition, clustering and dynamic modelling and prediction. These all are performed by means of various functions and even can be implemented using the Graphical User Interface (GUI). Many of the applications can be simulated using the "neural network tool" Simulink block present in Matlab–Simulink toolbox.Tools for building custom graphical user interfaces.Toolbox itself having several examples for all kind of operation mentioned above, it automatically generated scripts and model.

6.5.1 Data division (dividerand)

The easiest way to create a neural network is to use one of the network creation functions. To investigate how this is done, one can create a simple, two-layer feed forward network

6.5.2 Scaled conjugate gradient back propagation (trainscg)

It is a network training function that updates weight and bias values according to the scaled conjugate gradient method.

Trainscg can train any network as long as its weight, net input, and transfer functions have derivative functions. Back propagation is used to calculate derivatives of performance perf with respect to the weight and bias variables X.

Training stops when any of these conditions occurs:

- The maximum number of epochs (repetitions) is reached.
- The maximum amount of time is exceeded.
- Performance is minimized to the goal.
- The performance gradient falls below min_grad.
- Validation performance has increased more than max_fail times since the last time it decreased (when using validation).

6.5.3 Levenberg-Marquardt backpropagation (trainlm)

trainlm is a network training function that updates weight and bias values according to Levenberg-Marquardt optimization. It is often the fastest backpropagation algorithm in the toolbox, and is highly recommended as a first-choice supervised algorithm, although it does require more memory than other algorithms.

6.6 CLASSIFICATION USING ANN

Table 6.4 Performance measures of ANN

	Mean Squared error
Train Performance	0.8340
Test Performance	0.9114
Validation Performance	0.8781
Overall performance	0.8515

Table 6.4 Shows the Mean Squared Error (MSE) for training, testing and validation performance of ANN classifier.

6.6.1 Performance metrics for ANN classifier

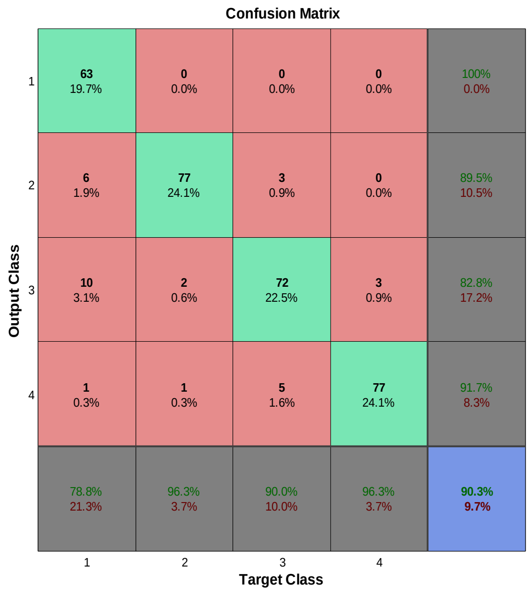

Figure 6.3 Performance obtained from confusion matrix for ANN

Figure 6.3 Shows the confusion matrix, classified and misclassified samples for four classes and its performance of ANN classifier.

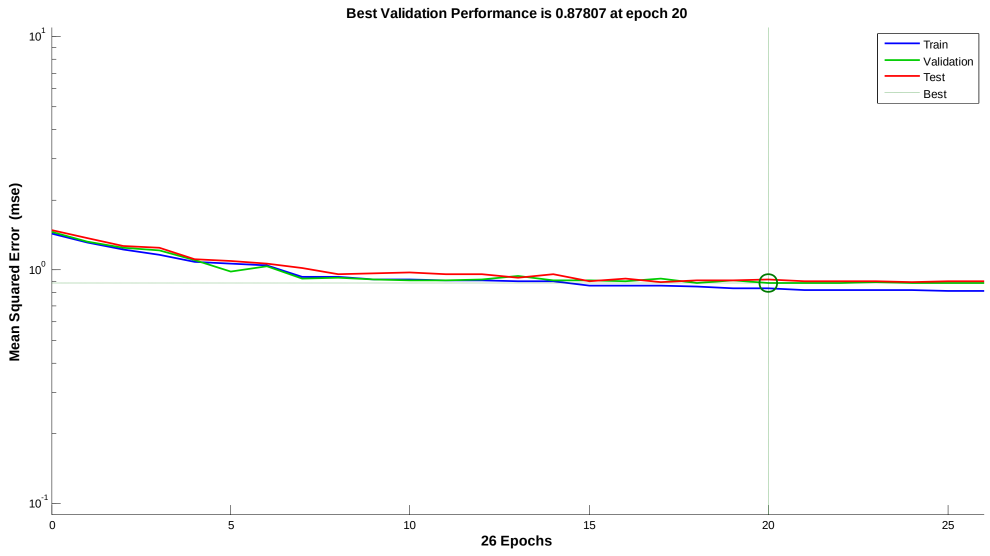

Figure 6.4 Best validation performance for ANN

Figure 6.4 Shows the best validation performance of 0.87807 at epoch 20 of ANN classifier

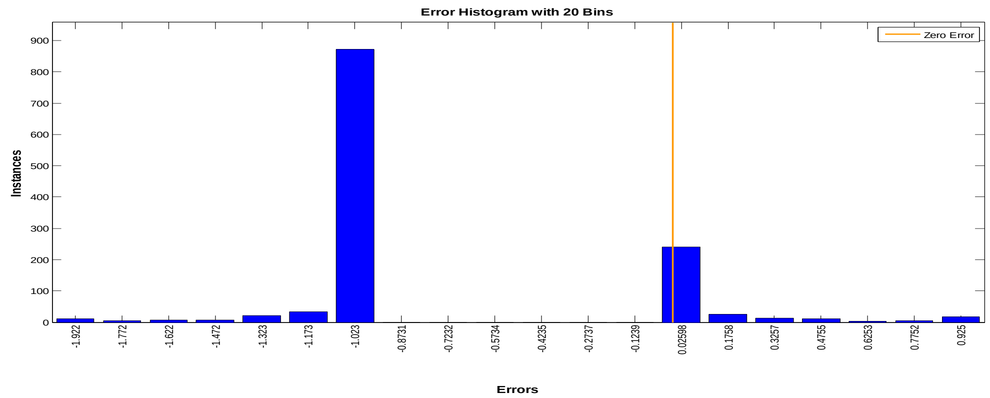

Figure 6.5 Error histogram plot for ANN

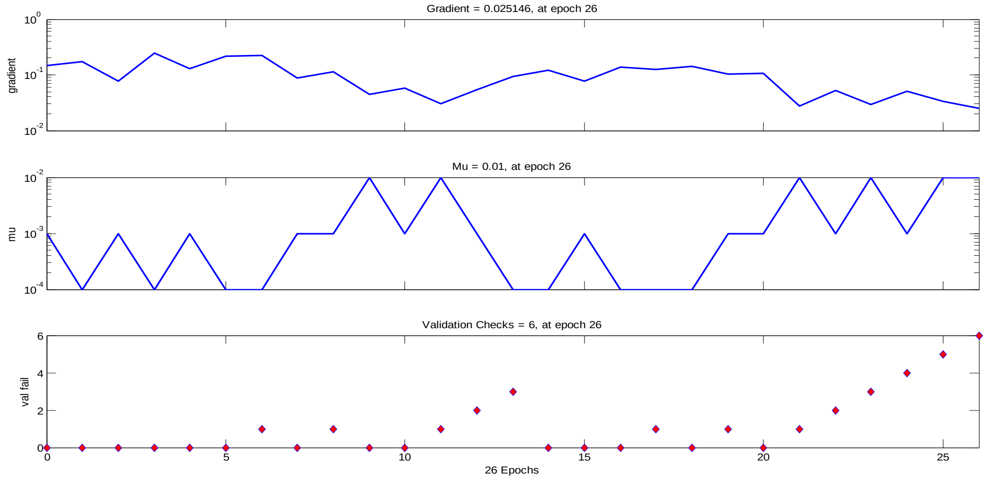

Figure 6.6 Best training state performance for ANN

Figure 6.6 shows the best training state performance of gradient 0.025146 at epoch 26 of ANN classifier.

6.7 ANN CLASSIFIER USING GA BASED FEATURES

Table 6.5 Performance measures of GA based ANN

	Mean Squared error
Train Performance	0.8138
Test Performance	0.8677
Validation Performance	0.8676
Overall performance	0.8293

Table 6.5 Shows the Mean Squared Error (MSE) for training, testing and validation performance of GA with ANN classifier.

6.7.1 Performance metrics for ANN classifier using GA based feature

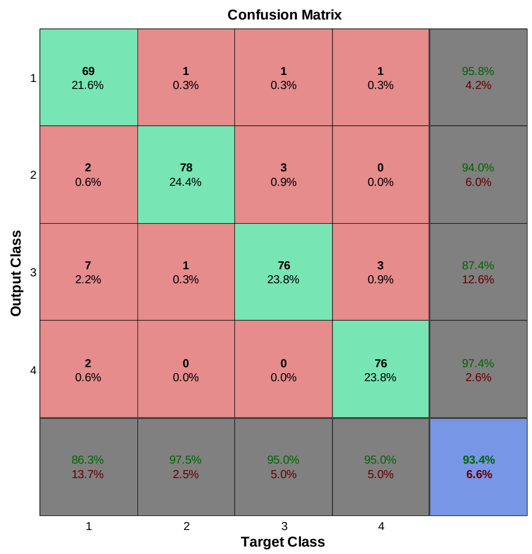

Figure 6.7 Performance obtained from confusion matrix for GA based ANN

Figure 6.7 Shows the confusion matrix, classified and misclassified samples for four classes and its performance of ANN classifier.

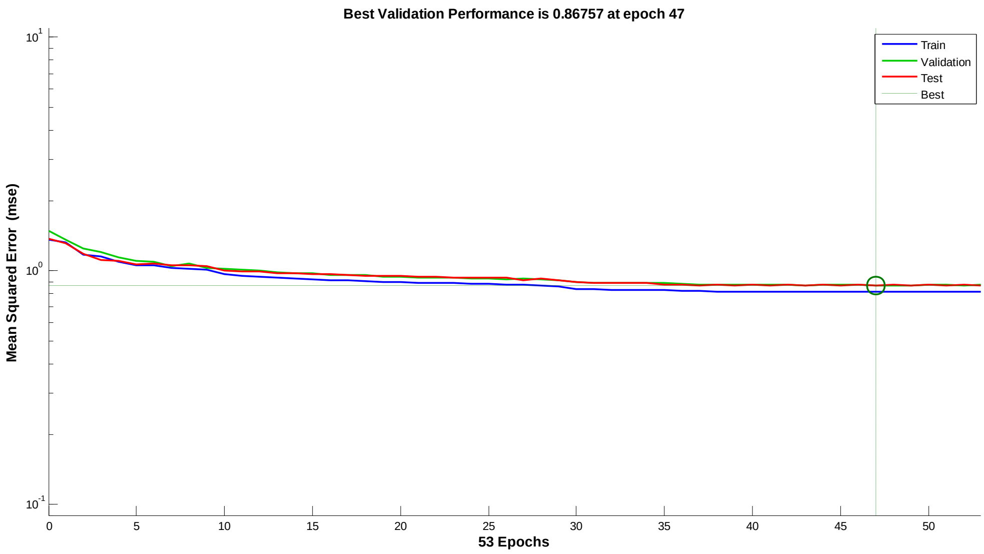

Figure 6.8 Best validation performance for GA based ANN

Figure 6.8 Shows the best validation performance of 0.88767 at epoch 47 of GA with ANN classifier.

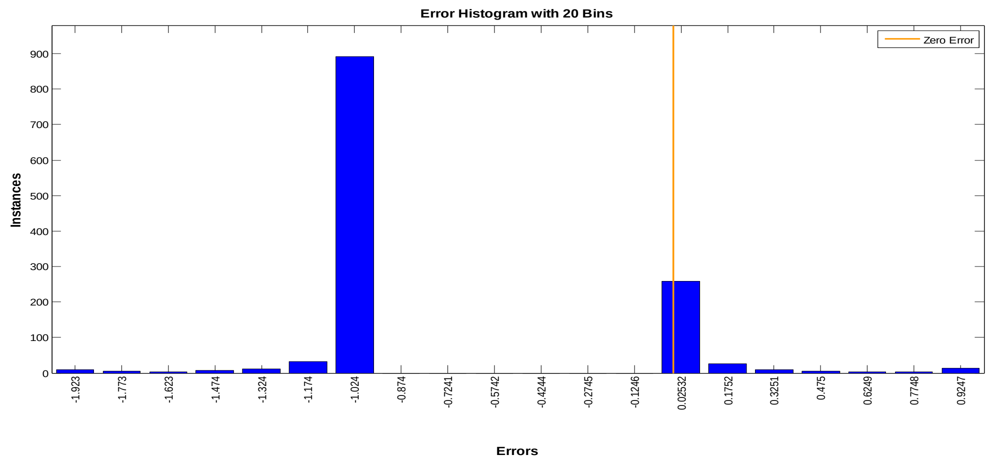

Figure 6.9 Error histogram plot for GA based ANN

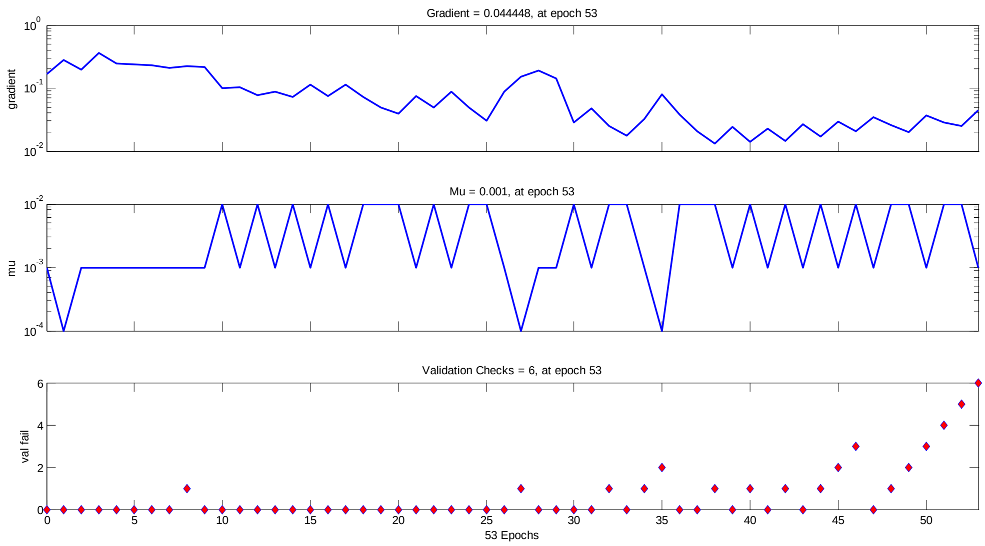

Figure 6.10 Best training state performance for GA based ANN

Figure 6.10 shows the best training state performance of gradient 0.044448 at epoch 53 of GA with ANN classifier.

6.8 ANN CLASSIFIER USING GP BASED FEATURES

Table 6.6 Performance measures of GP based ANN

	Mean Squared error
Train Performance	0.8375
Test Performance	0.8556
Validation Performance	0.8437
Overall performance	0.8399

Table 6.6 Shows the Mean Squared Error (MSE) for training, testing and validation performance of GP with ANN classifier.

6.8.1 Performance metrics for ANN classifier using GP based feature

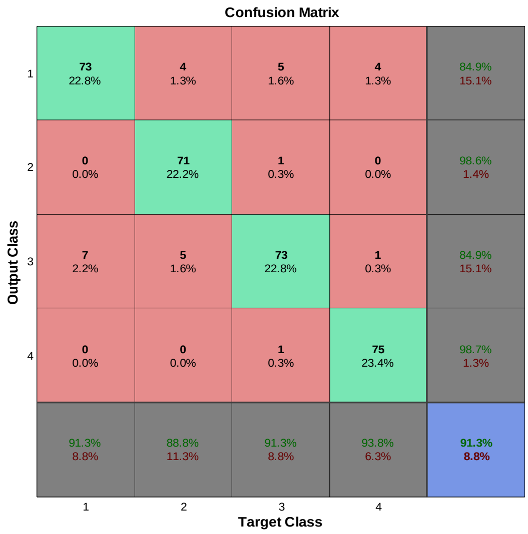

Figure 6.11 Performance obtained from confusion matrix for GP based ANN

Figure 6.11 Shows the confusion matrix, classified and misclassified samples for four classes and its performance of GP with ANN classifier.

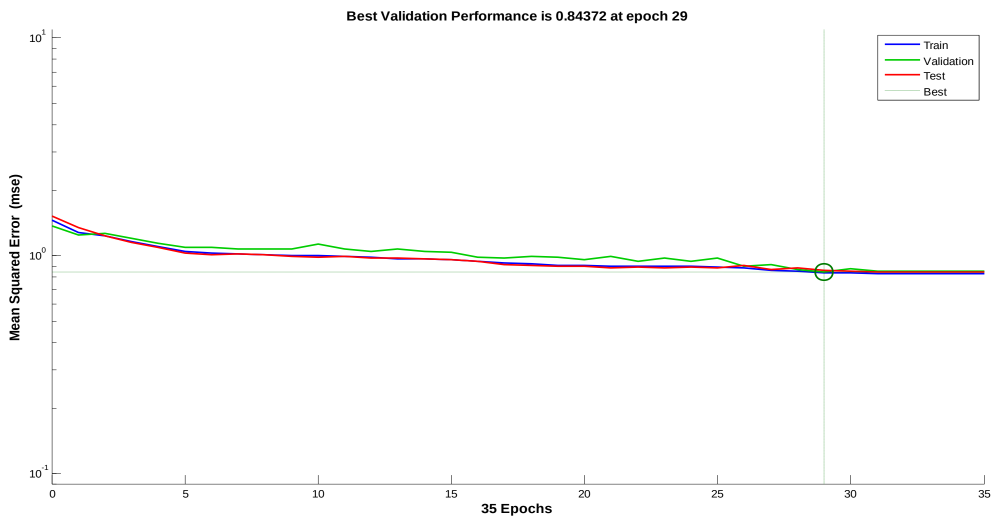

Figure 6.12 Best validation performance for GP based ANN

Figure 6.12 Shows the best validation performance of 0.84372 at epoch 36 of GP with ANN classifier

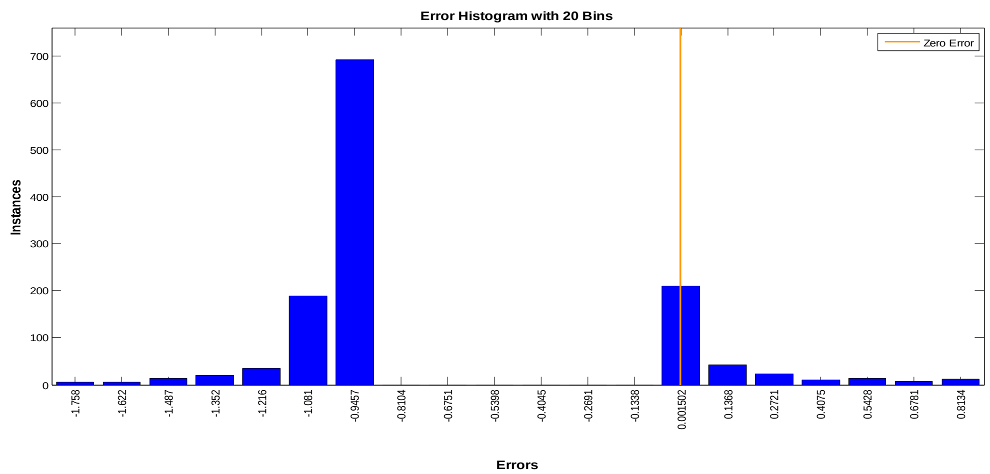

Figure 6.13 Error histogram plot for GP based ANN

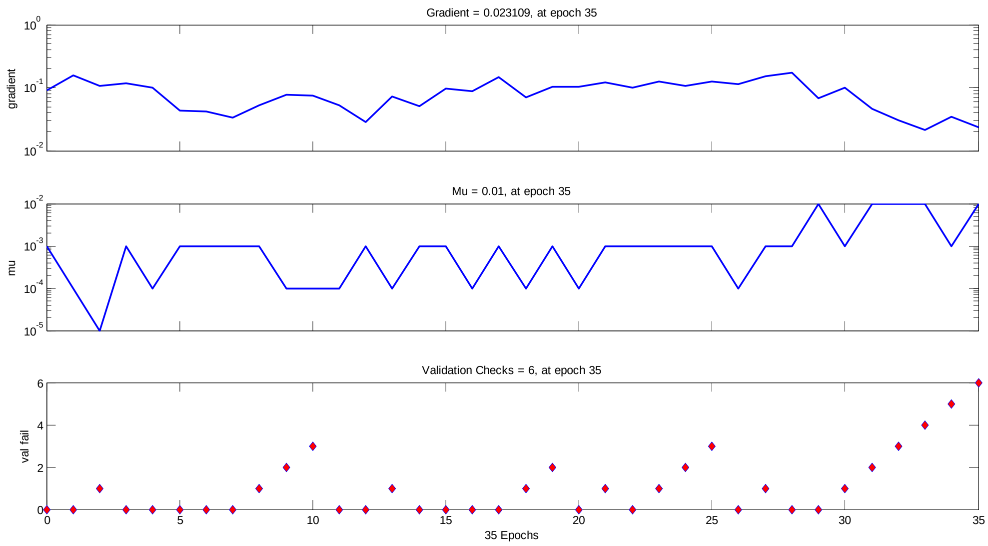

Figure 6.14 Best training state performance for GP based ANN

Figure 6.14 shows the best training state performance of gradient 0.023109 at epoch 35 of GP with ANN classifier.

6.9 PERFORMANCE MEASURES COMPARISON

Table 6.7 Overall Performance comparison measures

Performance meas	ANN	GA with ANN	GP with ANN
Accuracy	90.3%	93.4%	91.3%
Sensitivity	91.2%	94.7%	92.4%
Specificity	78.6%	81.2%	80.6%
Mean square error	0.8515	0.8123	0.8399

Table 6.7 Shows the performance measures (Accuracy, Specificity, Sensitivity, MSE) of ANN classifier, GA with ANN classifier, GP with ANN classifier. GA with ANN classifier gives higher performance than other two approaches.

CHAPTER 7

CONCLUSION AND FUTURE WORK

In this project, the normal and cardiac arrhythmias such as PVC, APC, BBB has been detected and classified using ANN classifier with Genetic Algorithm and Genetic Programming feature selection approaches. Initially six features were extracted from the samples, datasets were trained and tested, classified using Artificial Neural Network (ANN). Although ANN shows fair results, more irrelevant signals are incorporate in to the feature subsets that may affects the performance. Feature selection approach using genetic evolution (GA and GP) significantly reduced the redundant features and incorporate more information content to the input signals. On classification and evaluating the performance metrics ANN classifier provides 90.3% of accuracy, 91.2% of sensitivity and 78.6% of specificity, ANN classifier using GA based features provides 93.4% of accuracy, 94.7% of sensitivity and 81.2% of specificity, ANN classifier using GP based features provides 91.3% of accuracy, 92.4% of sensitivity and 80.6% of specificity. It is concluded that genetic evolution based ANN classifier provides better performance than ANN classifier. More number of cardiac disorders (Cardiac Arrhythmia) can be classified using Genetic evolution based ANN classifier using other databases like NSTDB and INCARTDB.

REFERENCES

[1] Osowski S., Markiewicz T. and Tran Hoai L., "Recognition and classification system of arrhythmia using ensemble of neural networks", Measurement, Vol. 41, No. 6, pp. 610-617, 2008.

[2] Tsipouras M. G. and Fotiadis D. I., "Automatic arrhythmia detection based on time and time- frequency analysis of heart rate variability", Computer Methods and Programs in Biomedicine, Vol. 74, No. 2, pp. 95-108, 2004.

[3] de Chazal P. and Reilly R. B., "A Patient- Adapting Heartbeat Classifier Using ECG Morphology and Heartbeat Interval Features"

[4] Pooja Gupta and Sulochana Wadhwani "Feature Selection by Genetic Programming, Artificial Neural Network-based Machine Condition Monitoring" International Journal of Engineering and Innovative Technology, Vol 1, No 4, 2012.

[5] Babatunde Oluleye, Armstrong Leisa and Jinsong Leng "A Genetic Algorithm-Based Feature Selection" International Journal of Electronics Communication and Computer Engineering, Vol 5, No 4, 2010.

[6] S.Karpagachelvi, M.Arthanari and M.Sivakumar "ECG Feature extraction technique - A survey approach" International Journal of Computer Science and Information Security, Vol 4, No 1, 2010.

[7] Cardiac Arrhythmia Classification Using Neural Networks with Selected Features - International Conference on Computational Intelligence: Modeling Techniques and Application, Vol 5, No 4, 2013.

[8] Masood Ahmad Arbab, Gul Muhammad Khan and Ali Mahmud Sahibzada "Cardiac Arrhythmia Classification using Cartesian Genetic Programming Evolved Artificial Neural Network, Vol 20, No 9, 2013.

[9] Krishna Prasad and J. S. Sahambi, "Classification of ECG Arrhythmias using Multi Resolution Analysis and Neural Networks," IEEE Transactions on Biomedical Engineering, vol. 1, pp. 227-231, 2003.

[10] Y. H. Hu, S. Palreddy, and W. Tompkins, "A Patient Adaptable ECG Beat Classifier Using A Mixture Of Experts Approach", IEEE Transactions on Biomedical Engineering vol. 44, pp. 891–900, 1997

CONFERENCE PUBLICATION

The paper presented on **"ECG Signal Enhancement Using Wavelet transforms"** in 2015 IEEE International Conference on Engineering and Technology (ICETECH'15) at Rathinam Technical Campus, Coimbatore.