Chapter 1
Introduction

1.1 General View

Electrocardiogram(ECG)is the most commonly used biomedical signal in clinical diagnostics of the heart, The word “electrocardiogram” is a combination of three words: electro, pertaining to electric signal; cardio, which translates into heart; and gram, which stands for recording. The recording of the electric activity of the heart is called ECG.[1]

A cardiac muscle contraction is a direct result of the cellular electric excitation described by the ECG. The depolarization initiates the shortening of each individual muscle cell. The electric activation of each cell is an indication of the functioning of that cell. Therefore, the ECG is the result of depolarization of the heart muscle in a controlled repetitive fashion By tracking the process of Electric depolarization of the cardiac muscle cells, an impression of the heart’s functionality can be formed and used to recognize regions in the heart structure that are not functioning to specifications and may require medical attention. Any deviation from the typical ECG observed in the recorded electric depolarization signal is analyzed and classified as a certain cardiac disorder.[1]

Electrocardiogram (ECG), a non-invasive technique is used as a primary diagnostic tool for cardiovascular diseases. Cardiac arrhythmias including heart attack, stroke, and hypertension, is caused by disorders of the heart and blood vessels and is by far the leading cause of death around the world. However, most heart attacks and strokes could be prevented if some method of pre-monitoring and pre-diagnostic can be provided. In particular, early detection of abnormalities in the function of

the heart, called cardiac arrhythmias, can be valuable for the clinicians. The electrocardiogram (ECG) plays an important role in the process of monitoring and preventing cardiac arrhythmias.[2]

The ECG signal provides key information about the electrical activity of the heart. This electrical activity is related to the impulse that travels through the heart, which determines its rate and rhythm. Physicians use this information to diagnose many cardiac disease conditions. The amplitude and duration of the wave contains useful information about the nature of disease afflicting the heart.

The electrical wave is due to depolarization and repolarization of Na+ and k- ions in the blood.[2]

1.2 Problem Statement

Detection of ECG arrhythmias is necessary for the treatment of patients for diagnosing the heart disease at the early stage. it is very difficult for doctors to analyze long ECG records in the short period of time, also human eye is poorly suited to detect the morphological variation of ECG signal this even for the most experienced cardiologist .

The basic complexity and mechanistic and clinical interrelationships of arrhythmias often brings about diagnostic difficulties for treating physicians and primary health care professionals creating frequent misdiagnoses.

Many works reported on arrhythmia beat classification shows that there is a need to improve the classification accuracy when used for huge database. Also most of the methods use complex mathematical features imposing lot of computational burden while evaluating these features .

1.3 Thesis Objectives

1.3.1 General Objective

To design an automated system for diagnosis of arrhythmia discords to assist the care provider in the diagnosing of cardiac arrhythmia with higher diagnostic accuracy with significant reduction in the cost .

1.3.2 specific objective

To develop a highly accurate program for ECG classification as normal and abnormal signals ,and to differentiation between APB and PB using MATLAB. This program will help the medical staff in diagnosing, treating and preventing of heart problems and therefore reducing the mortality.

1.4 Methodology

The thesis methodology is about how to develop an classification software tool for ECG waveform using MATLAB as a programming language and to test the accuracy of this software. The ECG Classification described in this research will consist of 4 modules. ECG signals used for the research were obtained from the Physionet Database (PhysionetBank). A set of programs from the Physionet was used to import ECG records each of which consists of data file, attribute file and header file in Matlab. Matlab and its Wavelet toolbox were used for ECG signal processing and analysis.

1. Preprocessing Module: The ECG signals were preprocessed by filtering it to remove the baseline wander, the power line interference, and the high frequency noise, hence enhancing the signal quality, and omitting the equipment and the environmental effects.

2. Feature Extraction Module: This module is important in the post processing of ECG readings, because a good set of features will have sufficient discriminative power to facilitate the diagnosis or investigation.

3. Classifier Module.

4. Output module: Gives the performance metrics of the Analyzer and diagnosis result .

1.5 Thesis Layout

The project thesis consists of six chapters which are :

Chapter one: introduces general view for research, problem statement, objectives and methodology, Chapter tow: literature review, Chapter three :describes theoretical background. Chapter four: illustrates research methodology, Chapter five : shows results and their discussions and Chapter sex : includes conclusion and future recommendations.

Chapter 2

Literature Review

2-1Classification of ECG Arrhythmias Using Discrete Wavelet Transform And Neural Networks

Sarkaleh, M, K, an expert system for Electro-Cardio Gram (ECG) arrhythmia classification is proposed. Discrete wavelet transform is used for processing ECG recordings, and extracting some features, and the Multi-Layer perceptron(MLP) neural network performs the classification task. Two types of arrhythmias can be detected by the proposed system. Some recordings of the MIT-BIH arrhythmias database have been used for training and testing this neural network based classifier. The simulation results show that the classification accuracy of this algorithm is 96.5% using 10 files including normal and two arrhythmias. In this paper, a neural network based system for automatic ECG arrhythmias classification was proposed. The system used 10 recordings from the MIT-BIH arrhythmias database for training as well as testing our classifier. The proposed system consists of two phases : the feature extraction phase and the classification phase. In the first phase, moving average filter is employed to eliminate the baseline noise from the ECG signals. Then the DWT is applied on filtered signal and some features from the wavelet coefficients are extracted. In the second phase, the extracted feature are used to train an MLP NN as the classifier. The simulation results demonstrated the proposed system could be employed for the classification of the ECG arrhythmias with a recognition rate of 96.5%, when 13 neurons were in the hidden layer in traingdx, 11 neurons in trainrp algorithm and 14 neurons in trainlm[3].

2-2 A novel Approach for Classification of ECG Arrhythmias : Type-2 fuzzy Clustering Neural Network

Ceylan, R; Ozbay, Y and Karlik, B, presents an improved classifier for automated diagnostic systems of electrocardiogram (ECG) arrhythmias. This diagnostic system consist of a combined fuzzy clustering neural network algorithm for classification of ECG arrhythmias using type-2 fuzzy c-means clustering(T2FCM) algorithm and neural network. Type-2 fuzzy c-mean clustering is used to improve performance of neural network. The aim of improving classifier's performance is to constitute the best classification system with high accuracy rate for ECG beats. Ten types of ECG arrhythmia(normal beat, sinus bradycardia, ventricular tachycardia, sinus arrhythmia, atrial premature contraction, paced beat, right bundle branch block, left bundle branch block, atrial fibrillation and atrial flutter) obtained from MIT-BIH database were analyzed. The classification accuracy of an improved classifier in training and testing, namely type-2 fuzzy clustering neural network(T2FCNN), was compared with neural network(NN) and fuzzy clustering neural network(FCNN).In T2FCNN architecture, decision making has two stages: forming of the new training set obtained by selection of the best arrhythmia for each arrhythmia class using T2FCM and classification using neural network trained on the new training set. The results are demonstrated that the proposed diagnostic systems achieved high accuracy rate (99%).we hope that the performance of the method will be better, if the number of the beats is increased for the training. This technique is obtained by incorporating the techniques of type-2 fuzzy c-means clustering method and back propagation learning and combining their advantages. So, it can be said that the structure, which is more beneficial structure than conventional BPNN and FCNN to recognize and classify ECG signals, is obtained.[4]

2-3 Automated Detection Of Arrhythmias Using Different Intervals Of Tachycardia ECG Segments With Convolutional Neural Network

Acharya, U, and Rajendra,, Our cardiovascular system weakens and is more prone to arrhythmia as we age. An arrhythmia is an abnormal heartbeat rhythm which can be life-threatening atrial fibrillation (A fib), atrial flutter(A fl), and ventricular fibrillation (V fib) are the recurring life-threatening arrhythmias that affect the elderly population. An electrocardiogram (ECG) is the principal diagnostic tool employed to record and interprets ECG signals. These signals contain information about the different types of arrhythmias. However, due to the complexity and non-linearity of ECG signals, it is difficult to manually analyze these signals. Moreover, the interpretation of ECG signal is subjective and might vary between the experts.

Hence, a computeraided diagnosis (CAD) system is proposed. In this work, we present a convolutional network (CNN) technique to automatically detect the different ECG segments. In this work, we have used ECG signals of two seconds and five seconds durations without QRS detection. We achieved an accuracy, sensitivity, and specificity of 92.5%, 98.09%, and 93.13% respectively for two seconds of ECG segments.

We obtained an accuracy of 94.90%, the sensitivity of 99.13%, and specificity of 81.44% for five seconds of ECG duration. Further, the robustness of the proposed system can be improved by using large arrhythmia database with more number of V fib, A fib, A fl, and Nsr ECG segments.[5]

2-4 Automatic Classification Of ECG Signal With Features Extracted Using Wavelet Transform and Support Vector Machines

Sambhu, D and Umesh, A, C, accurately classified and differentiated normal and abnormal heartbeats such as left bundle branch block (LBBB), right bundle branch block (RBBB), atrial premature contractions (APC) and premature ventricular contractions (PVC), atrial premature beat (APB), paced beats and fusion beats with adequate levels of accuracy.

At first the multi resolutions analyses of ECG signal is done to de-noised and extract 25 features. The mother wavelet used for decomposition was db4.

The classification is implemented by using one against one (OA0) and support vector machine (SVM).7 SVM's were trained and final grouping is done by maximum voting.ECG signals are obtained from the open source MIT-BIB cardiac arrhythmia database.

Experiments reveal that the overall classification accuracy is well above 97% for all the classes. Some models gave high amount of accuracy while others gave less accuracy. The reason is that all the feature do not work equally well for all the classes.

Some features that work well for some type of classes will affect like noise for some other classes. The solution for this is implementing feature selection algorithm. The performance of the wavelet based feature extraction was a suitable one. The advantage of SVM classifier using the given feature vector is its ease of implementation and simplicity.

In future we will try to develop this project in fully open platform. We can improve the accuracy by including more features .[6]

2-5 Classification of ECG Arrhythmias Using Adaptive Neuro-Fuzzy Inference System and Cuckoo Optimization Algorithm

Ebrahimi, A and Addeh, J, presents a hybrid method for automated diagnostic system of electrocardiography arrhythmias. The proposed method includes three main modules including the de-noising module, the classifier module and the optimization module. In the de-noising module, the stationary wavelet transform is proposed for noise reduction of the electrocardiogram signals. In the classifier module, the adaptive neuro-fuzzy inference system is investigated. In the test stage, 3-fold cross validation method has been applied to the MIT-BIH arrhythmia database for evaluating the capability of the proposed method.

The simulation results show that the proposed method has high recognition accuracy. In the past decades, many automatic ECG arrhythmia classification system have been developed using computational intelligence techniques. In this research, we proposed a method for ECG classification based on wavelet transform and ANFIS. The complexity of the recognition system is very low in comparison with other works.

The highest level of accuracy obtained by ANFIS using unprocessed data was 94.26%. the proposed method improves the accuracy up to 97.14% by using approximation coefficients of wavelet transform as the classifier inputs.[7]

2-6 Detection of Ventricular Fibrillation Using Hilbert Transforms, Phase-space Reconstruction, and Time-domain Analysis

Sang-hong; Chung; Kyung-Yong; LIM, and Joon S, proposes feature extraction using Hilbert transforms, phase-space reconstruction, and time domain analysis to detect ventricular fibrillation and normal sinus rhythm from electrocardiogram (ECG) episodes. We implemented three per-processing steps to extract features from ECG episodes. In the first step, we use Hilbert transforms to

extract peaks. In the second step, we use statistical methods and extract four features from the peaks. In the final step, we extract four features using statistical methods based on the Euclidean distance between the origin (0, 0) and the peaks after the peaks are plotted in two-dimensional phase-space diagram. By applying time domain analysis directly to the series of successive peak-to-peak interval values, we extract seven additional features.

Using a neural network with weighted fuzzy membership functions(NEWFM), the applied the nonoverlaparea distribution measurement method, and from 15 initial features, we selected 11 minimum features exhibiting the highest accuracy. Then, we applied the 11 minimum features as inputs to the NEWFM and recorded sensitivity, specificity, and accuracy values of 79.12, 89.58, and 87.51%, respectively.[8]

2-7 ECG Beats Classification Using Mixture of Features

DAS, Manab K, ARI, Samit, proposes the design of an efficient system for classification of the normal beet (N), ventricular ectopic beat (V), super-ventricular ectopic beat (S), fusion beat (F), and unknown beat (Q) using a mixture of features. In this paper, two different feature extraction methods are proposed for classification of ECG beats : (i) s-transform based features along with temporal features and (ii) mixture of ST and WT based features along with temporal features. The extracted feature set is independently classified using multilayer perceptron neural network.

The performances of proposed features are compared with the other existing methods. Experimental results demonstrate that the proposed features provide better detection sensitivity than WT based features. The overall results of the proposed extracted feature methods also show an effective and efficient approach in computer-aided diagnosis of heart disease based on ECG signals.

As a result, automatic detection and classification of cardiac electrophysiology using biomedical signal processing techniques have become a critical aspect of clinical monitoring. [9]

2-8 Novel ECG Signal Classification Based on KICA Nonlinear Feature Extraction

LI and Hongqiang, Electrocardiogram (ECG) signal feature extraction is important in diagnosing cardiovascular diseases. This paper presents a new method for nonlinear feature extraction of ECG signal by combining principal component analysis (PCA) and kernel independent component analysis (KICA).

The proposed method first uses PCA to decrease the dimensions of ECG signal training set and then employs KICA to calculate the feature space for extracting the nonlinear features. Support vector machine (SVM) is utilized to determine the nonlinear features of ECG signal classification. Genetic algorithm is also used to optimize the SVM parameters. The proposed method is advantageous because it does not require a huge amount of sampling data, and this technique is better than traditional strategies to select optimal features in the multi-domain feature space. computer simulations reveal that the proposed method yields more satisfactory classification results on the MIT-BIH arrhythmia database, reaching an overall accuracy of 97.78%.This paper presents a novel method of ECG signal classification, which is based on KICA and PCA nonlinear feature extraction. the proposed method does not require a large amount of samples and is better than traditional methods in selecting optimal features in the multi-domain space vector. Thus, this is new strategy can be used efficiently for ECG diagnosis. [10]

2-9 Optimization of ECG Classification By Means of Feature Selection

MAR and Tanis, tackles the ECG classification problem by means of a methodology, which is able to enhance classification performance while simultaneously reducing the computational resources, making it specially adequate for its application in the improvement of ambulatory settings. For this purpose, the sequential forward floating search (SFFS) algorithm is applied with a new criterion function index based on linear discriminates. This criterion has been devised specifically to be a quality indicator in ECG arrhythmia classification. Based on this measure, a comprehensive feature set is analyzed with the SFFS algorithm, and the most suitable subset returned is additionally evaluated with a multilayer perceptron (MLP) to assess the robustness of the model. The first set contained features used in a previous study; it served to evaluate the capability of the method to improve previously proposed classifier systems. The second had a twofold purpose : first, the capability of the method to narrow down comprehensive datasets was assessed; second, by including only features suitable for online monitoring in the examined set, further insight into which are the most suited features for this kind of setting was accomplished. In what follows, the returned subsets were used to carry out ECG classification using two different classifier paradigms. The achieved performance results on both models prove the suitability of the index driven SFFS algorithm to improve performance while reducing classifiers in the heartbeat classification field. [11]

2-10 A Multi-Stage Neural Network Classifier For ECG Events

Hosseini, H; Reynolds, K. J; Powers, D, a multi-stage network including two multilayer perceptron (MLP) and oneself organizing map (SOM) networks is present. The input of the network is a combination of independent features and the compressed Electro Cardio Gram (ECG) data.The proposed network as a form of data fusion performs better than using the raw data or individual features. We classified six common ECG wave forming using ten ECG records of the MIT-BIH arrhythmia database. An a verge. Recognition rate of 0.883 was achieved within a short training and testing time. An a verge Recognition rate of 0.883 was achieved combining multi-stage networks and performing local tests on the detected waveforms. [12]

2-11 Wavelet/Mixture of Expert Network Structure for EEG Signals Classification

Ubeyli, E, D, Mixture of experts (ME) is modular neural network architecture for supervised learning. This paper illustrates the use of ME network structure to guider model selection for classification of electroencephalogram (EEG) signals. Expectation-maximization (EM) algorithm was used for training the ME so that the learning process is decoupled in a manner that fits well with the modular structure. The EEG signals were decomposed into time-frequency representations using discrete wavelet transform and statistical features were calculated to depict their distribution. The ME network structure was implemented for classification of the EEG signals using the statistical features as inputs. To improve classification accuracy, the outputs of expert networks were combined by a gating network simultaneously trained in order to stochastically select the expert that is performing the best at solving the problem. The EEG signals were classified with the accuracy

of 93.17% by the ME network structure, which were higher than that of the stand-alone neural network models. [13]

2-12 ECG Beat Classification Using Neuro-fuzzy Network

Engin, M, have studied the application on the fuzzy-hybrid neural network for electrocardiogram (ECG) beat classification. Instead of original ECG beat, we have used; autoregressive model coefficients, higher-order cumulate and wavelet transform variances as features. Tested with MIT/BIH arrhythmia database, we observe significant performance enhancement using proposed method .[14]

2-13 Analysis of ECG signal for Detection of Cardiac Arrhythmias

Prakash, J, Electrocardiogram (ECG), a noninvasive technique is used as a primary diagnostic tool for cardiovascular diseases.

A cleaned ECG signal provides necessary information about the electrophysiology of the heart diseases and ischemic changes that may occur. It provides valuable information about the functional aspects of the heart and cardiovascular system. The objective of the thesis is to automatic detection of cardiac arrhythmias in ECG signal.

Recently developed digital signal processing and pattern reorganization technique is used in this thesis for detection of cardiac arrhythmias. The detection of cardiac arrhythmias in the ECG signal consists of following stages: detection of QRS complex in ECG signal; feature extraction from detected QRS complexes; classification of beats using extracted feature set from QRS complexes. In turn automatic classification of heartbeats represents the automatic detection of cardiac arrhythmias in ECG signal. Hence, in this thesis, we developed the automatic algorithms for classification of heartbeats to detect cardiac arrhythmias in ECG signal.

QRS complex detection is the first step towards automatic detection of cardiac arrhythmias in ECG signal. A novel algorithm for accurate detection of QRS complex in ECG signal is proposed in chapter 2 of this thesis. The detection of QRS complex from continuous ECG signal is computed using autocorrelation and Hilbert transform based technique. The first differential of the ECG signal and its Hilbert transformed is used to locate the R-peaks in the ECG waveform. The autocorrelation based method is used to find out the period of one cardiac cycle in ECG signal.

The advantage of proposed method is to minimize the large peak of P-wave and T-wave, which helps to identify the R-peaks more accurately. Massachusetts Institute of Technology Beth Israel Hospital (MIT-BIH) arrhythmias database has been used for performance analysis. The experimental result shows that the proposed method shows better performance as compared to the other two established techniques like Pan-Tompkins (PT) method and the technique which uses the difference operation method (DOM).

For detection of cardiac arrhythmias, the extracted features in the ECG signal will be input to the classifier. The extracted features contain both morphological and temporal features of each heartbeat in the ECG signal. Twenty six dimension feature vector is extracted for each heartbeat in the ECG signal which consist of four temporal features, three heartbeat interval features, ten QRS morphology features and nine T-wave morphology features.

Automatic classification of cardiac arrhythmias is necessary for clinical diagnosis of heart disease. Many researchers recommended Association for the Advancement of Medical Instrumentation (AAMI) standard for automatic classification of heartbeats into following five beats: normal beat (N), super-ventricular ectopic beat (S), ventricular ectopic beat (V), fusion beat (F) and unknown beat (Q). The beat classifier system is adopted in this thesis by first training a local-classifier using the

annotated beats and combines this with the global-classifier to produce an adopted classification system. The Multilayer perceptron back propagation (MLP-BP) neural network and radial basis function (RBF) neural network are used to classify the cardiac arrhythmias. Several experiments are performed on the test dataset and it is observed that MLP-BP neural network classifies ECG beats better as compared to RBF neural network.[15]

2-14 Analysis of Electrocardiograph (ECG) Signal for the Detection of Abnormalities Using MATLAB

Ojha, D, K, The proposed method is to study and analyze Electrocardiograph (ECG) waveform to detect abnormalities present with reference to P, Q, R and S peaks. The first phase includes the acquisition of real time ECG data. In the next phase, generation of signals followed by pre-processing. Thirdly, the procured ECG signal is subjected to feature extraction.

The extracted features detect abnormal peaks present in the waveform Thus the normal and abnormal ECG signal could be differentiated based on the features extracted. The work is implemented in the most familiar multipurpose tool, MATLAB.

This software efficiently uses algorithms and techniques for detection of any abnormalities present in the ECG signal. Proper utilization of MATLAB functions (both built-in and user defined) can lead us to work with ECG signals for processing and analysis in real time .applications. The simulation would help in improving the accuracy and the hardware could be built conveniently.[16]

2-15 Application of Artificial Neural and Fuzzy-Neural Networks to QRS Detection and PVC Diagnosis

Prof. Al-Alaoui, M, A, The Electrocardiogram (ECG) is the most clinically used biological signal and it is the means of detecting several cardiac diseases and abnormalities.

Among various abnormalities related with functioning of the human heart, Premature Ventricular Contraction (PVC) is the contraction of the lower chambers of the heart (the ventricles) which occur earlier than usual, because of abnormal electrical activity of the ventricles. The QRS complex detection is a substantial procedure in analyzing the ECG; however, it is not always straightforward.

The great variety of QRS detection algorithms reflects the need for a reliable QRS detection in cardiac signal processing, however, the currently achievable detection rates reflect only the overall performance of the detectors which generally hide the problems still present in case of noisy or abnormal ECG signals which indicates that a satisfying solution to these problems is still not found.

In our final year project we have designed and implemented an enhanced artificial neural network QRS detection and PVC diagnosis algorithms in MATLAB.

In our implementation, the MIT- BIH Arrhythmia Database, which is the most popular standard database and the most widely used in QRS publications, was used.[17]

2-16 Matlab Implementation of ECG Signal Processing

V.Viknesh and P. Ram Prashanth, Signal processing today is performed in vast majority of systems for ECG analysis and interpretation. The objective of ECG signal processing is manifold and comprises the improvement of measurement accuracy and reproducibility and the extraction of information not readily available from the signal through visual assessment.

In many situations, the ECG is recorded during ambulatory or strenuous conditions such that the signal is corrupted by different types of noise, sometimes originating from other physiological process of the body. Hence noise reduction represents another important objective of ECG signal processing. The paper mainly focuses on implementing the present day trends and procedures in the processing of ECG signals using software (MATLAB).

The implementation process helps us to understand the drawbacks and difficulties of such methods and gives us an opportunity to work out towards finding a better solution. Such a solution would satisfy the scope of improvement expected in the technologies, used at present. Keywords: ECG, baseline wander, power line interference, QRS detection.[18]

Chapter 3

Theoretical Background

This section aims to demonstrate a quick glance into the heart anatomy and physiology, ECG wave's characteristics and a few of cardiac arrhythmias and their waves which is essential to the understanding of the use of diagnostic ECG.

3.1 Heart Anatomy

The heart contains four chambers that is right atrium, left atrium, right ventricle, left ventricle and several atrioventricular and sinoatrial node as shown in the figure3.1. The two upper chambers are called the left and right atria, while the lower two chambers are called the left and right ventricles. The atria are attached to the ventricles by fibrous, non-conductive tissue that keeps the ventricles electrically isolated from the atria. The right atrium and the right ventricle together form a pump to the circulate blood to the lungs. Oxygen-poor blood is received through large veins called the superior and inferior vena cava and flows into the right atrium. The right atrium contracts and forces blood in to the right ventricle, stretching the ventricle and maximizing its pumping (contraction) efficiency. The right ventricle then pumps the blood to the lungs where the blood is oxygenated. Similarly, the left atrium and the left ventricle together form a pump to circulate oxygen-enriched blood received from the lungs (via the pulmonary veins) to the rest of the body [2].

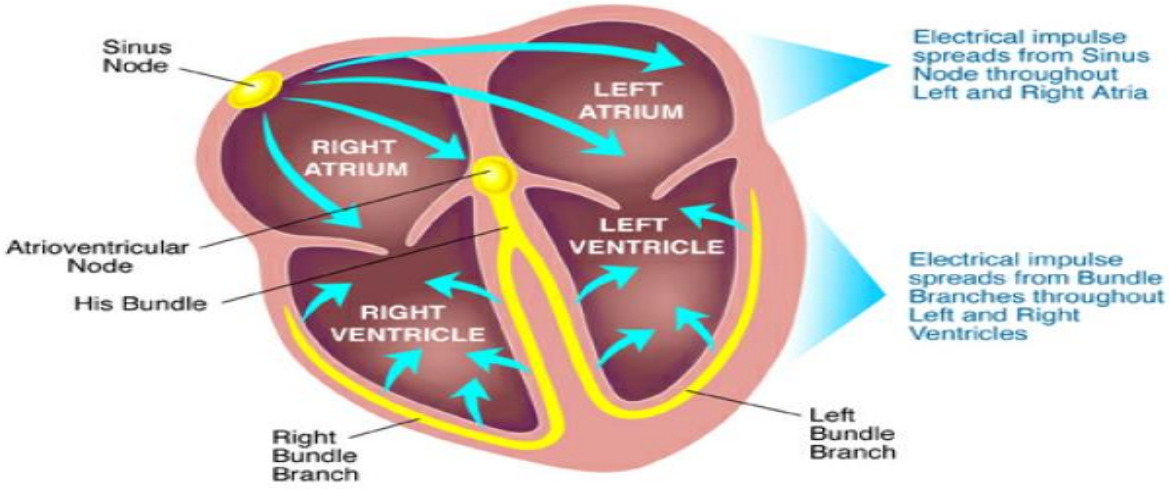

Figure 3.1 the Heart conduction system.[2]

3.2 Normal Heart Function and the Electrophysiology of the heart

To fully understand the electrical and mechanical problems that cause arrhythmias and Cardiovascular diseases, it is important to understand the electrical and mechanical functions of a normal, healthy heart. A concrete understanding of these differences will provide the information necessary to diagnose, detect, and treat these conditions. Electrical activity in the body drives the mechanical function of the heart. Action potentials are responsible for the contraction of cardiac muscle cells, which is essential for pumping blood through the body. Some cardiac cells, unlike nerve and skeletal muscle cells, are able to generate their own action potentials in order to achieve steady and rhythmic contraction. These cells are known as auto-rhythmic cardiac cells and work with contractile cardiac cells to pump the heart.[2]

These contractile cells, which make up the majority of cardiac cells, are responsible for heart contraction. Auto-rhythmic cardiac cells display a pacemaker like activity; their membrane potential depolarizes after each action potential until

they reach threshold and generate another action potential that will cause cardiac contraction.[1]

The auto-rhythmic cells do not all have the same rate of depolarization. Those cells with the fastest rate of action potential initiations are found on the sinoatrial (SA) node, which is a small region in the right Arial wall. Because the SA node has the fastest rate of action potential initiation, it is known as the pacemaker region of the heart and is known to be the driving force for the rest of the heart. After an action potential is initiated in the SA node, the excitation travels through the remainder of the heart for a full contraction. The action potential first spreads through the atria. The atria must contract before the ventricles can contract, as Arial contraction is responsible for pumping the blood to the ventricles which then contract to pump the blood through the rest of the body. The electrical pathway through which the action potentials follow is demonstrated in Figure 3.2

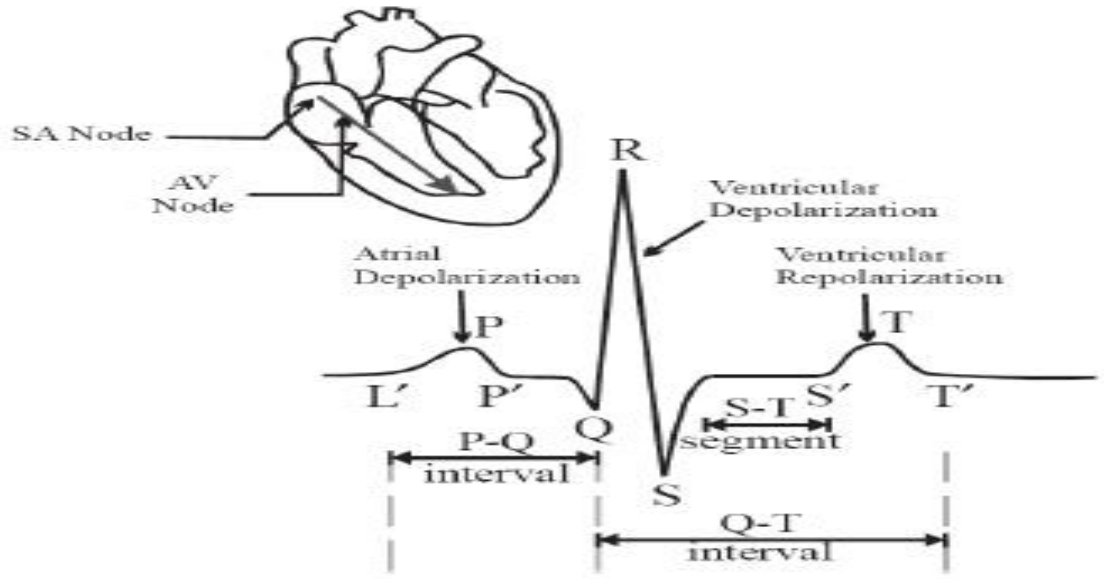

Figure 3-2 Ideal ECG signal. The SA node causes atrial depolarization(P complex). The AV node Causes ventricular depolarization (QRS complex). The T complex indicates ventricular repolarization.[2]

An electrocardiogram (ECG) is a recording of this electrical activity of the cardiac cycle. As mentioned, when cardiac muscle cells depolarize and repolarize, electrical currents are generated and spread through the chambers and the tissues that surround the heart. Some of this electricity reaches the surface of the body, and the ECG is able to detect these electrical impulses. An ECG is obtained through electrode placement on a person's skin that detects this electricity and outputs a voltage-time reading. It is important to note that the reading is not a recording of the actual electricity activity of the heart; it is a representation of the overall spread of electricity during cardiac cell depolarization and repolarization.[2]

The ECG signal provides the following information of a human heart :

• heart position and its relative chamber size
• impulse origin and propagation
• heart rhythm and conduction disturbances
• extent and location of myocardial ischemia
• changes in electrolyte concentrations
• drug effects on the heart.

ECG does not afford data on cardiac contraction or pumping function [1].

Early ECG systems were just recording the signal by printing it in a paper strip. Slight changes in the amplitude and time of the ECG signal from a predefined pattern have been used routinely to detect the cardiac abnormality. Because of the difficulty to elucidate these changes manually, a computer-aided diagnosis system can help in monitoring the cardiac health status. Computeraided cardiac arrhythmia detection and classification can play an important role in the management of cardiovascular diseases [2].

Classification of electrocardiogram (ECG) is an important area in biomedical signal processing. Several algorithms have been developed for classification of

ECG beats. These techniques extract features, which are either temporal or trans-fanned representation of the ECG waveforms, On the basis of these features classification is performed by template-matching , the neural networks or by the other recognition systems [3].The analysis of ECGs can benefit from the wide availability of computing technology.

This study presents some results achieved by carrying out the classification tasks of equipment integrating the most common features of the ECG analysis: arrhythmia, myocardial ischemia, chronic alterations. Several ANN architectures are implemented, tested, and compared with competing alternatives.

The approach, structure, and learning algorithm of ANNs are designed according to the features of each particular classification test .ECG signals are frequently non-stationary meaning that their frequency content changes over time. These changes are the events of interest. Wavelets decompose signals into time-varying frequency (scale) components. Because signal features are often localized in time and frequency, analysis and estimation are easier when working with sparser (reduced) representations. The QRS complex consists of three deflections in the ECG waveform. The QRS complex reflects the depolarization of the right and left ventricles and is the most prominent feature of the human ECG[3].

3.3.1 ECG Leads

The standard ECG has 12 leads: which includes 3 - bipolar leads, 3 - augmented unipolar leads and 3 - chest (precordial) leads. A lead is a pair of electrodes (+ve & -ve) placed on the body in designated anatomical locations & connected to an ECG record [17].

Bipolar leads: record the potential difference between two points (+ve & -ve poles).

Unipolar leads: record the electrical potential at a particular point by means of a single exploring electrode.

Leads I, II and III are commonly referred to bipolar leads as they use only two electrodes to derive a view. One electrode acts as the positive electrode while the other as the negative electrode (hence bipolar) [2].

Table 3.1 Types of leads used in ECG monitoring.[2]

Standard Leads	Limb Leads	Chest Leads
Bipolar leads	**Unipolar leads**	**Unipolar leads**
Lead I	**AVR**	**V1**
Lead II	**AVL**	**V2**
Lead III	**AVF**	**V3**
		V4
		V5
		V6

Einthoven leads:

Lead I: records potentials between the left and right arm.

Lead II: between the right arm and left leg.

Lead III: those between the left arm and left leg .

Goldberger leads are unipolar augmented limb leads in the frontal plane.

Unipolar Limb leads: (when the +ve terminal is on the right arm: **aVR**, left arm **aVL**, or left leg, **aVF)** .[2]

One lead connected to +ve terminal acts as the different electrode, while the other two limbs are connected to the –ve terminal serve as the indifferent (reference) electrode [19]. Wilson leads **(V1–V6)** are unipolar chest leads positioned on the left side of the thorax in a nearly horizontal plane. The indifferent electrode is obtained by connecting the 3 standard limb leads. When used in combination with the unipolar limb leads in the frontal plane, they provide a three dimensional view of the integral vector.

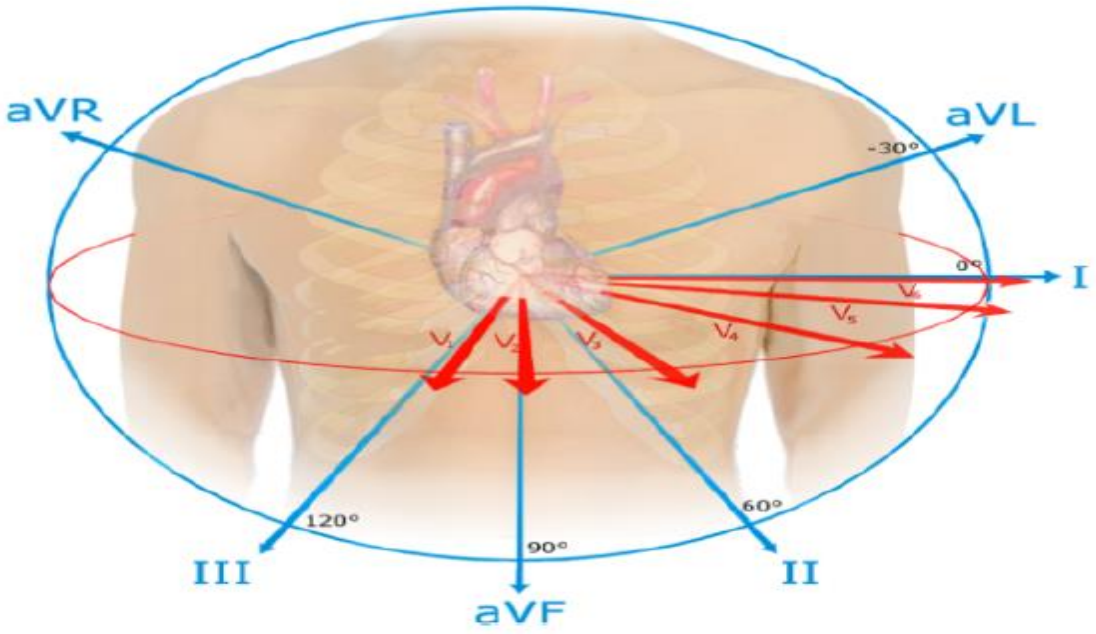

Figure3.3 ECG Leads.[2]

Three augmented unipolar extremity leads (aVR, aVL, and aVF). A unipolar lead records voltages at one point relative to zero potential. The unipolar extremity leads can also be represented by a triaxial diagram.

They are related by the equation aVR + aVL + aVF = 0.

Unipolar Limb leads:

aVR Lead: +ve terminal is on the right arm.

aVL Lead: +ve terminal is on the left arm.

aVF Lead: +ve terminal is on the left leg.

As a general rule, the P-QRS-T pattern in lead I resembles that in lead aVL. Leads aVR and II usually show reverse patterns. Lead aVF usually resembles lead III. One lead connected to +ve terminal acts as the different electrode, while the other two limbs are connected to the–ve terminal serve as the indifferent(reference) electrode [2].

Chest (precordial) leads

The six chest-leads (V1 to V6) record voltages from the heart as directed onto the horizontal plane of the body, from the front and the left side. The chest leads attach to six positions on the chest overlying the 4th and 5th rib spaces.

V1: 4th intercostal space, right sternal edge.

V2: 4th intercostal space, left sternal edge.

V3: between the 2nd and 4th electrodes.

V4: 5th intercostal space in the midclavicular line.

V5: on 5th rib, anterior axillary line.

V6: in the midaxillary line.

To make recordings with the chest leads (different electrode), the three limb leads are connected to form an indifferent electrode with high resistances. The chest leads mainly detect potential vectors directed towards the back. These vectors are hardly detectable in the frontal plane [1]. Since the mean QRS vector is usually

directed downwards and towards the left back region, the QRS vectors recorded by leads V1–V3 are usually negative, while those detected by V5 and V6 are positive [19]. In leads V1 and V2, QRS = -ve because, the chest electrode in these leads is nearer to the base of the heart, which is the direction of electro negativity during most of the ventricular depolarization process. In leads V4, V5, V6, QRS = +ve because the chest electrode in these leads is nearer the heart apex, which is the direction of electro positivity during most of depolarization [18].

3.3.2 Colors in ECG

It's important that you know if you're using a European or American patient cable for monitoring because the colors of the wires will differ.

3.3.3 ECG waves and interval

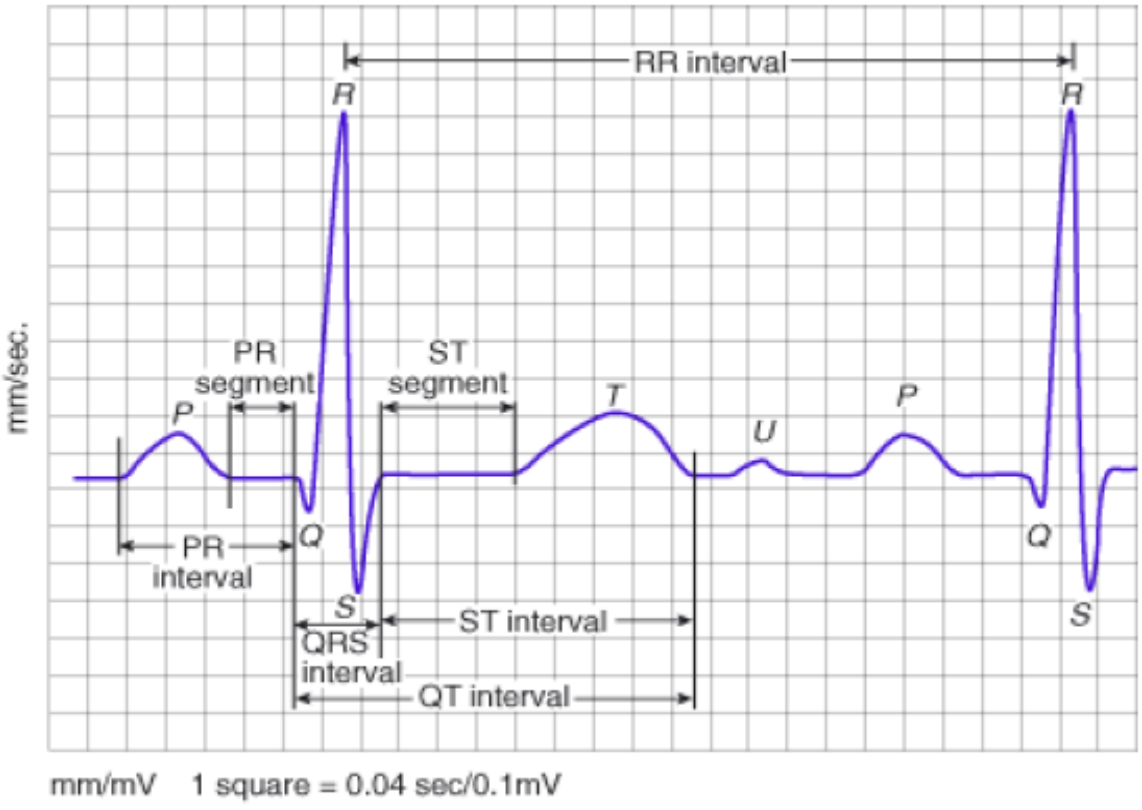

Figure 3.4 Schematic representation of normal ECG waveform.[2]

Waves Representation:

•**P wave:** the amplitude level of this voltage signal wave is low(approximately1 mV) and represent depolarization and contraction of the right and left atria. A clear P wave before the QRS complex represents sinusrhythm. Absence of P waves may suggest a trial fibrillation, junction rhythm or ventricular rhythm. It is very difficult to analyze P waves with a high signal-to-noise ratio in ECG signal.

•**QRS complex:** The QRS complex is the largest voltage deflection of approximately 10-20 mV but may vary in size depending on age, and gender.

The voltage amplitude of QRS complex may also give information about the cardiac disease . Duration of the QRS complex indicates the time for the ventricles to depolarize and may give information about conduction problems in the ventricles such as bundle branch block.

•**T wave:** Represents ventricular repolarization. Large T waves may represent ischemia, and Hyperkalaemia.

Table 3-2 Amplitude and duration of waves, intervals and segments of ECG signal.[2]

No	Features	Amplitude(mA)	Duration (ms)
1	1 P wave	0.1-0.2	60-80
2	PR-segment	-	50-120
3	PR-interval	-	120-200
4	QRS complex	1	80-120
5	ST-segment	-	100-120
6	T-wave	0.1-0.3	120-160
7	ST-interval	-	320
8	RR-interval	-	(0.4-1.2)s

The Table3.3 shows features of P-wave, QRS complex and T wave in maximum amplitude and its duration. According to medical definition, the duration of each RR-interval is about (0.4-1.2)s[1].

3.3.4 the noise in ECG signal

Generally the recorded ECG signal is often contaminated by different types of noises and artifacts that can be within the frequency band of ECG signal, which may change the characteristics of ECG signal. Hence it is difficult to extract useful information of the signal .

Electrocardiogram traces used for identification are obtained using surface electromyography (EMG), where electrodes are placed on the skin in the vicinity of the heart. Potential differences of 1 to 3 mV generated at the body surface by the current sources in the heart are picked up by the electrodes and are amplified in order to improve the signal to noise ratio (SNR). The ECG waveform is observed on an oscilloscope or is digitized for further processing by a computer (as will be the case for recognition purposes). The digitization process should use a sampling rate of at least 1 kHz to ensure that the ECG trace is of a high enough resolution as required for biometric purposes [19].

ECG measurements may be corrupted by many sorts of noise. The ones of primary interest are:

1. power line interference.
2. electrode contact noise.
3. motion artifacts.
4. EMG noise.
5. Instrumentation noise.

An idealized block diagram of each of these noise sources is shown in Figure3.5. The various noise signals presented in the figure will be characterized in greater detail in this section.

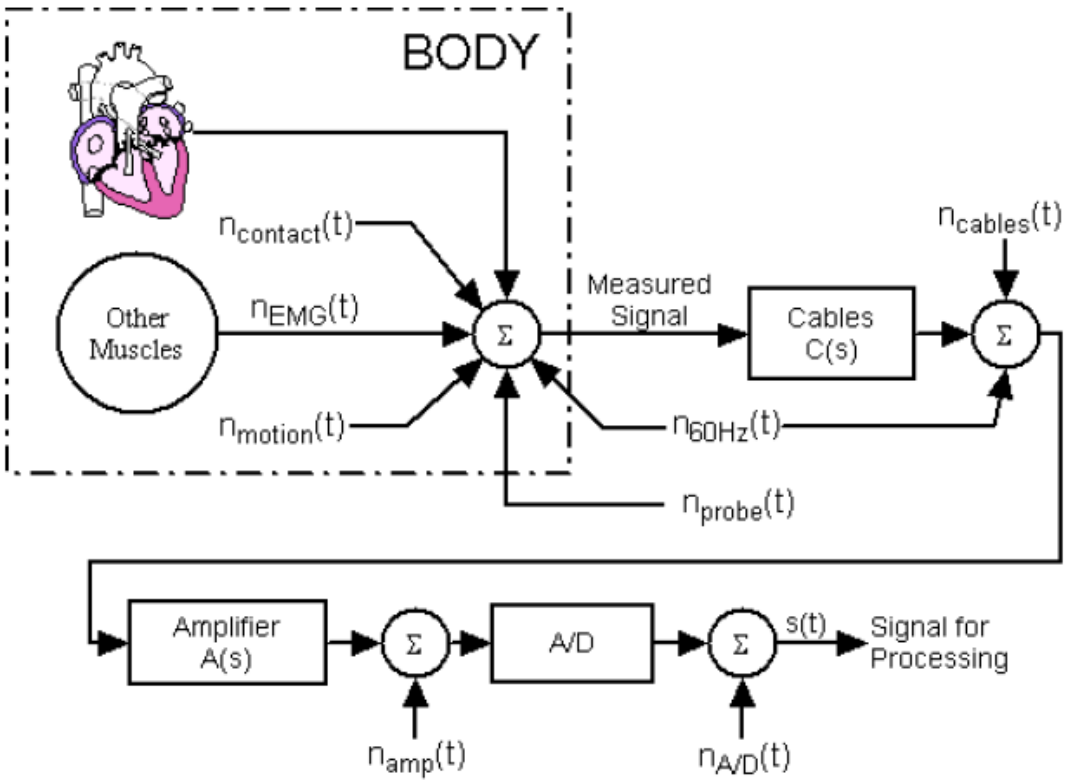

Figure3.5 block diagram showing the principal noise sources in electro cardiology.[19]

3.3.4.1 Power Line Interference

Power line interferences contains 60 Hz pickup (in U.S.) or 50 Hz pickup (in India) because of improper grounding [20]. It is indicated as an impulse or spike at 60 Hz/50 Hz harmonics, and will appear as additional spikes at integral multiples of the fundamental frequency. Its frequency content is 60 Hz/50 Hz and its

harmonics, amplitude is up to 50 percent of peak-to-peak ECG signal amplitude [21]. A 60 Hz notch filter can be used remove the power line interferences.

Plotting a Fourier power spectrum of a typical ECG signal figure2.6 reveals various common ECG frequency components. Several interesting features are readily identifiable:

•The 1.2 Hz heart beat information (approximately 72 beats per minute)

•The 60 Hz power line interference

The remainder of the frequency components represents the subject information (situated between 0.1 Hz and 40 Hz) and contributions of other noise sources.

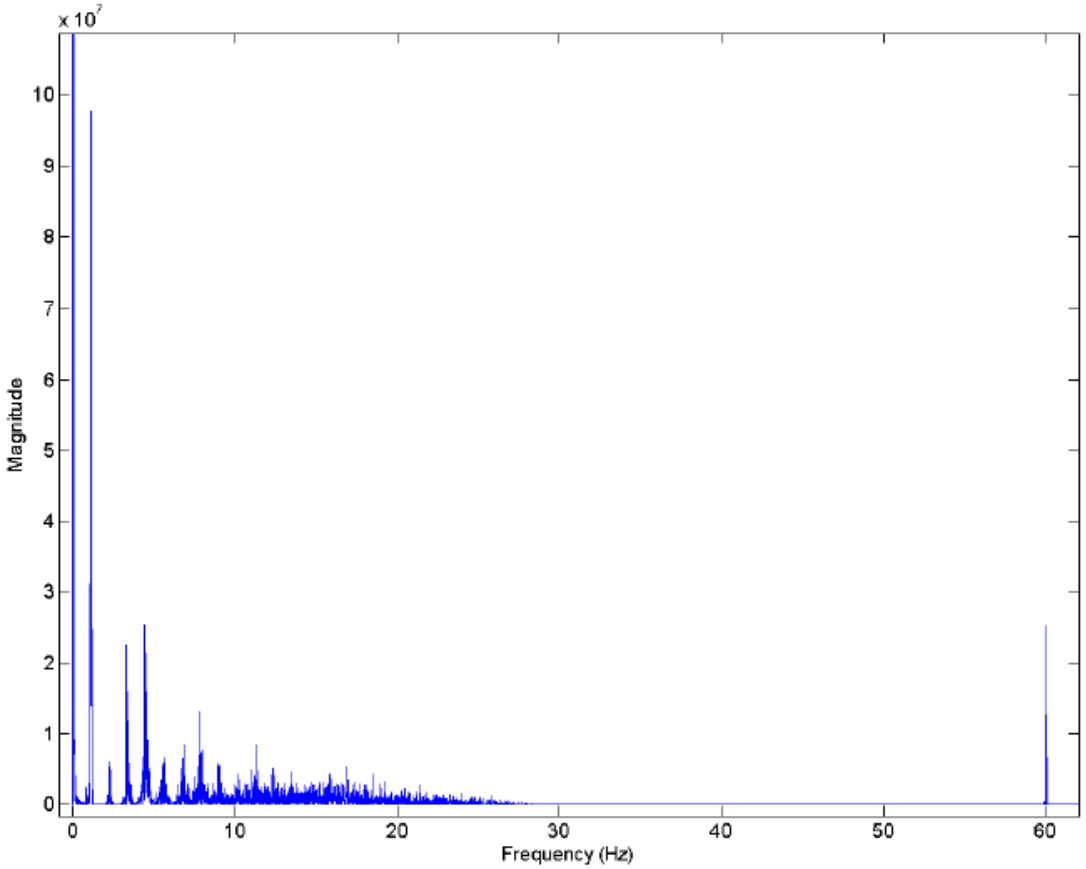

Figure 3.6 Fourier power spectrum of an ECG trace. The 60 Hz power line interference and the baseline potential drift noise (at approximately 0 Hz) are identifiable.[21]

Power line interferences contain 60 Hz pickup because of improper grounding. It is indicated as an impulse or spike at 60 Hz/50 Hz harmonics, and will appear as additional spikes at integral multiples of the fundamental frequency. Its frequency content is 60 Hz/50 Hz and its harmonics, amplitude is up to 50 percent of peak-to-peak ECG signal amplitude. A 60 Hz notch filter can be used remove the power line interferences.

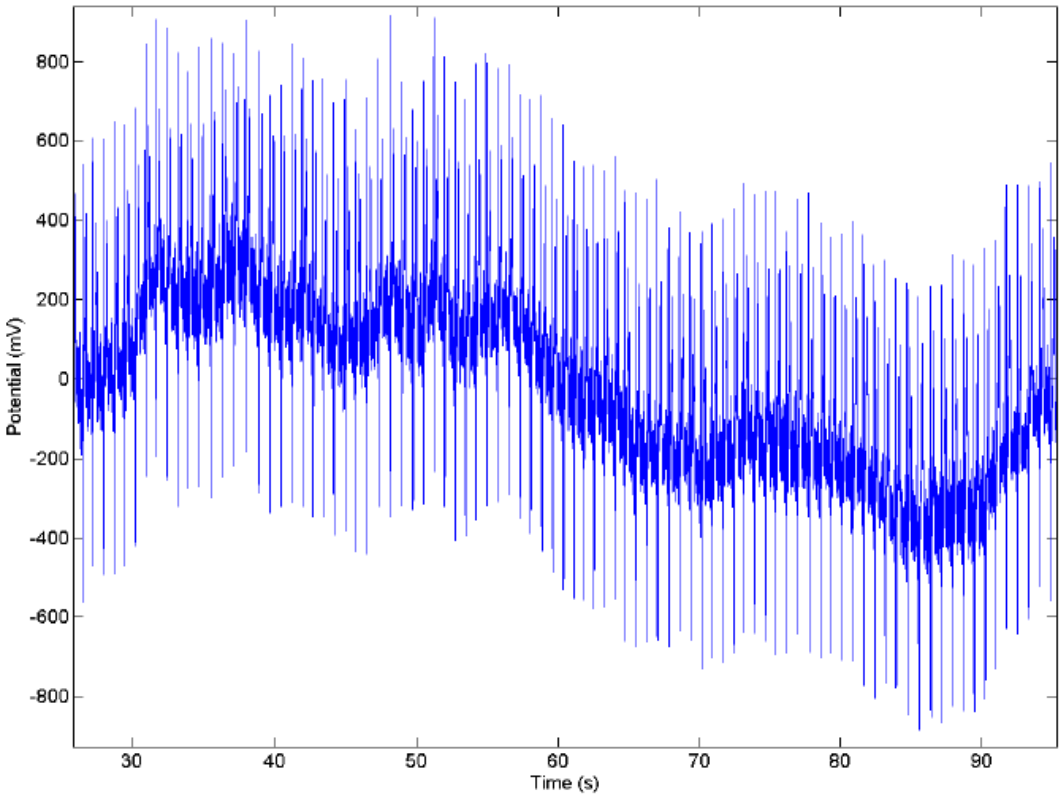

Figure3.7 Seventy seconds of ECG data. The x-axis is time in seconds, and y-axis is the electrical potential in mill volts. A baseline potential drift is present in the ECG trace [21].

3.3.4.2 Electrode Contact Noise and Motion Artifacts

Electrode contact noise is caused by variations in the position of the heart with respect to the electrodes and changes in the propagation medium between the heart and the electrodes. This causes sudden changes in the amplitude of the ECG signal,

as well as low frequency baseline shifts. In addition, poor conductivity between the electrodes and the skin both reduces the amplitude of the ECG signal and increases the probability of disturbances (by reducing SNR)[22]. The underlying mechanism resulting in these baseline disturbances is electrode-skin impedance variation. The larger the electrode-skin impedance, the smaller the relative impedance change needed to cause a major shift in the baseline of the ECG signal.

If the skin impedance is extraordinarily high, it may be impossible to detect the signal features reliably in the presence of body movement. Sudden changes in the skin-electrode impedance induce sharp baseline transients which decay exponentially to the baseline value. This transition may occur only once or rapidly several times in succession. [23].

Motion artifacts are transient (but not step) baseline changes caused by electrode motion. The usual causes of motion artifacts are vibrations, movement, or respiration of the subject. The peak amplitude and duration of the artifact are random variables which depend on the variety of unknowns such as the electrode properties, electrolyte properties (if one is used between the electrode and skin),skin impedance, and the movement of the patient [24]. Fig. 3-7 shows a 70 second segment of a high resolution ECG trace, where the baseline drift varies from approximately -400mV to 400mV. In this ECG signal, the baseline drift occurs at an unusually low frequency (approximately 0.014Hz), and most likely results from very slow changes in the skin-electrode impedance. This noise can also be observed on the Fourier power spectrum ; the large peak nearest to DC is the result of very low frequency base line shifts. The noise artifacts introduced by subject motion are modeled by nmotion(t) in figure3.8.

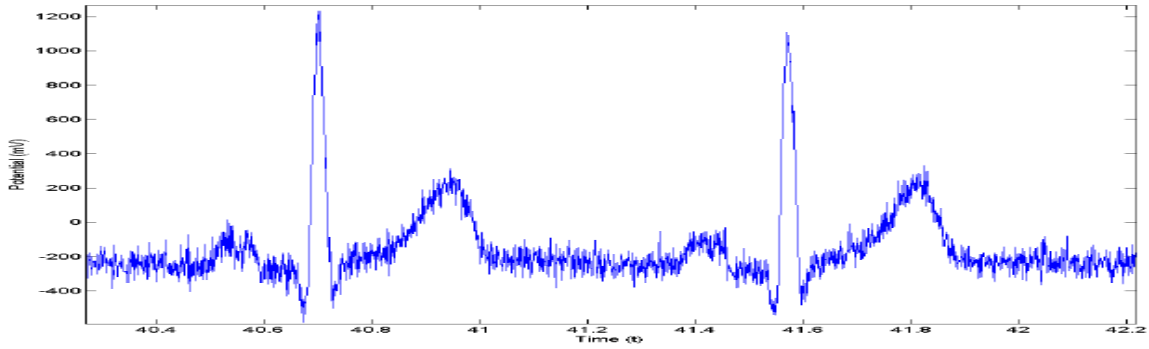

Figure3.8 Two second segment of an ECG trace. The x-axis is time in seconds, and the y-axis is the electrical potential in millivolts. Exact positions of P and T complexes are obscured be presence of EMG noise.[24]

3.3.4.3 EMG Noise

EMG noise is caused by the contraction of other muscles besides the heart. When other muscles in the vicinity of the electrodes contract, they generate depolarization and repolarization waves that can also be picked up by the ECG.

The extent of the crosstalk depends on the amount of muscular contraction (subject movement), and the quality of the probes.

It is well established that the amplitude of the EMG signal is stochastic (random) in nature and can be reasonably modeled by a Gaussian distribution function.[25] The mean of the noise can be assumed to be zero; however, the variance is dependent on the environmental variables and will change depending on the conditions. Certain studies have shown that the standard deviation of the noise is typically 10% of the peak-to-peak ECG amplitude [26]. While the actual statistical model is unknown, it should be noted that the electrical activity of muscles during periods of contraction can generate surface potentials comparable to those from the heart, and could completely drown out the desired signal. The effects of typical EMG noise can be observed in the ECG signal, and is particularly problematic in the areas of the P and T complexes.

3.3.4.4 Instrumentation Noise

The electrical equipment used in ECG measurements also contributes noise. The major sources of this form of noise are the electrode probes, cables, signal processor/amplifier, and the Analog-to-Digital converter, represented respectively by n probe(t), n cables(t), n amp(t), and n A/D(t) in Figure3.8. Since this form of noise is usually defined by a white Gaussian distribution, adequately represents its effects on the ECG signal. Unfortunately instrumentation noise cannot be eliminated as it is inherent in electronic components, but it can be reduced through higher quality equipment and careful circuit design.[27]

One type of electrical noise is resistor thermal noise (also known as Johnson noise). This noise is produced by the random fluctuations of the electrons due to thermal agitation. The power spectrum of this noise is given by

$$Vn^{-2} = 4KTR, \qquad (1\text{-}3)$$

where k is the Boltzmann's constant, T is the temperature, and R is the resistance [28]. This type of noise is generated in the electrodes, in the wire leads connecting electrodes to the amplifier, and in all the resistive electronic components internal to the ECG instrumentation. Since the magnitude of this noise component is substantial relative to the measured signal, its effects are most noticeable in the electrodes and any other electronic equipment prior to the amplifier.

Another form of noise, called flicker noise, is very important in ECG measurements, due to the low frequency content of ECG data. The actual mechanism that causes this type of noise is not yet understood, but one widely accepted theory is that it is caused by the energy traps which occur between the interfaces of two materials. It is believed that the charge carriers get randomly

trapped/released and cause flicker noise. The power spectral density of flicker noise is given by:

$$\bar{V}\ 1/f = 4KT/WLCoxF, \qquad (2\text{-}3)$$

where k is the Boltzmann's constant, T is the temperature, Cox is the silicon oxide capacitance, WL is the transistor area, and f is the frequency [29]. As the equation suggests, flicker noise is inversely proportional to frequency, indicating that it becomes dominant at lower frequencies. It can be found in any electronic equipment which utilizes bipolar or metal oxide transistors, such as the amplifier used for signal amplification (or more specifically any device which has material junctions).

3-4 Arrhythmias in ECG signal

The normal rhythm of the heart where there is no disease or disorder in the morphology of ECG signal is called Normal sinus rhythm (NSR). The heart rate of NSR is generally characterized by 60 to 100 beats per minute. The regularity of the R-R interval varies slightly with the breathing cycle.

When the heart rate increases above 100 beats per minute, the rhythm is known as sinus tachycardia. This is not an arrhythmia but a normal response of the heart which demand for higher blood circulation [1]. If the heart rate is too slow then this is known as bradycardia and this can adversely affect vital organs. When the heart rate is too fast, the ventricles are not completely filled before contraction for which pumping efficiency drops, adversely affecting perfusion.

3-5 Artificial Neural Networks (ANNs)

3-5-1 General Overview of Artificial Neural Networks

Work on artificial neural networks, commonly referred to as "neural networks "has been motivated right from its inception by the recognition that human brain computes in an entirely different way from the conventional digital computer. An

artificial neural network, which is the formal name for the term neural networks used here, is one of many attempts to build an intelligent machine or to create artificial intelligence. It is based on biological neural networks. The basic idea to model this is to make a very simplified model of biological neurons and their synapses.[30]

The novelty of Artificial Neural Network (ANN) theory could undoubtedly be attributed to the experiment performed by “McCulloh” and “Pitts” in modeling bio-systems using nets of simple logical operations back in 1943. The idea was to find a simple parametric nonlinear model for a real neuron. Ever since this innovation, there has been a great interest from various researchers and scientists, thus several ANN-based models in different fields were discovered. The technology though had lost its momentum in the late 1969 till 1986 when the back-propagation of error was discovered. To date, ANN-based models have been successfully implemented in a number of industries ranging from: Aerospace, automotive, defense, electronics, entertainment, financial and so on. Artificial neural Networks have been also successfully applied in medical fields[31].

ANN is a part and parcel of intelligent based systems, designed distinctively to improve the performance of conventional computing techniques. The biggest drawback associated with the so called conventional methods is the inability to learn and identify patterns in dynamic systems. Thus the need to eliminate this shortcoming through learning is proven essential.

Artificial Neural Network is an information processing paradigm inspired by the way biological nervous systems, such as the brain, process information.[32] The human brain has 100 billion biological neurons with about 100 000 connections per neuron. A simplified biological neuron is illustrated in Figure 3.9

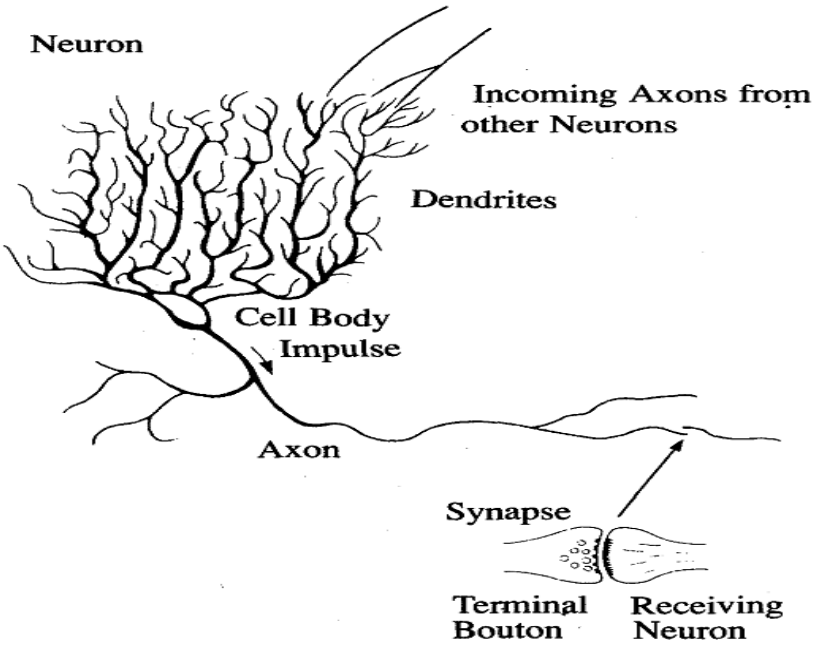

Fig. 3.9: A biological neuron[32]

Biological neurons receive spikes through synapses located on the dendrites of the neuron. When the spikes received are strong enough and exceed a definite threshold, the neuron is activated and fires a signal though the axon. This signal travels from the body, down the axon, to the next neuron(s)[33]. Learning arises by adjusting the effectiveness of the input (synapses) so that it influences one neuron on other changes. Humans and highly trained animals use the same configuration and summing up to extremely complex networks.

Similarly, an artificial network is made up of simple interconnected processing elements called neurons. The neurons are arranged in a layered structure to complete a network competent of executing parallel and distributed computations. Architecture of a simple ANN is shown in Figure 3.10. The attraction of ANN-based models comes with the network's ability to learn, recognize data patterns, and adapt to a changing environment like the human brain. This adaptive characteristic is often called "the human-like reasoning.

The architecture illustrated in Figure 3.10, presents a three layered feed-forward network. ANN has a remarkable capability to develop sense from convoluted or imprecise data, extract patterns and detect trends that are too complex often only noticeable by either humans or other computer techniques. [34]

In broad terms, ANN-based models offer a variety of benefits namely: adaptive learning, self organization, real time operation, fault tolerance via redundant information coding. Thus neural network processes information in the similar way the human brain does.[35]

The neurons are organized in a way that defines the network structure. The most concerned structure is the multilayer perceptron (MLP) type, in which the neurons are organized in layers. The neurons in each layer may share the same inputs, but are not connected to each other. If the architecture is feed-forward, the outputs of one layer are used as the inputs to the following layer. The layers between the input neurons and the output layer are called the hidden layer.

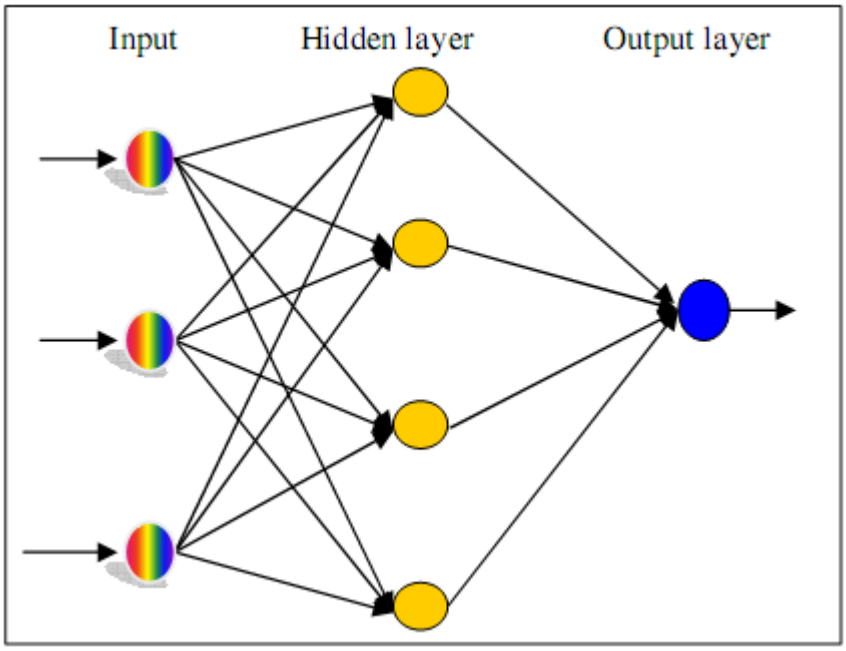

Fig. 3.10: Architecture structure of a feed forward neural network .[35]

3-5-2 Basic Architecture of a Feed-forward Network

The feed-forward network topology illustrated in Figure 3.10 permits signals to travel one way only, from the input through the hidden layer to the output layer. These types of networks are somehow straight forward and associate inputs with outputs. They are extensively used in pattern recognition. This kind of organization is also referred to as bottom-up or top-down and commonly used in pattern recognition. Figure 3.10 also shows the commonest type of artificial neural network which consists of two layers. The hidden layer neurons are connected to the output layer neurons. The functions of each layer in the network are defined below:

a) The input layer neurons represent the pre-processed data fed into the network.

b) The input of each hidden layer neuron is defined by the sum of the input vector set and the connection weights between the input layer and hidden layer.

c) The input of the output neuron is determined by the weighted sum of outputs of the hidden layer neurons.

d) The output of a neuron is defined by the type of the transfer function used in that specific layer.

This type of network is attractive because the hidden neurons are free to develop their individual representations from the input set.[36]

3-5-3 ANN Training and Generalization

There are various processes involved in developing a supervised ANN-based model. Amongst others, training process and validation process are some of the vital steps. The input-output patterns are repeatedly presented to the network during the training process. Through this process, the network learns the subjected environment or patterns, and eventually yields the desired output. The desired output is attained as a result of adjusting weights of all interconnections between neurons to establish the correct set of input-output response. However, during the

training process, optimal training time is required to avoid overtraining. Overtraining of the network can be prevented by employing complex stopping criterion. Early stopping (the most common), regularization, pruning, Information Criterion Pruning (ICP), Cross-Validation Pruning (CVP) are some of the stopping methods.

Training the network at infinitum normally results in a reduced error function for a given set of inputs. Though, this does not guarantee better accuracy and robustness of the network because the error could be extremely big if the network is presented with the data it has not seen before. This characteristic is explored during the validation process. Moreover, the validation process also improves the network reliability and generalization.

3-5-4 Network Coupled Errors

Broadly speaking, there exist two types of errors during the training process: training error and generalization error. The general error presentation can be seen in Figure 3.11. At the beginning of the training, the training error is at maximum, and gradually decreases as time elapses.[30] Optimum training time should be within the circled area in order to maintain network generalization tendency. In essence, the desired goal should not be an error-free output, as the network will tend to memorize the input-output patterns if trained at infinitum.

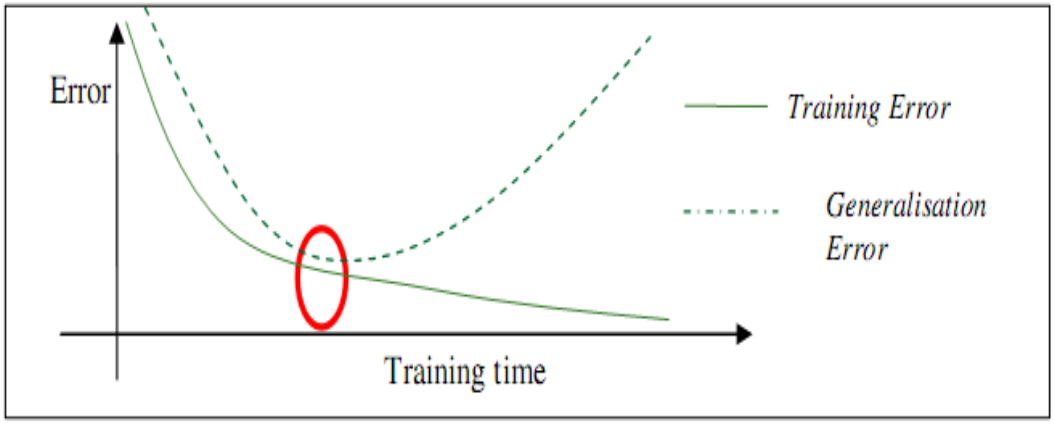

Fig. 3.11: Errors vs. Optimal network training time.[30]

Chapter 4

Methodology

This chapter introduces the research methodology which is performed using MATLAB to classify ECG signals as normal or abnormal signals. These steps represent only one part of the whole classification system. The program will provide the user with a framework for ECG classification. In this project the ECG signals are collected from Physionet bank (MIT-BIH Arrhythmia Database).

The proposed methodology is depicted in the block diagram in Figure 3.1. Initially ECG Signals are preprocessed for removal of power line noise and high frequency interference. Then deflections in the ECG Signal Q, R, S and statistical features are identified ,which is a very important feature in identifying arrhythmias. A neural network is trained with 1500 dataset containing features of QRS complex, intervals and the statistical features .

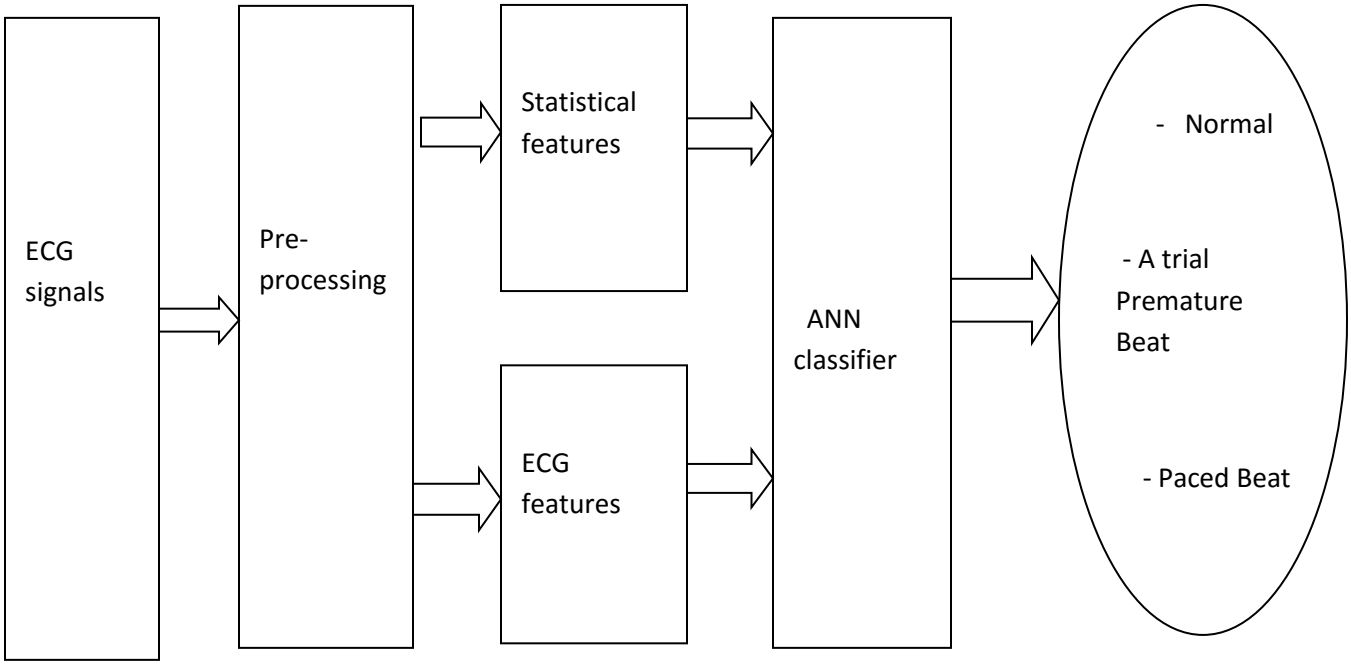

Figure 4-1 : block diagram of proposed method for ECG classification

4.1 Preprocessing Module

The main aim of ECG signal preprocessing is to prepare a compact description of the QRS complex, ECG intervals and segment and heart rate, for input to the classification device (Artificial Neural Network) with minimum loss of information.

4.1.1 Used signals

The MIT-BIH arrhythmia database contains number of normal and abnormal ECG signal collected from people with different condition from different gender and ages. This database consists of 48 ECG recordings each recording is 30:06min long. In this study, we used 10 records from this database, each record of the MIT-BIH database is 30:06 minutes selected from 24 hours. The sampling frequency of the ECG signals in this database is360Hz; each beat is described by a label called an annotation. Typically an annotation files for an MIT-BIH record contains about 2000 beat annotations, and smaller numbers of rhythm and signal quality annotation. In our design we selected 3 records we have divided the signal into 500 samples from each record thus we have 1500 samples ready to be as input for the system. In below figures that show the three types of signals they are : normal signal, a trial premature beat signal (APB), and paced beat signal (PB) :

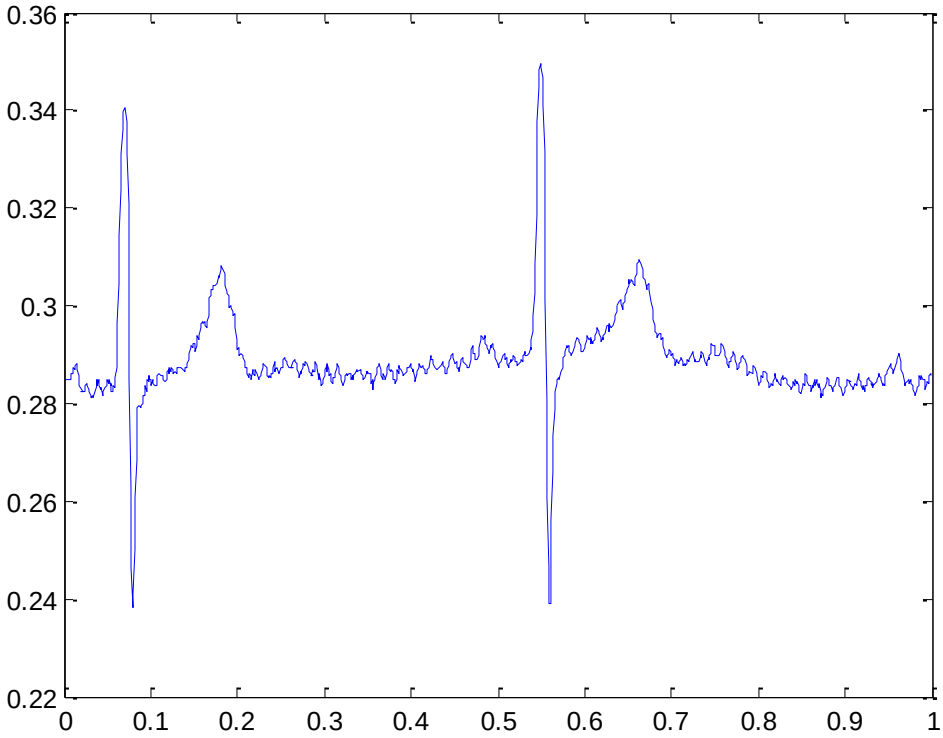

Figure 4-2 : normal signal

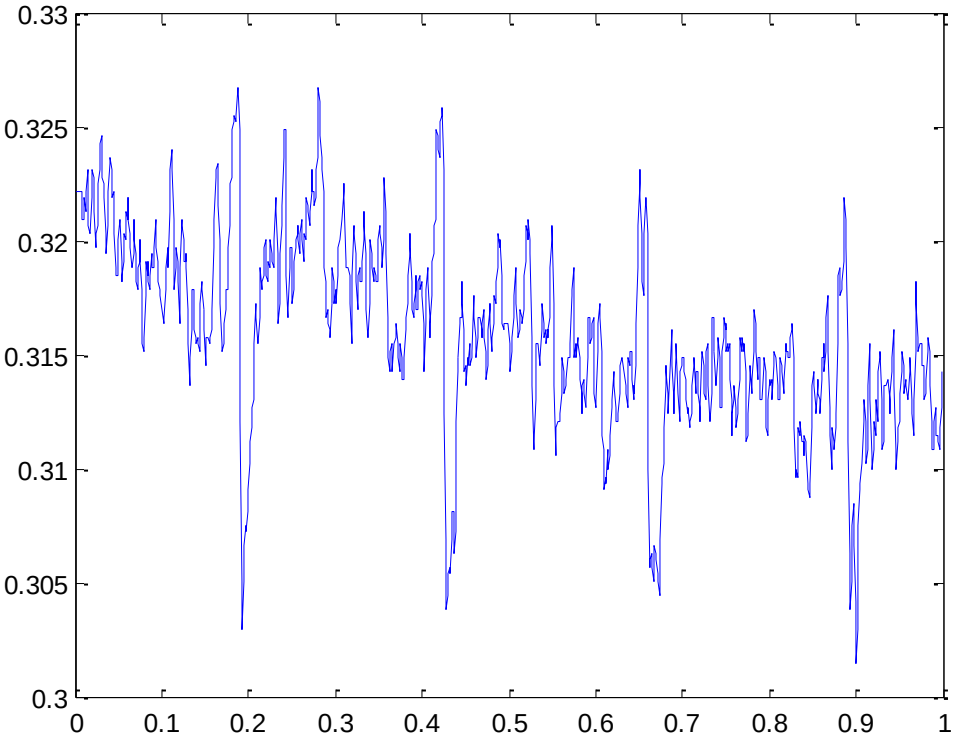

Figure 4-3 : a trial premature beat signal (APB)

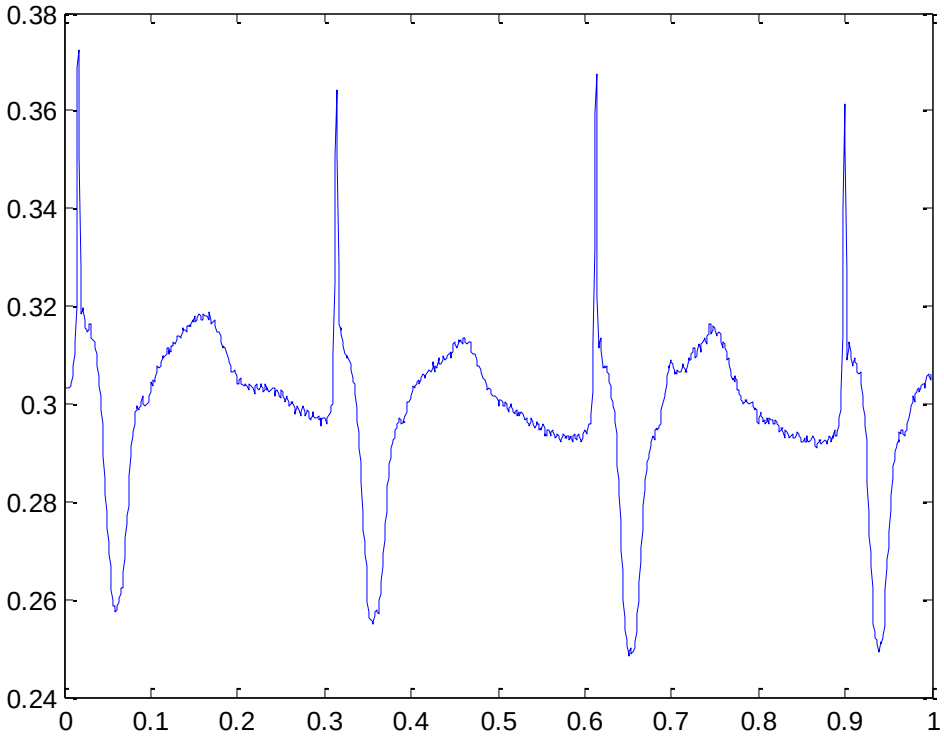

Figure 4-4: paced beat signal (PB)

4.1.2Band path filter

In this section we cover the first preprocessing stage of our ECG signals, this stage includes denoising, smoothing, and normalization. To perform the following operations, we have used different tools on MATLAB. Denoising was performed using the band path filter. Since the extracted signals are noisy in many cases, denoising was an essential step before further processing of the signals. After searching for and trying several methods, we chose to perform our de- noising using the band path filter, that provides the necessary function for de- noising signals that convey properties similar to our ECG signals, filter of pass band frequencies of is used to remove the power-line interference and high

frequency noises from the original signal, this filter can be created by combining a law-pass filter with a high-pass filter. Band pass is an adjective that describes a type of filter or filtering process; it is to be distinguished from pass band. which refers to the actual portion of affected spectrum.

4.2 Feature Extraction Module

QRS Complex Identification & Feature Extraction

The ECG signals from database were preprocessed for removal of power line noise and high frequency interference. Deflection Identification is then applied to the data thus obtained. Deflection indices found works as input to feature extraction and then neural network is applied to train the system.

4.2.1 R Peak Detection

The first stage is the extraction of suitable metrics form the signal of interest.

Before these can be extracted from the ECG signal, the Q, R, S deflections in each beat were identified. This is performed with an algorithmic script with the following methodology: The first goal is the detection of the R Peak because once the R-Peak is detected; it can be used to detect the Q and S points easily. Due to the idiosyncratic nature of the QRS complex & the distinctive characteristics of the R peak, this is readily identified even in the most distorted ECG readings. Thus it is used as the basis for ECG feature identification.

The QRS complex is the most striking waveform within the electrocardiogram. Since it reflects the electrical activity within the heart during the ventricular contraction, the time of its occurrence as well as its shape provide mucli information about the current state of the heart. Due to its characteristic shape it serves as the basis for the automated determination of the heart rate, as an entry point for classification schemes of the cardiac cycle. In that sense, QRS detection provides the fundamentals for almost all automated ECG analysis algorithms.

Typical frequency components of a QRS complex range from about 14 to about 25Hz. Therefore, almost all QRS detection algorithms use a filter stage prior to the actual detection in order to attenuate other signal components and artifacts, such as P wave, T wave, baseline drift, and in coupling noise. In this thesis a fixed interval thresholding is used to segment the signal into three intervals and the maximum value of each interval was identified as the R peak of this time interval. The values if R peak can be tested by using the equation:

$$R\ peak=max(ecg)*0.6 \quad (1\text{-}4)$$

Where (ecg) is the signal itself. All the point equal and above the R peak value are consider when finding the peak as the peak itself or a point near it. The fixed thresholding intervals vary from 1:350, 350:600 and 600:850. It was found after testing that most of the signal can be segmented in this way.

4.2.2 P, Q, S and T detection:

After the R peak detection the other key points are easily identified by using the suitable interval thresholding. For the Q detection we used the equations:

$$range=[Rpos-20:Rpos]; \quad (2\text{-}4)$$

Where R pos is the position of R peak. After determining the range where Q should exist it can be easily defined as the lowest point there. In the same way the P, S and T is determined.

P point range:

$$range=[Qpos1-50:Qpos1]; \quad (3\text{-}4)$$

S point range:

$$range=[Rpos1:Rpos1+45]; \quad (4\text{-}4)$$

T point range:

$$Tmax1=max(clean(rang \quad (5\text{-}4)$$

Once all the points are determined in position and value other features could be acquired easily such as: RR interval, QRS duration, QT segment and so on.

4-2-3 statistical features

The cardiologist depends only on the wave form and the morphology of the signal in order to give his diagnosis, while in fact the ECG signal carries more hidden information and characteristic these characteristics can be detected by extracting the statistical features. We used The statistical features were extracted represented in "**mean, median, standard deviation, ketosis, skewness**"

4-3 feature selection and reduction

By applying the ttest2 function to all features , the features that effective on differentiation between the normal and abnormal signals are eight features, they are: (Ppos1, Qpos1, Rpos1, Tpos1, mean, median, ketosis and skewness), And the effective features that differentiation between a trial premature beat (APB) and paced beat (PB) signals are six features, they are : (Qpos1, Rpos1, Spos1, mean, median and standard deviation).

4.4 Classifier Module

4.4.1 ARTIFICIAL NEURAL NETWORKS (ANN)

In modern software implementations of artificial neural network, the approach inspired by biology has more or less been abandoned for a more practical approach based on statistics and signal processing. In some of these systems neural networks, or parts of neural networks (such as artificial neurons) are used as components in larger systems that combine both adaptive and non-adaptive elements. While the more general approach of such adaptive systems is more suitable for real-world problem solving, it has far less to do with the traditional artificial intelligence connectionist models. [17] What they do, however, have in common is the principle of non-linear, distributed, parallel and local processing and adaptation.

Artificial neural networks have been widely applied in nonlinear signal processing, classification, and optimization. In many applications their performance was shown to be superior to classical linear approaches. The current medical knowledge does not allow any reliable definition of rules to help the traditional automatic system. The self-learning property of neural network allows to recover the feature, that implicitly characterize the signal. The large variety of topological architectures and learning paradigma makes hard the choice for each problem.

In this thesis artificial neural network tool where used by using the command(nprtool). After preprocessing all of our samples and then extract their feature. All the data collected from this step is presented in two excel sheets. The first represent the target of the ANN: 1 as the normal ECG signals and 0 as the abnormal ones. The second sheet contains 24 ECG features that were used to classify each signal a normal or abnormal. The features hypothesis were first tested and then used in the construction of the network. The 1500 signals were divided into:70% training signals,15% for testing and 15% as validation. [17] Training signals are presented to the network during training and the network is adjusted according to its error, validation samples used to measure network generalization and to halt training when network generalization stops and testing samples have no effect in training and so provide an independent measure during and after training stops. 20 hidden layer neurons are used in the end training process were performed and results were collected. The following widows represent the ANN steps for creating ECG classification tool:

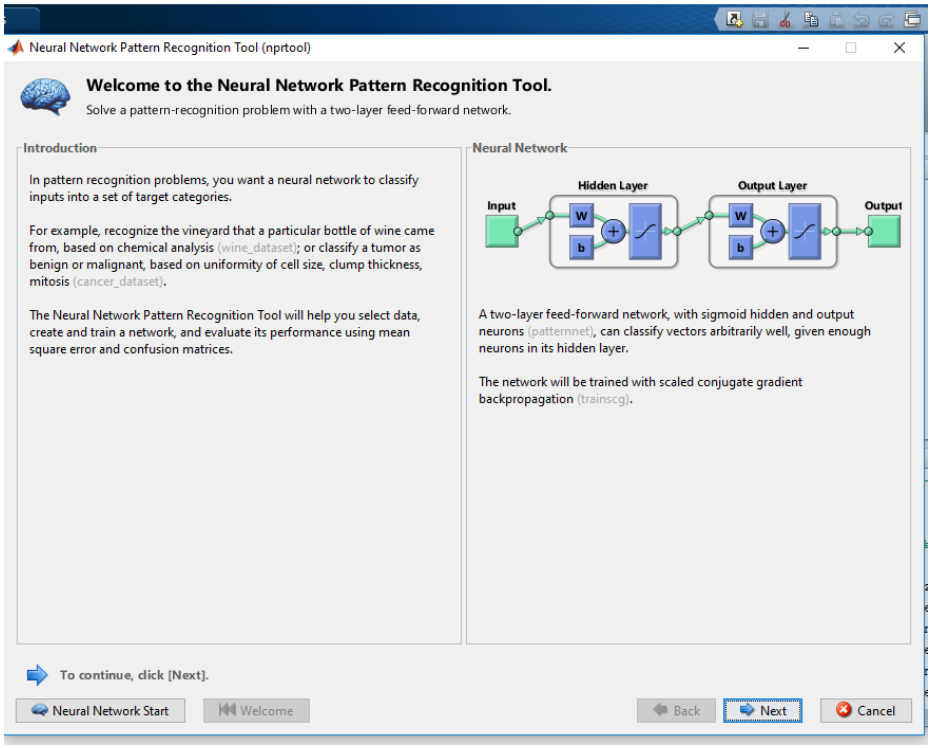

Figure 4.5 : The main window of the tool which is open using (nprtool) command

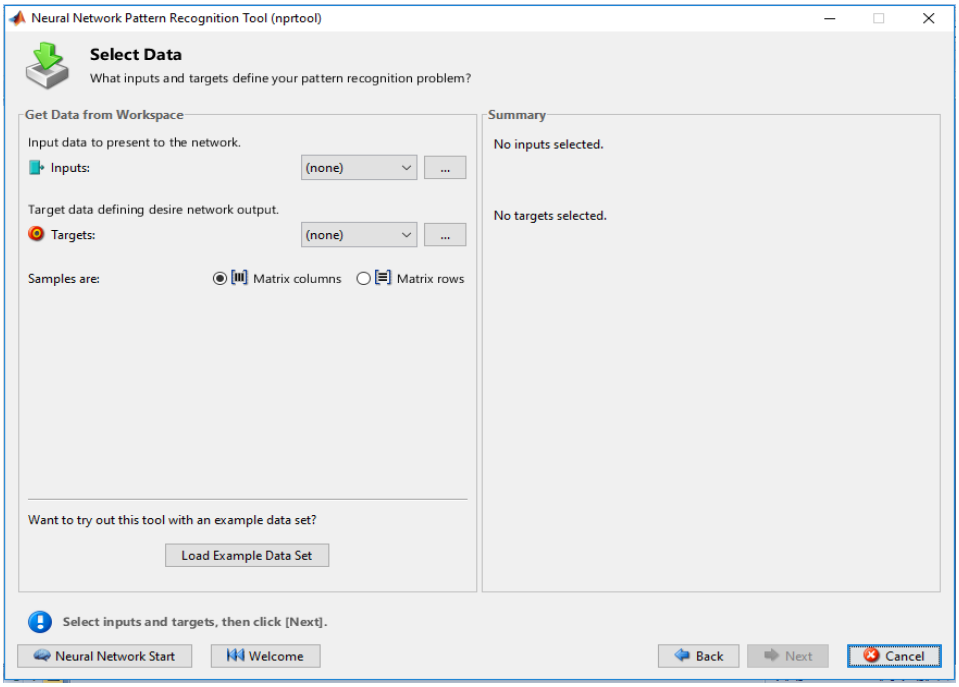

Figure 4.6: widow shows where to input the feature elements and the target for classifications

4.5 Output module

In order to evaluate the classifiers performances, we compute the percentages of Sensitivity, Specificity, Precision and classification rate; defined by the following equations:

The sensitivity (SE): $SE=TP/(TP+FN)$ (6-4)

The precision (PR): $PR=TP/(TP+FP)$ (7-4)

The specificity (SP): $SP=TN/(TN+FP)$ (8-4)

Classification percentage (CC):

$$TN+TP/(TN+TP+FN+FP) \quad (9\text{-}4)$$

Where: FP = False Positives; FN = False Negatives; TP = True Positives; TN = True Negatives; and The actual values are presented in the result and discussion chapter.

Chapter 5
Results and Discussion

This section describes and discusses the results that obtained from applying of all the steps.

5.1 ECG Data

The ECG data used in this study were obtained from Arrhythmia Database, PHYSIONET website. In this website we found short term (00:10 sec) ECG signals for healthy and sick individuals. Information about each individual health, gender, age and medications used were provided. Also information about the ECG signal itself was found. The MIT-BIH arrhythmia database contains number of normal and abnormal ECG signal collected from people with different condition from different gender and ages. This database consists of 48 ECG recordings each recording is 30:06min long. In this study, we used 10 records from this database, each record of the MIT-BIH database is 30:06 minutes selected from 24 hours. The sampling frequency of the ECG signals in this database is360Hz; each beat is described by a label called an annotation. Typically an annotation files for an MIT BIH record contains about 2000 beat annotations, and smaller numbers of rhythm and signal quality annotation. In our design we selected 3 records we have divided the signal into 500 samples from each record thus we have 1500 samples ready to be as input for the system.

5.2 Main Windows

After starting the training progress the Neural Network Training window pop up to show the training in progress which stopped once the validation checks reached 6, as shown in Figure 5.1.

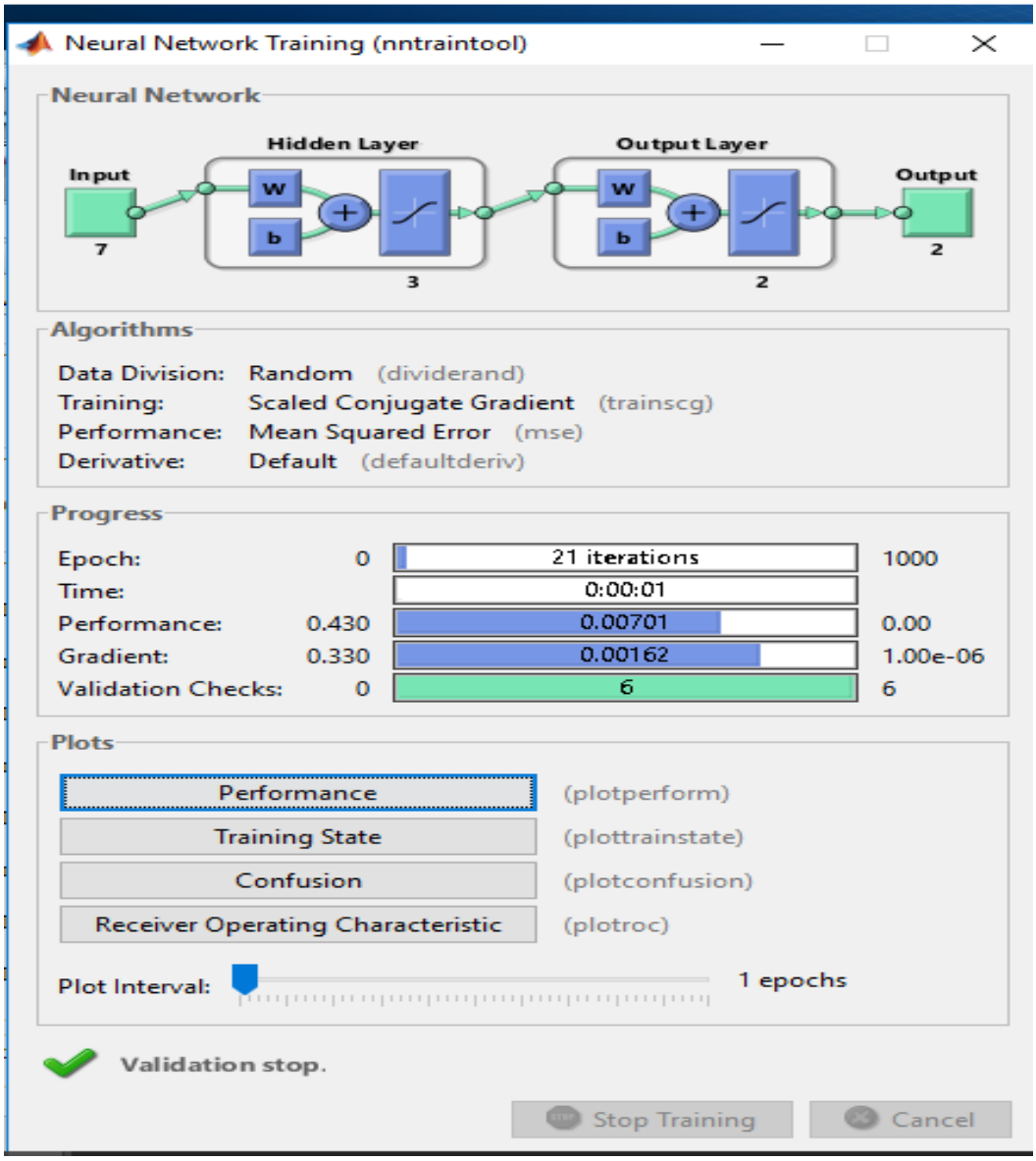

Figure 5.1: Neural Network Training Window (nntaintool)

5-3 Result between the Normal to abnormal

Figure 5-2confusion window

The confusion matrix for training, testing and validation samples and the three data combined is shown in the above Figure5.2. The network output is very accurate, is seen by the high number of correct responses in the green square and the low number of incorrect responses in the red squares. The lower right blue squares illustrate the overall accuracies. Due to low accuracy value resulted in the first

training process, a retraining is performed again and results are obtained as shown in figure5.3 and 5.4 below

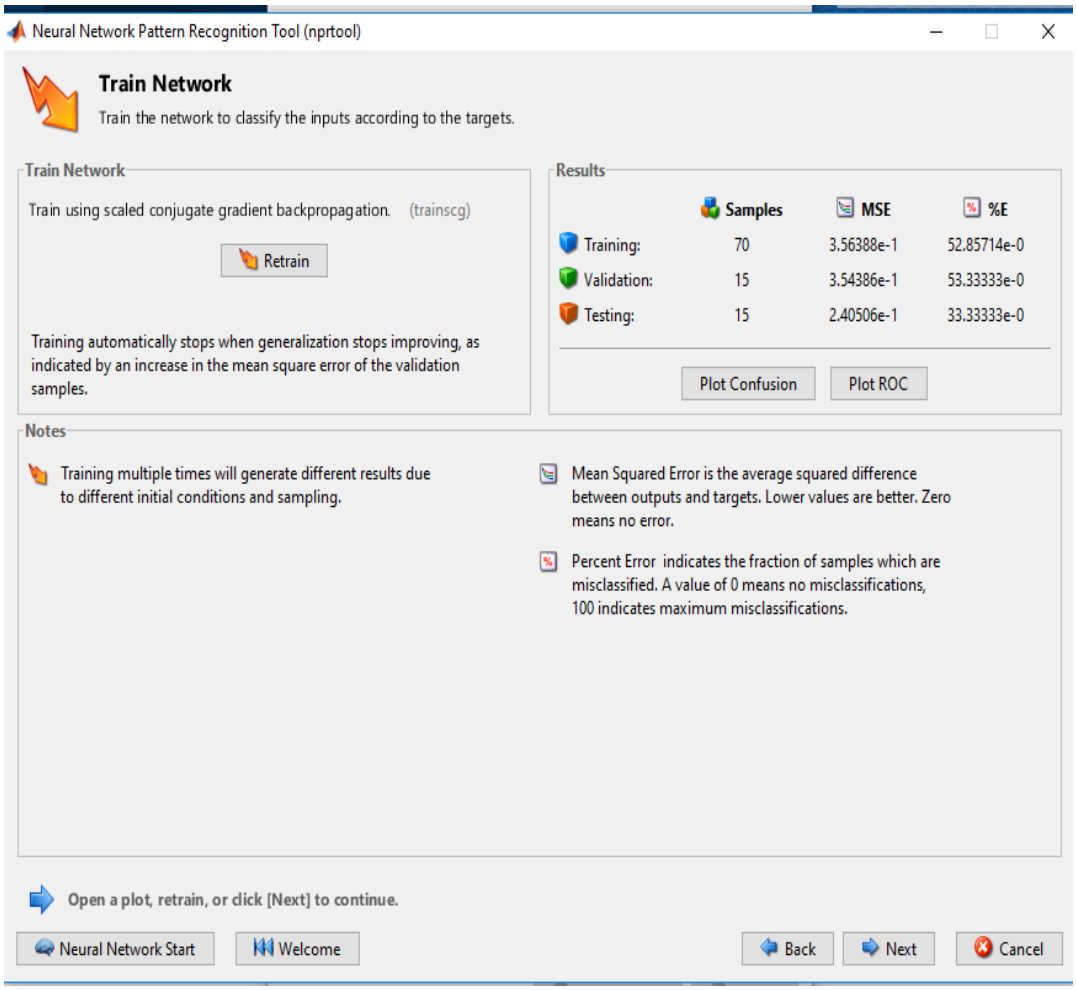

Figure 5.3: retraining window

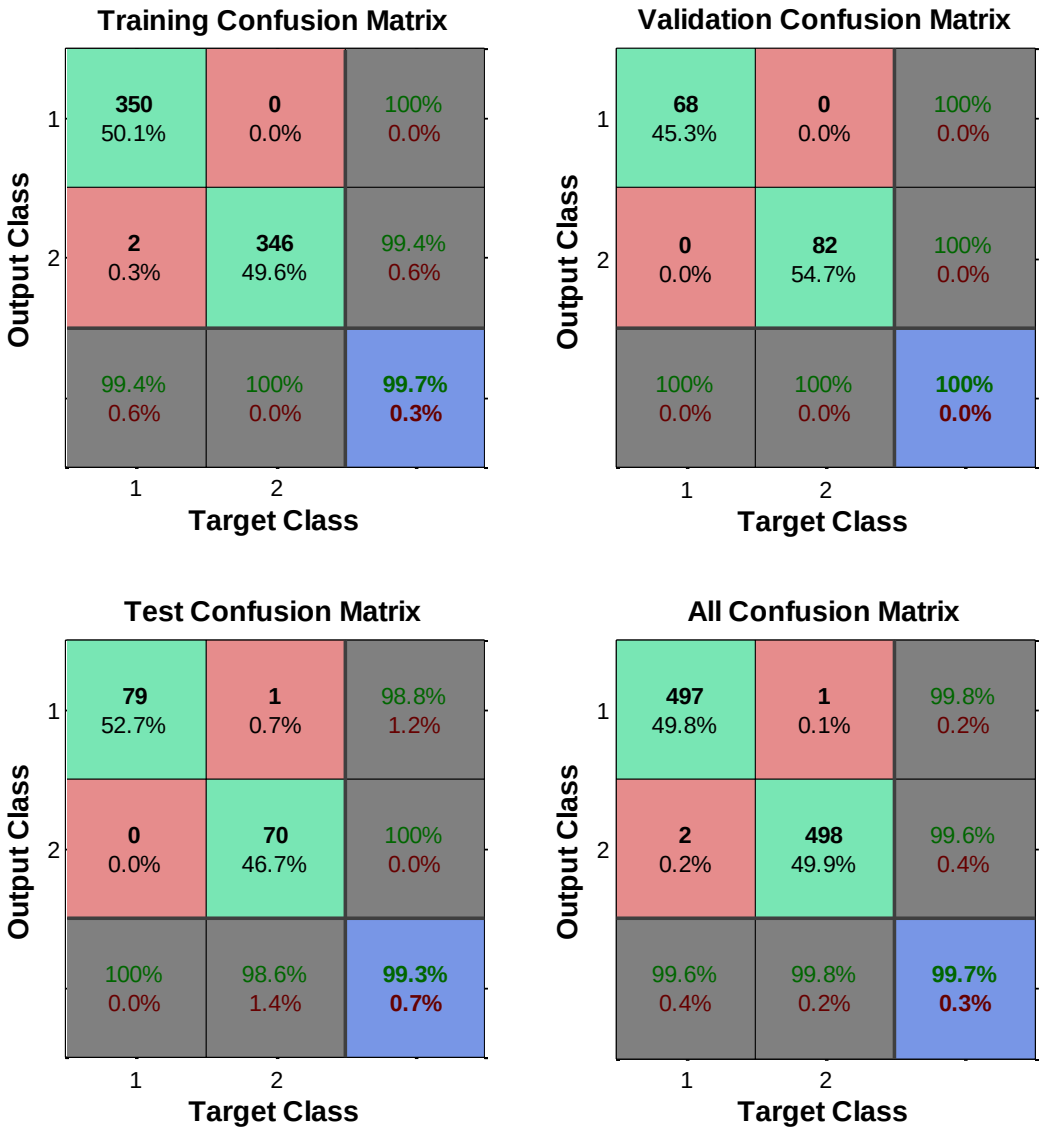

Figure 5-4 : second confusion window

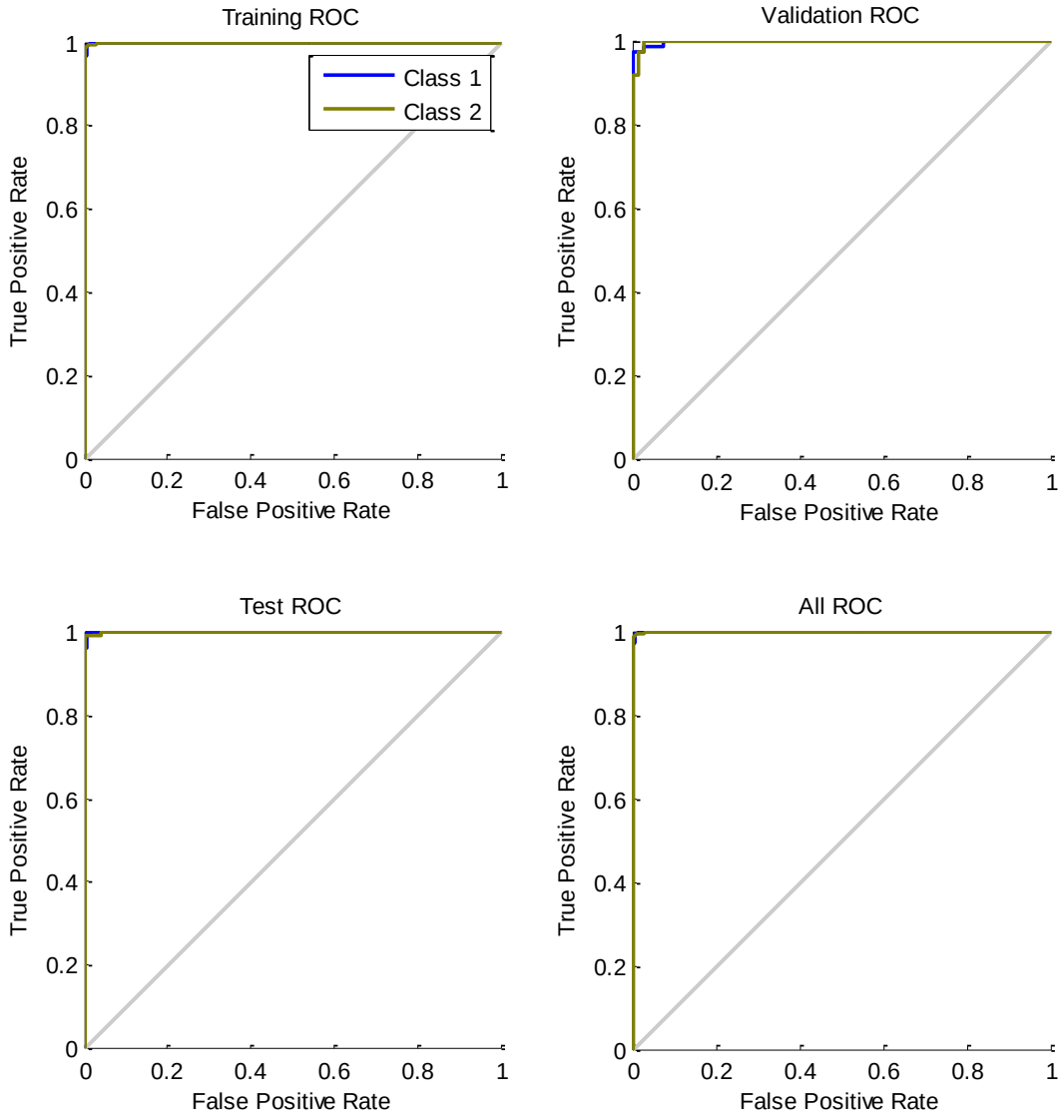

Figure5.5 Receiver operating characteristics(plotroc)(ROC)

In the above figure the colored line in each axis represent the ROC curves. The ROC curve is a plot of the true positive rate (sensitivity) versus the false positive rate (1-specificity) as the threshold is varied. A perfect test would show points in the upper-left corner with 99.8% sensitivity and 99.6% specificity. For this problem the network in training and validation samples performed well, therefore the all ROC curve did relatively well.

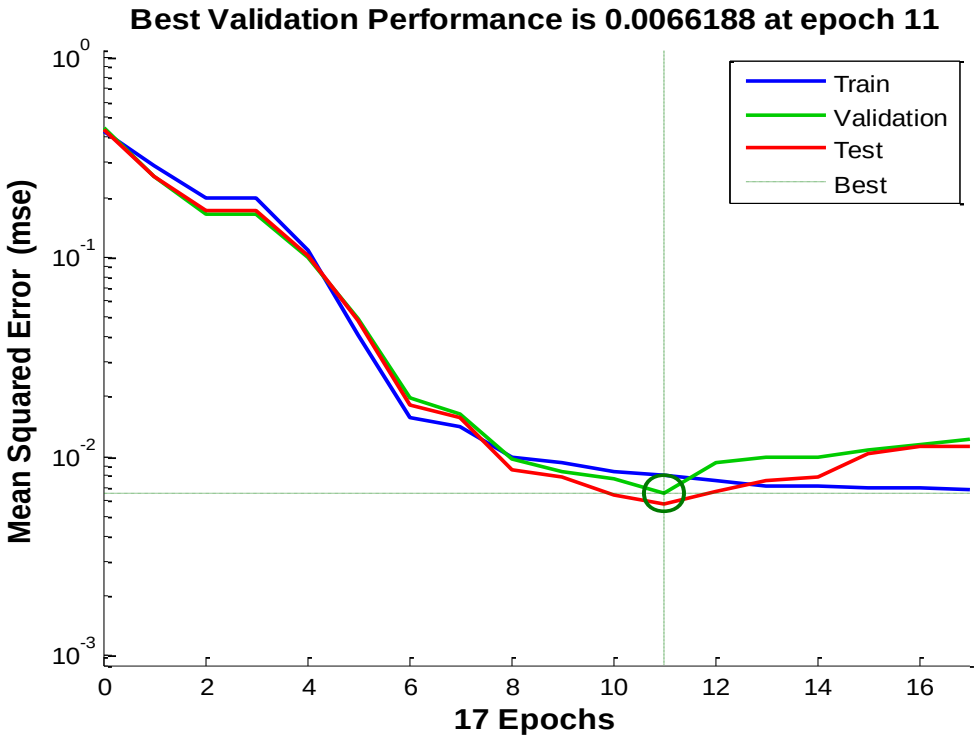

Figure 5.6 Performance (plot perform)

The figure mentioned before shows the performance of each the sample categories. As seen here the validation, train and test sample did their best performance at most points.

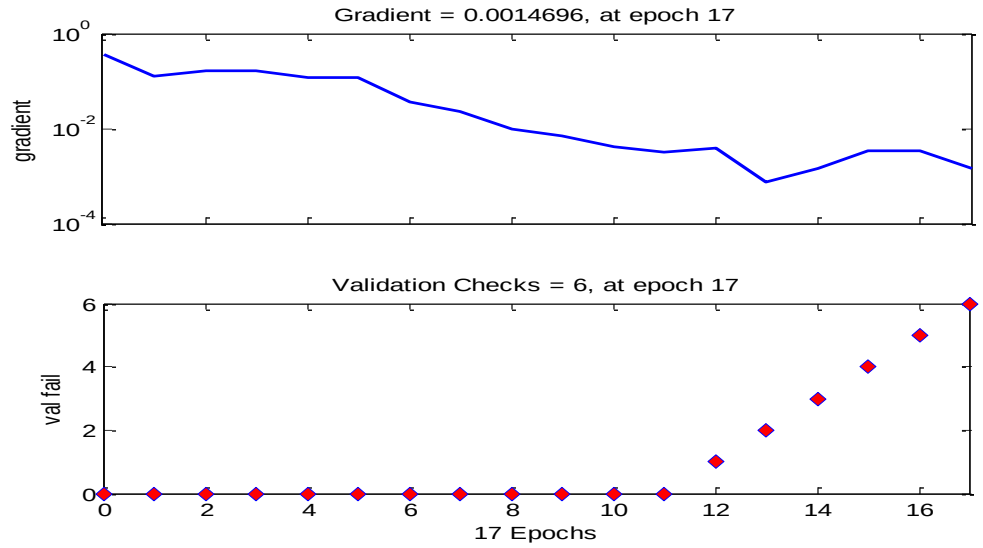

Figure 5.7 training state

5-4 Result between APB and PB

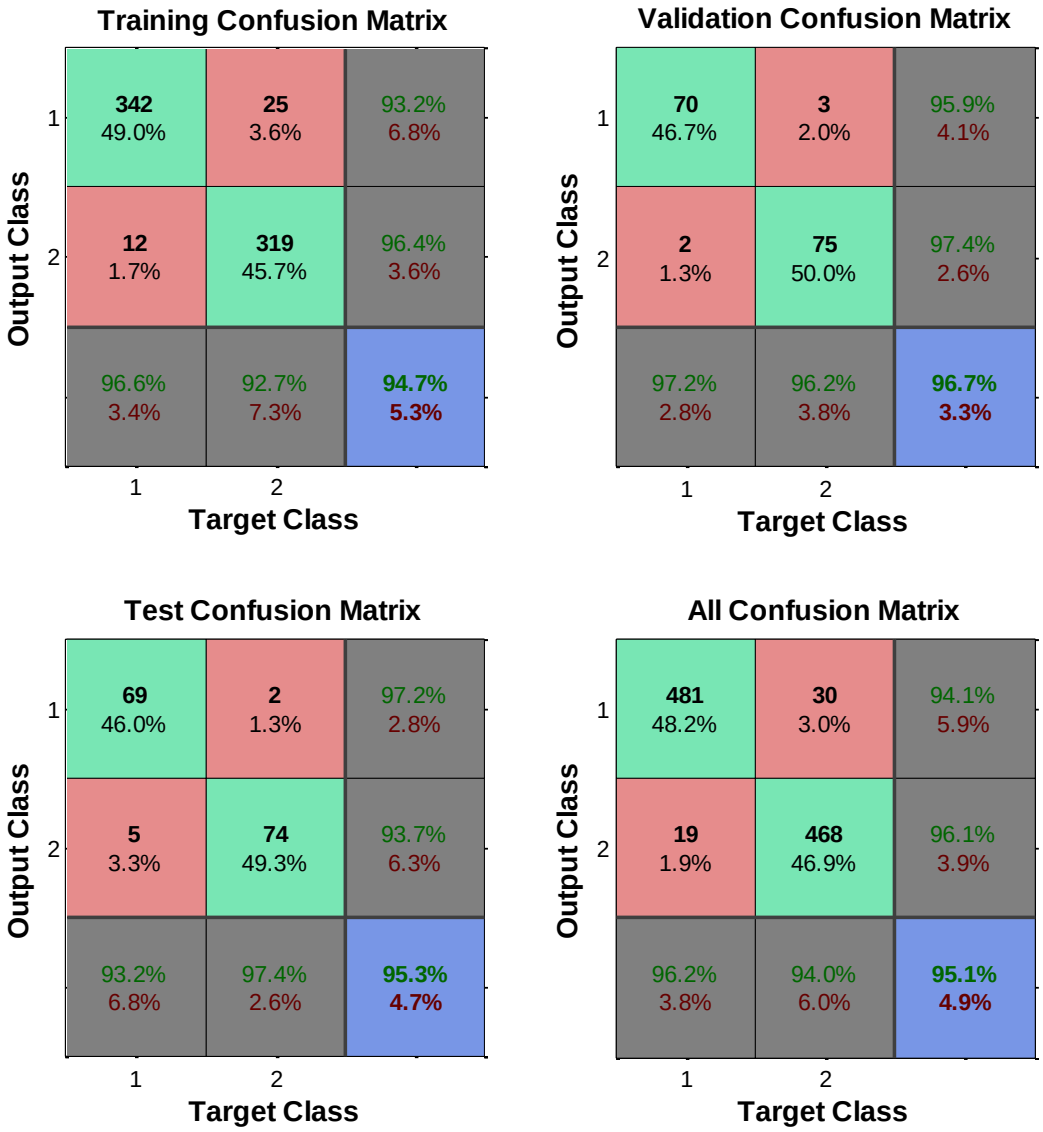

Figure 5.8 confusion window

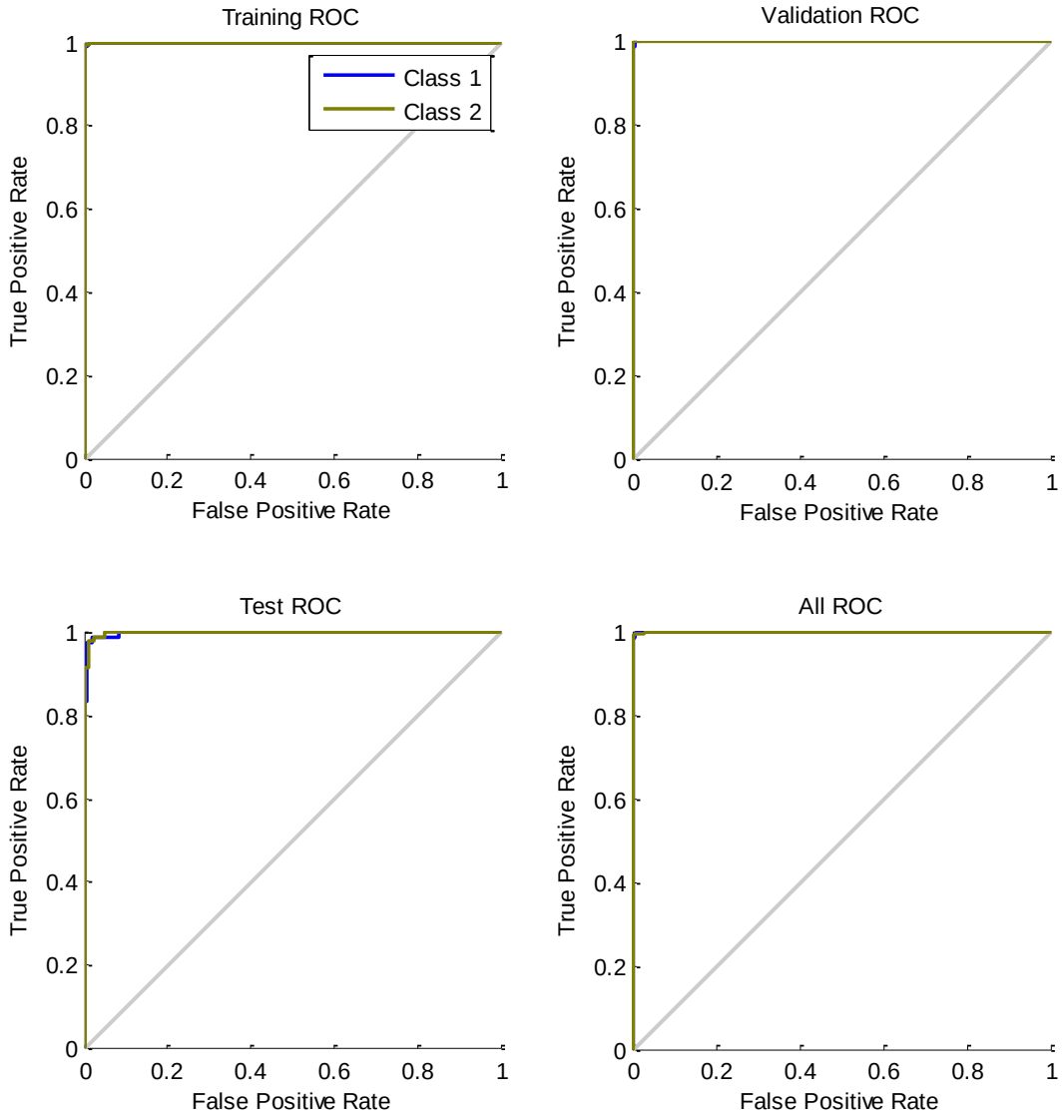

Figure 5.9 ROC

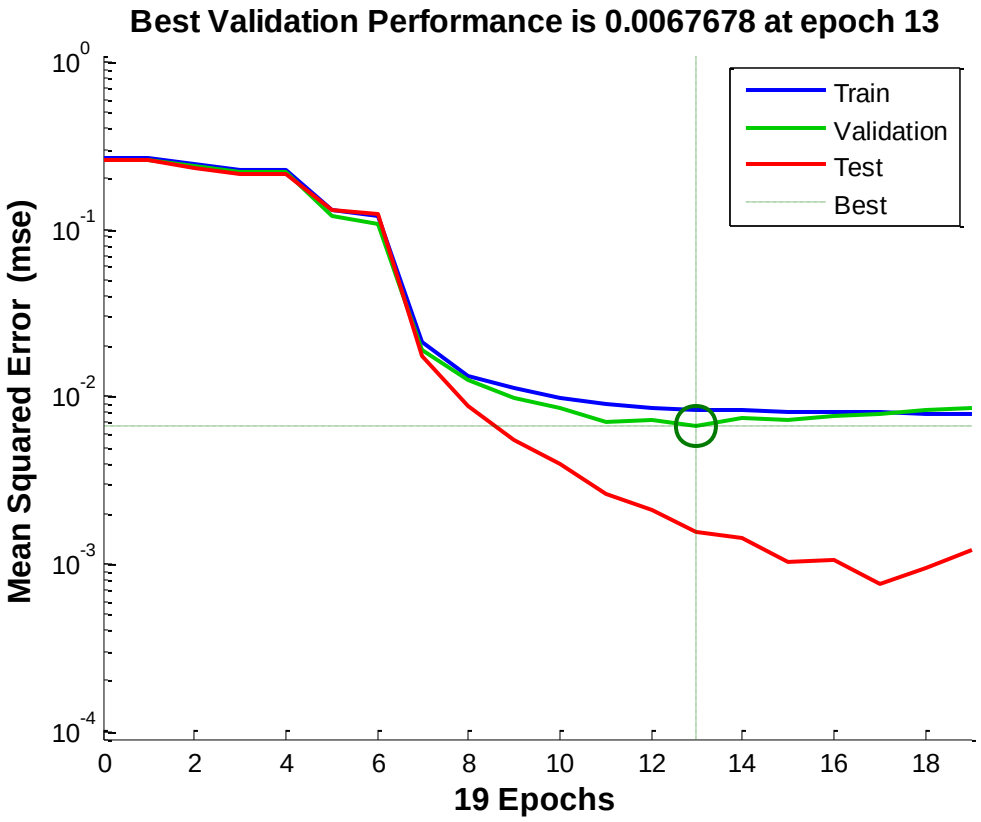

Figure 5.10 Performance (plotperform)

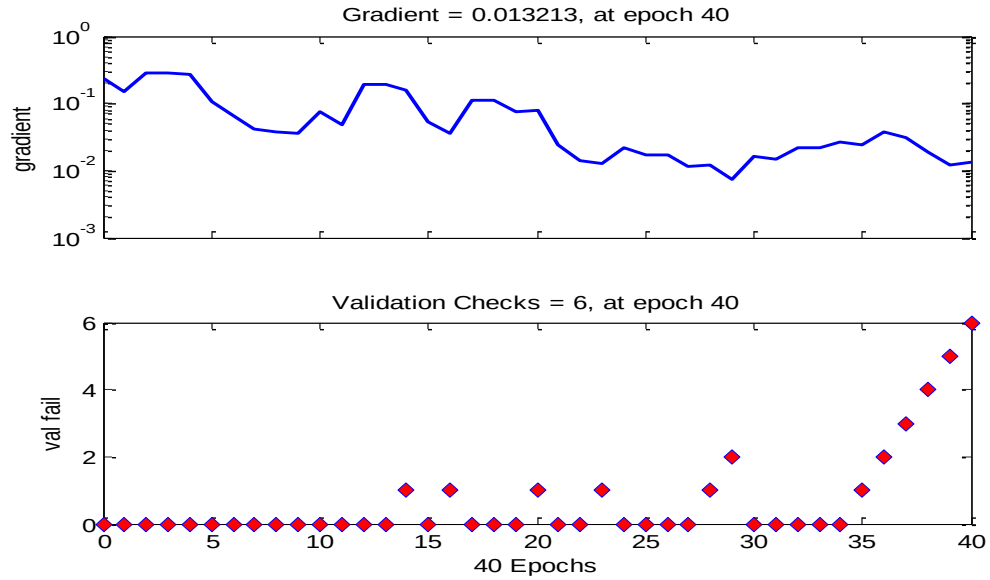

Figure 5.11 training state

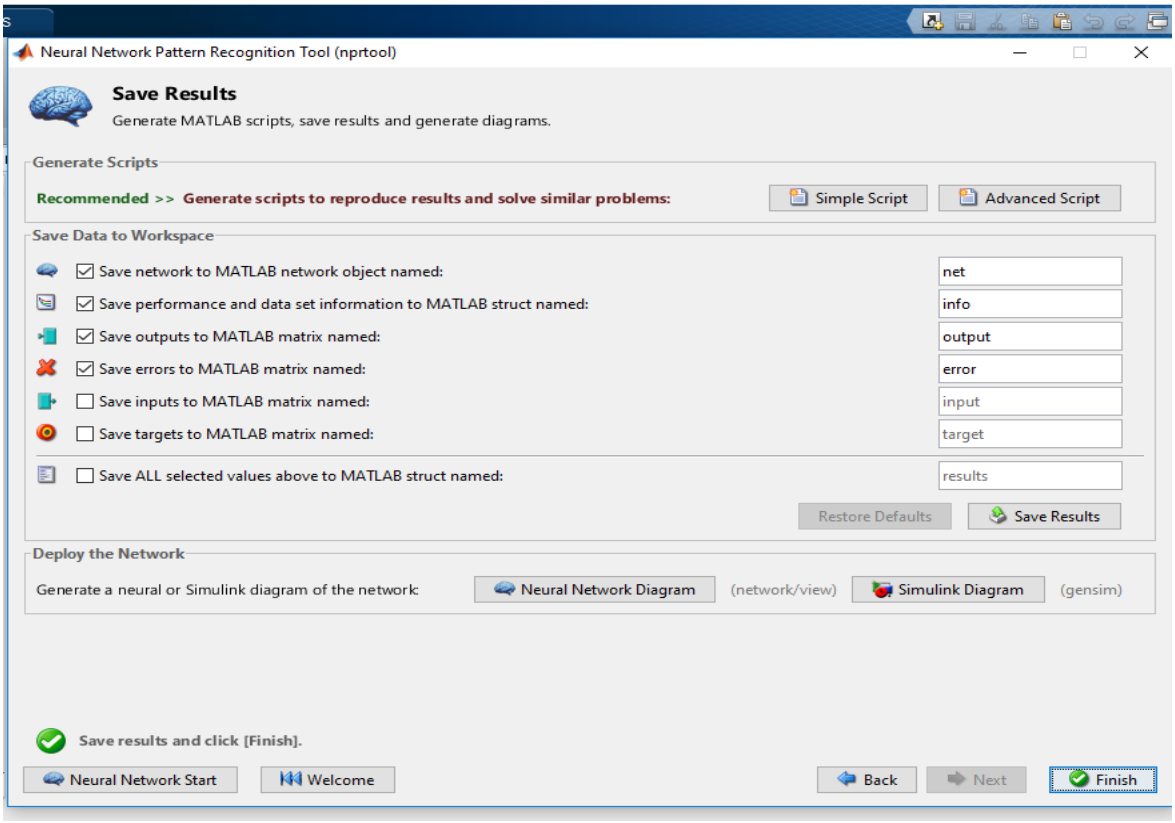

Figure5.12 Neural Network Pattern Recognition Tool

In the above window the network is saved as an M-file in the workshop for further testing.

5-5 Discussion

After all features were presented in an excel file and used in the development of the artificial neural network (ANN), all these data were used in the training of ANN with 1500 dataset containing features of QRS complex, intervals and the statistical features, the classification accuracy has been improved and the system yield high predictive accuracy of 99.7%, the sensitivity is 99.8%, and the specificity is 99.6% between the normal and abnormal signal ,and between APB and PB the accuracy was 95.1%, the sensitivity of 97.6%, and the specificity of 93.5% .

Chapter 6

Conclusion and Recommendations

6.1 Conclusion

Arrhythmia is one of fetal heart disease. Diagnosis of arrhythmia is critical to direct proper treatment. The ECG is the main bio signal used for such diagnosis , and the diagnosis still remain a complex process used to be decide by expert cardiologist.

The project succeeded in developing a program which provides ECG de-noising , feature extraction and classification as normal and abnormal sample. We used segmented samples which were obtain from the internet , also In this thesis we used simple matlab programming language in the feature detecting step and finally created the ANN using the (nprtool) for easier understanding of how the ANN works.

When we test the proposed system yield high predictive accuracy of 99.7%. the sensitivity is 99.8%, and the specificity is 99.6%.

6.2 Recommendations

- We used in this project two types of abnormal ECG signal so that we recommend for cover other ECG diseases.
- we can use different classification technique and compare between the results.
- We can improve the accuracy by including more features .this will result in prolonged training and testing, including more features has advantages of improved accuracy since the chance of finding an optimal subset of relevant features is always more, but the problem is the intense computation that might be carried out to do the classification; this again can be tackled by using various dimensionality reduction techniques.

REFERENCES

[1] Najarian, K, (2012). 'Electrocardiogram', '*Biomedical Signal And Image Processing' (2nd Edition),*171-193.

[2] Sahoo, P, J, (2011). *'Analysis of ECG signals for Detection of Cardiac Arrhythmias',* 'National Institute Of Technology',(769 008),204-244.

[3] Sarkaleh, M, K, (2012). *'Classification of ECG arrhythmias using discrete wavelet transform and neural networks',* (2.1:1).

[4] Ceylan, R; Ozbay, Y and Karlik, B, (2009). *'A novel approach for classification of ECG arrhythmiasm',* (36.3:6), 721-6726.

[5] Acharya, U. Rajendra, et al, (2017).*'Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network'* , (405), 81-90.

[6] Sambhu, D and Umesh, A, C, (2013). '*Automatic classification of ECG signals with features extraction using wavelet transform and support vector machines',*(2), 235-241.

[7] Ebrahimi, A and Addeh, J, (2015). *'Classification of ECG arrhythmias using adaptive neuro-fuzzy inference system',* (104), 134-140.

[8] LEE, Sang-hong; Chung; Kyung-Yong; LIM, and Joon S, (2014). '*detection of ventricular fibrillation using Hilbert transform',*(8.6),1315-1324.

[9] DAS, Manab K, ARI, Samit, (2014). ' *ECG beats classification using mixture of features. International scholarly research notices'.*

[10] LI, Hongqiang, et al. (2016).*' Novel ECG signal classification based on KICA nonlinear feature extraction',*(35.4), 1187-1197.

[11] MAR, Tanis, et al, (2011). '*optimization of ECG classification by means of feature selection',* (58.8), 2168-2177.

[12] Hosseini, H; Reynolds, K. J; Powers, D, (2001). '*A multi-stage neural network classifier for ECG events'*, 1672-1675.

[13] Ubeyli, E, D, (2008). *'Wavelet/mixture of experts network structure for ECG signals classification',* (34.3),1954-1962.

[14] Engin, M, (2004). '*ECG beat classification using neuro-fuzzy network. Pattern recognition letters'*,(25.15),1715-1722.

[15] Prakash, J, (2009). '*Analysis of ECG signal for Detection of Cardiac Arrhythmias',* (209),(C117).

[16] Ojha, D, K, (2005). '*Analysis of Electrocardiograph (ECG) Signal for the Detection of Abnormalities Using MATLAB'*.

[17] Prof. Al-Alaoui, M, A, (2006). *'Application of Artificial Neural and Fuzzy-Neural Networks to QRS Detection and PVC Diagnosis', 'department of electrical and computer engineering faculty of engineering and architecture American university of Beirut'*.

[18] V.Viknesh and P. Ram Prashanth, (2013). *'Matlab implementation of ECG signal processing, Final year ECE Panimalar Institute Of Technology Chennai'*, 2319 – 4197.

[19] Urban, P. and, E. Jothi, S, J, (2014*). 'Approach To Automatically Detect Cardiac Arrhythmia. International Journal for Research in Science and Engineering Technology.EIE Department, Karunya University*'.

[20] Harikumar, R. and Shivappriya, S, N, (2014). '*Analysis of QRS Detection Algorithm for Cardiac Abnormalities. International Journal of Soft Computing and Engineering (IJSCE)'*.

[21] Bailey, E. ; Sorrenno, J. ; Bolkhovsky, J , and Anderson, L, (2011). *'Automatic Detection of Atrial Fibrillation and Atrial Flutter'*.

[22] Moss, A , J, and Stern, S, (1996). *'Noninvasive Electro cardiology', 'Clinical Aspects of Holter'*.

[23] Morris, F; Edhouse, J; William, J, B, and Camm, J, (2003). *'ABC of Clinical Electrocardiography'*.

[24] Eng. Owis, M, (2001). '*Novel Techniques For cardiac arrhythmia detecction'*. Systems and Biomedical Engineering Department Faculty of Engineering'.

[25] Acharya, J, S; Suri, J, A, E; Spaan ,J ,and Krishnan, S, M (2000). *'Advances in Cardiac Signal Processing'*, 1-50.

[26] Israel, S, A; Irvine, J, M A, and heng, M. D, (2005). *'ECG to Identify Individuals*," Pattern Recognition', Elsevier, Vol. 38(1),133-142.

[27] Friesen, G, M, and Jannett, T .(1990). *'A comparison of the noise sensitivity of nine QRS detectionalgorithm',*. Vol. 37, 85–98.

[28*] Characterization of ECG Noise Sources*.

[29] Tikkanen, P, (1999). *'Characterization and Application of Analysis Methods for ECG and Time Interval Variability Data'*, 230-243.

[30] Carlo, J, D, (2002). '*Surface Electromyography: Detection and Recording'*, DelSys Incorporated, 2-6.

[31] Razavi, B, (2001). *'Design of Analog CMOS Integrated Circuits'*,'New York: McGraw Hill', 209-218.

[32] Viknesh,V, and Prashanth,, P,R, (2004). '*Matlab implementation of ECG signal processing, Final year ECE Panimalar Institute Of Technology Chennai'*.

[33] Tamil, E, M, and Kamarudin, N, H, (2008*). 'A Review on Feature Extraction & Classification Techniques for Biosignal Processing '(*Part I: Electrocardiogram), Proceedings 21, 107–112.

[34] Ilic, S, (2007). *'Comparison of Compression Ratios for ECG Signals by Using Three Time-Frequency Transformations'*, Vol. 20, no. 2, 223-232.

[35] John, R, H, (2007). *'the ECG Made Easy'(6th ed)*. Churchill, Livingstone.

[36] LI, Hongqiang, et al. (2016). '*novel ECG signal classification based on KICA nonlinear feature extraction. Circuits, Systems, and signal processing'*, (35.4),1187-1197.

Appendix

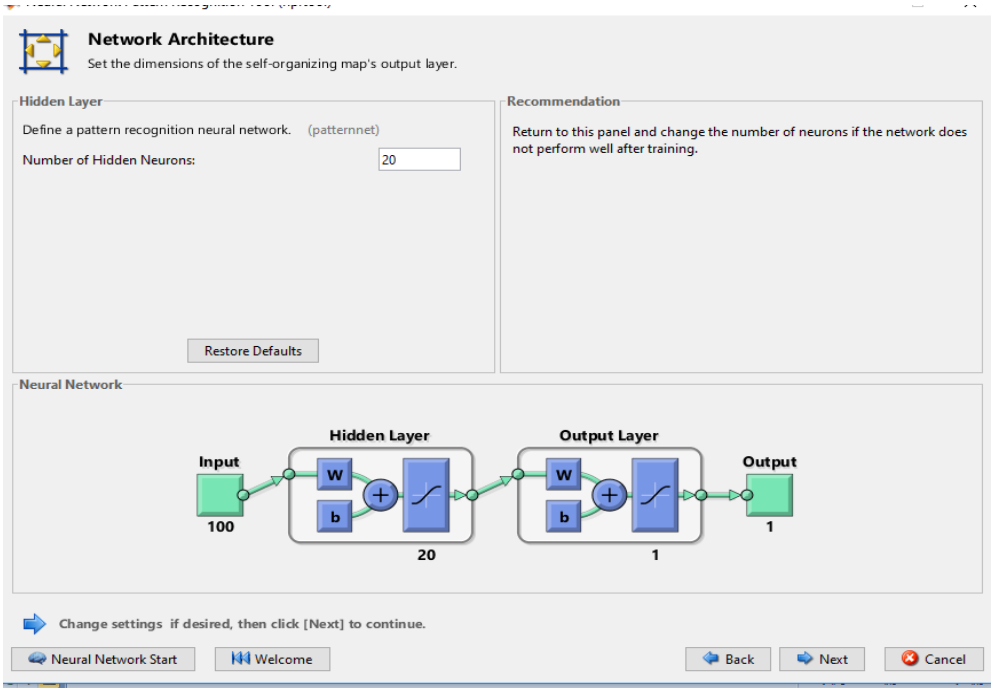

Figure 1: 20 neurons in the hidden layer is selected

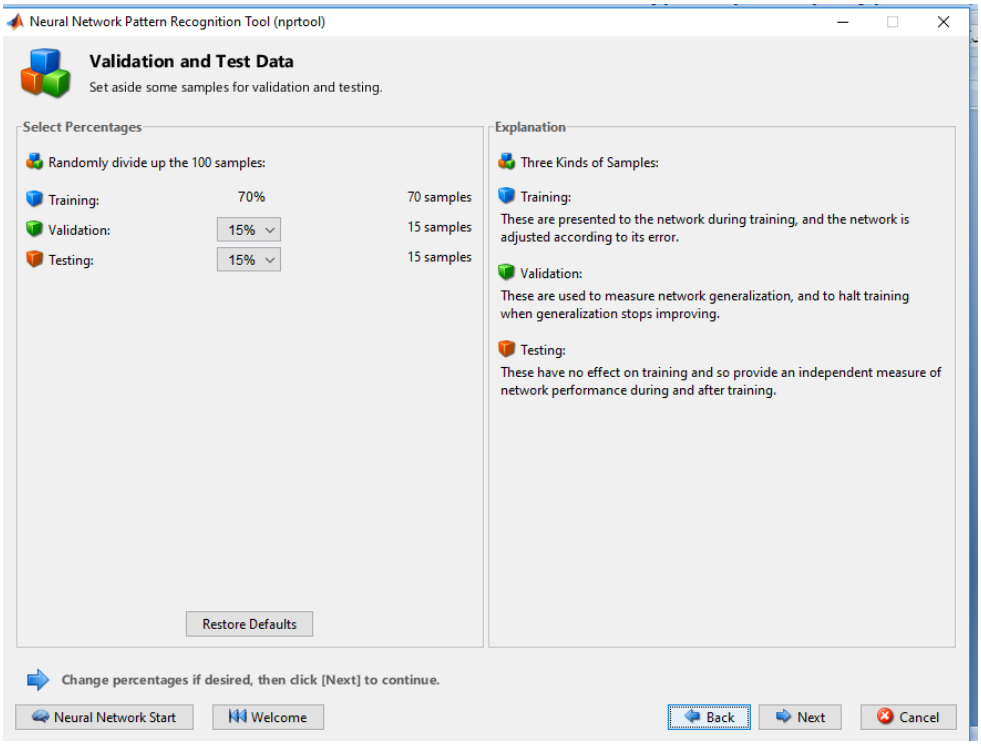

Figure 2: then the percentage of the training, testing, and validation samples are selected.

After entering all the data required for creating the ANN the next step is to start the training process.

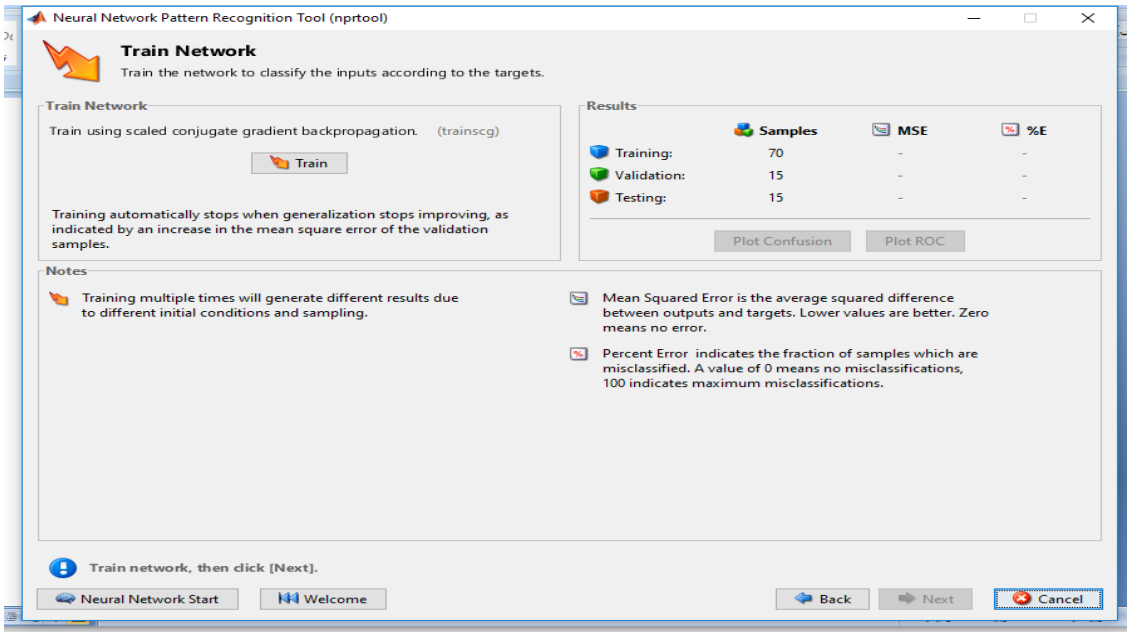

Figure 3: training window

After this step results are reviewed and decided whether further training is required or not.