Study on Biomedical Signals Using Signal Processing Techniques

A Dissertation submitted

in partial fulfillment of the requirements for
the degree of

Master of Engineering

in

Thermal Engineering

by

Aditya Pathania

Registration No.: 801683003

Under Supervision of

Dr. A. Mittal (Assistant Professor)

THAPAR INSTITUTE
OF ENGINEERING & TECHNOLOGY
(Deemed to be University)

MECHANICAL ENGINEERING DEPARTMENT
THAPAR INSTITUTE OF ENGINEERING AND TECHNOLOGY, PATIALA

June, 2018

CERTIFICATE

I hereby declare that the thesis entitled **"Study on Biomedical Signals Using Signal Processing Techniques"** is an authentic record of my work carried out as requirements for the award of **Master of Engineering in Thermal Engineering** at **Thapar Institute of Engineering and Technology, Patiala** under the supervision of **Dr. Anu Mittal (Assistant Professor)**, from 21st July 2016 to 15th June 2018. No part of the matter embodied in this report has been submitted to any other university or institute for the award of any degree.

Aditya

Date: 15th June 2018 Student's Signature

It is certified that the above statement made by the student is correct to the best of my /our knowledge and belief.

Anumittal 18-06-18

Dr. Anu Mittal
Assistant Professor
Mechanical Engineering Department
Thapar Institute, Patiala-147004

Acknowledgements

I would like to express my deepest gratitude and indebtedness to my supervisor, Dr. A. Mittal for their excellent guidance, caring, patience, and providing me an excellent atmosphere for doing my research work. The opportunity, support, exposure and atmosphere provided by the Thapar University, Patiala, to carry out my studies are highly appreciated. A special debt of gratitude is owed to the authors whose works I have consulted and quoted in this work. Last but not least, I am forever grateful to my parents, family and friends for their unconditional support and best wishes.

I am grateful to almighty for those joyful moments I enjoyed and painful instances which made me tough and strong to face situations in life to come and for the exceptional journey and memories at Thapar Institute of Engineering and Technology, Patiala.

Aditya Pathania

Abstract

Electrocardiogram (ECG) has been used by the doctors for a long time to monitor the condition of the heart and is popular due to its ease of use and non-invasive nature. Electrocardiogram (ECG) is a record of electrical activity of a heart. Detection of heart diseases at an early stage can prolong life through appropriate treatment, and the minute variations in the signal are not always easily discernible to a naked eye. Thus, our objective regarding ECG signals is to use different signal processing and pattern recognition techniques to extract useful information from them and to try to increase the classification accuracy, even in noisy signals.

An attempt has been made to detect the R-peaks present in the QRS complex. This is done because R-peaks are easier to detect and further help us, to extract features which will help in classification. The signals used in this work, have been downloaded from a very popular dataset MIT-BIH database, which has been extensively used in prior scientific research. When the algorithm was applied on the dataset, it was found that almost all the peaks were being detected in the signal. Further the techniques of noise reduction and baseline-wander reduction were also applied on the signals. It was achieved by converting the signal from time domain to frequency domain.

After noise filtering, the wavelets transform was used to extract features from the signal. Wavelet transform was used because, it works in both time domain and frequency domain, which allows to extract more information from the signal. First, the signal was filtered to remove noise and then different features were extracted from the wavelet coefficients like mean, mode, variance, standard deviation, Shannon entropy, Spectral energy etc. Then these features were used to see, which ones could help in classification of the signal. It was found out that for most of the feature set could not classify the ‘Disease’ signal from ‘Non-Disease’ signal. Only mean vs log energy graph showed some non-overlapping dataset, and thus could be used for further classification.

Finally, ANN (Artificial Neural Network) and SVM (Support Vector Machine), was attempted to classify the signals. Both are Machine Learning algorithms, but work on different principles. At first a model was created using SVM, to classify the signals. However, the model could provide only 50% accuracy. The reason for less accuracy was due to the dataset, which was used for training, was highly overlapped. Next, a model was created using ANN, which

provided better results. The training and testing accuracy was about 95%-96%, and the algorithm was tested on 9 disease and 9 non-disease signals, in which all the signals were successfully classified.

Contents

1. Introduction

2. Literature Review

3. Feature Extraction and R peak Detection from the ECG Signal

4. Wavelet Analysis

5. Classification of ECG Signals (ANN and SVM)

6. Conclusion and Scope of Further Research

References

List of Figures

List of Tables

Nomenclature

$R(k)$ - Autocorrelation function

H - Spectral Entropy

p_f - probability density function

$F(n)$ - Integrated time series

f - frequency

ω - Natural frequency

ξ - Damping coefficient

μ - Forcing input

$F(u)$ - Fourier Transform

$|F(u)|^2$ - Power Spectrum

Ψ - Wavelet Function

a - Scaling parameter

b - Shifting parameter

$\varphi(v)$ - Activation Function

e_i - error

w_{ij} - weight

α - learning rate

v_i - weighted sum of the output node

$J(w)$ - cost function

Acronyms

ACF	Autocorrelation Function
ANN	Artificial Neural Network
ART	Adaptive Resonance Theory
AV	Auriculo-Ventricular
BBS	Best Basis Selection
BPM	Beats-per-minute
CC	Centered Core
CVD	Cardiovascular diseases
CWD	Choi-Williams Distribution
DomF	Dominant frequency
DWT	Discrete Wavelet Transform
ECG	Electrocardiogram
EEG	Electroencephalogram
EMG	Electromyogram
FFT	Fast Fourier Transform
FM	Median frequency
HF	High frequency
HFA	High fuzzy automatons
LDB	Local Discriminant Basis
LF	Low frequency
MATLAB	Matrix Laboratory
ML	Machine Learning

MLBS	Multi-Level Basis Selection
MOE	Mixture Of Experts
MRA	Multiresolution Analysis
OAO	One Against One
PCA	Principal Component Analysis
PCG	Phonocardiographic
PKS	Peak size
PSD	Power Spectral Density
PVC	Premature Ventricular Contraction
SAN	Sinoatrial Node
SE	Spectral Entropy
SP	Sliding Fourier Transform
STFT	Short Time Fourier Transform
SVM	Support Vector Machine
SWD	Smoothed Wigner Distribution
TQWT	Tuneable Q Wavelet Transform
ULF	Ultra low frequency
VF	Ventricular Fibrillation
VLF	Very low frequency
VT	Ventricular Tachycardia
WD	Wigner Distribution
WT	Wavelet Transform

Chapter 1

Introduction and Objectives

1.1 Introduction

Cardiovascular diseases (CVDs) are a collection of diseases like stroke, heart failure, heart arrhythmia, congenital heart disease, valvular heart disease, etc. Tobacco use, alcohol consumption, unhealthy diet, physical lethargy, etc. are the main culprits that may increase the risk of heart attacks. These life decisions can cause problems like obesity, high blood pressure, high blood sugar and high blood cholesterol, etc.

Heart diseases are the major cause of death globally. Approximately, around 17.5 million people (31%) died from heart diseases in 2012 (WHO Report, 2013).

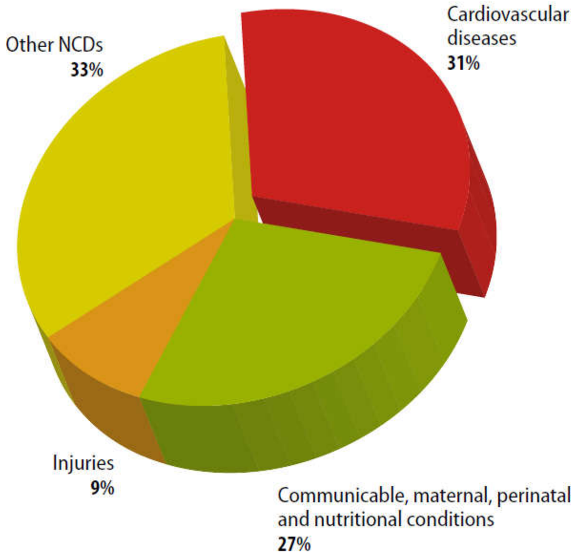


Figure 1.1 Distribution of major causes of death including Cardiovascular Diseases (CVD) (WHO Report, 2013)

According to the WHO, the South Asian region has one of the highest cardiovascular death rates in the world. In India, CVDs cause up to 45% of all non-communicable diseases deaths. WHO report of 2002 shows that cardiac diseases (CVDs) will be the major cause of fatality by year 2020-2021 in India. Around 2.5 million Indians are projected to die due to heart diseases which makes up 54 % of deaths by heart problems (WHO Report, 2002). These deaths are most likely to happen in individuals aged between 30-70 years. Presently Indians experience cardiac deaths 10 years prior than in other countries. Because CVDs have a high rate of mortality especially in today's time, research needs to be done in this field.

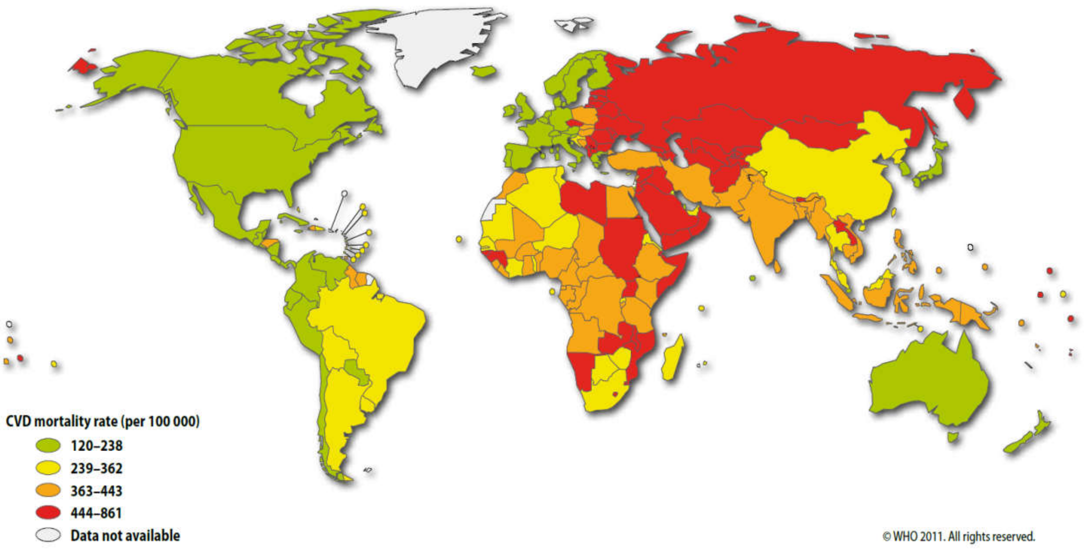


Figure 1.2 World map showing the global distribution of CVD mortality rates (age standardized, per 100 000) (WHO Report, 2013)

The automation of CVD detection is becoming of increasing importance due to number of potential causes such as:

- Subtle variations in ECG signal might not be easily conspicuous to the naked eye.
- Signal might have noise and baseline wander embedded in it.
- Diagnostic tests like the treadmill test are risky because the patient might develop tachycardia and eventual heart failure.
- Catheterisation is invasive and takes considerable time
- Most of the methods being used require trained technicians and doctors.
- The equipment of diagnosis is expensive.
- ECG diagnosis is subjective and may vary depending on the experience and expertise of the medical practitioner.

Therefore, manual diagnosis is always associated with errors and automation of the process is required for accurate diagnosis. One of the major hurdles in classifying ECG signals is that in some instances identical defects have dissimilar wave characteristics. And in some cases, patients having different diseases might have similar ECG waves, hence complicating the diagnosis. Therefore, further research is required in order to obtain more precise and accurate diagnosis. Most of the research is being performed in the studying biomedical signals because of the potential benefits such as:

- Heart signals are non-stationary in nature and contain a lot of information.
- ECG can be useful for detection of other diseases such as diabetes, depression, etc.
- The process is cost effective and easy to use.
- Easily accessible in villages and cities.

Because so many people die all around the word due to CVDs, it is necessary to study and research in medication and diagnostics to decrease the mortality rate. Huge part of the research has been concentrated on the processing of biomedical signals. Daily clinical practice generates any number of biomedical signals during monitoring of patients and for diagnostic purposes. Therefore, automatic processing systems are frequently used in medical data analysis. New methods can simplify and speed up the processing of large volumes of data. The physician very frequently has to decide a patient’s diagnosis on the basis of a number of numerical values measured during examination. The aim of this study is to help the cardiologist in diagnosis through different signal processing techniques.

Most of the research work being done as on date is based upon detecting and classifying ECG signals. Several ongoing studies are being performed on different types of pattern recognition and machine learning techniques to analyse biomedical signals. The different methods being used are Support Vector Machine (SVM), Artificial Neural Network (ANN), Markov algorithm, operating curves, root mean square of successive differences, etc (Kadir et al. 2016). The features extracted for the above-mentioned techniques are usually done in time domain, frequency domain and time frequency domain. Features extracted in time domain consisted of R-R intervals, QRS complex, peak position, etc. In frequency domain features like Fast Fourier Transform (FFT), Power Spectral Density (PSD), etc and in time- frequency domain features like Wavelet Transform (WT), Short Time Fourier Transform (STFT), Z transform, etc were experimented with. However, so far people have worked on features extracted from time domain and frequency domain, not much work has been done regarding the frequency peaks, we get after applying FFT (Fast Fourier Transform). Also, the accuracy achieved so far was limited to 90-91 % (Anuradha et al. 2008) and required further research to improve the accuracy of the model being used for feature extraction.

In order to classify and characterise the signals it is vital to understand the basic anatomy of the heart and the mechanism with which the ECG signals are generated. Therefore, in the following section a very brief description has been provided on heart anatomy and signal generation and subsequently some of the techniques, used for extracting useful information from the signals, have been discussed.

1.2 Heart Anatomy:

The heart is an organ which helps to pump blood via blood vessels so that it can reach all body tissues. This process is required to facilitate exchange of materials between blood and body organs.

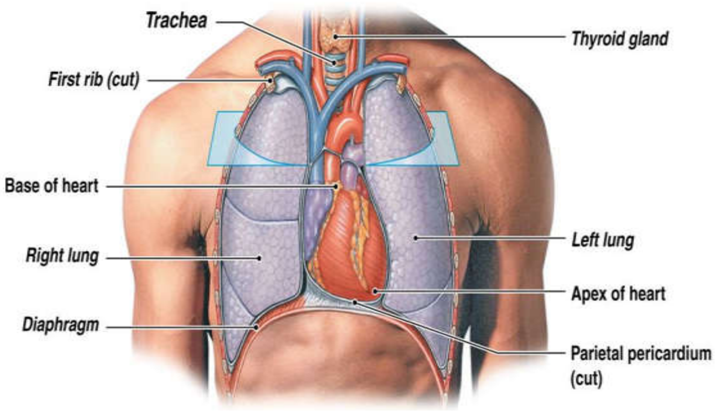


Figure 1.3 Location of Heart (Ganong, 1995)

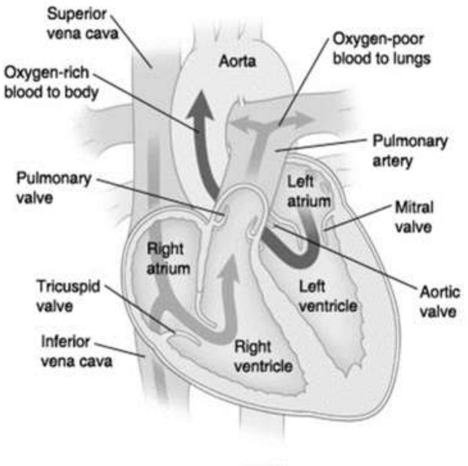


Figure 1.4 Chambers and Valves of Heart (Ganong, 1995)

The heart has four chambers and many arteries and veins which help in circulation of blood throughout the body. Blood travels from the right atrium into the right ventricle via the tricuspid valve. From here, blood is taken to the lungs for oxygenation via pulmonary artery. After oxygenation, the blood is carried into the left atrium by pulmonary vein and from left atrium, the oxygenated blood is transported to left ventricle through a bicuspid valve. This oxygenated blood then moves from left ventricle to other parts of the body by an artery called aorta.

1.3 Conduction system of the Heart:

Heart consists of special muscle fibres called autorhythmic fibres which are self-excitable. They generate action potential, which causes heart to contract. These fibres act as pacemakers, and provide a conduction pathway, for the signal to travel throughout the heart muscles, which enables it to contract in a co-ordinated manner.

1.4 Action Potential:

Action potential (AP) is a sequence of events, which occur very quickly, and in which the cell membrane potential is reversed from -70 mV to +30mV (called depolarisation), and vice versa (repolarisation). These processes occur because of the transfer of Na+ and K+ ions, through channels present in the cell membrane. Action potential initiates in the Sinoatrial Node (SAN), causing atria to contract. Then the AP reaches the atrioventricular node and travels via right and left Auriculo-Ventricular (AV) bundles to reach the Purkinje fibres, leading to contraction of the ventricles (Barrett et al. 2015).

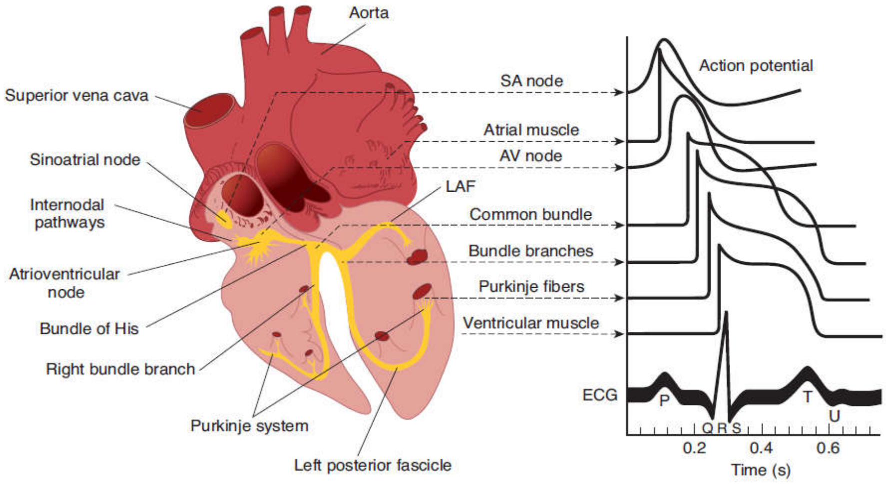


Figure 1.5 Conduction Pathway of the Signal (Ganong, 1995)

1.5 Electrocardiogram (ECG):

The fluctuation in AP of heart muscles can be recorded from outside the body, as the body fluids are decent conductors. This recording over time is known as ECG (Barrett et al. 2015). A positive deflection occurs when the action potential moves towards the active electrode and vice versa.

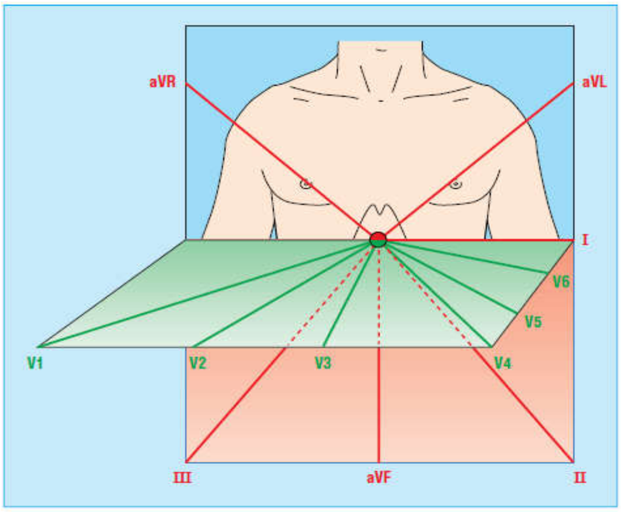


Figure 1.6 12-Lead System (Hampton and Adlam, 2013)

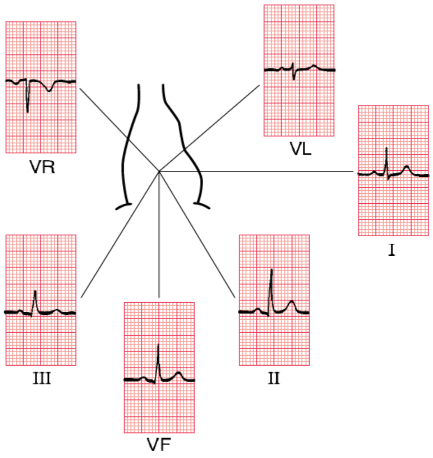


Figure 1.7 Figure 7 Signals in Different Leads (Hampton and Adlam, 2013)

The signal which is seen on the machine, is the mean average of the deflection in each lead, as shown in figure 1.7.

1.6 ECG Signal Generation:

In a typical ECG signal three peaks are easily discernible to a naked eye as shown in the Figure 1.8. The first wave is the P-wave, which occurs when atrial depolarization takes place.

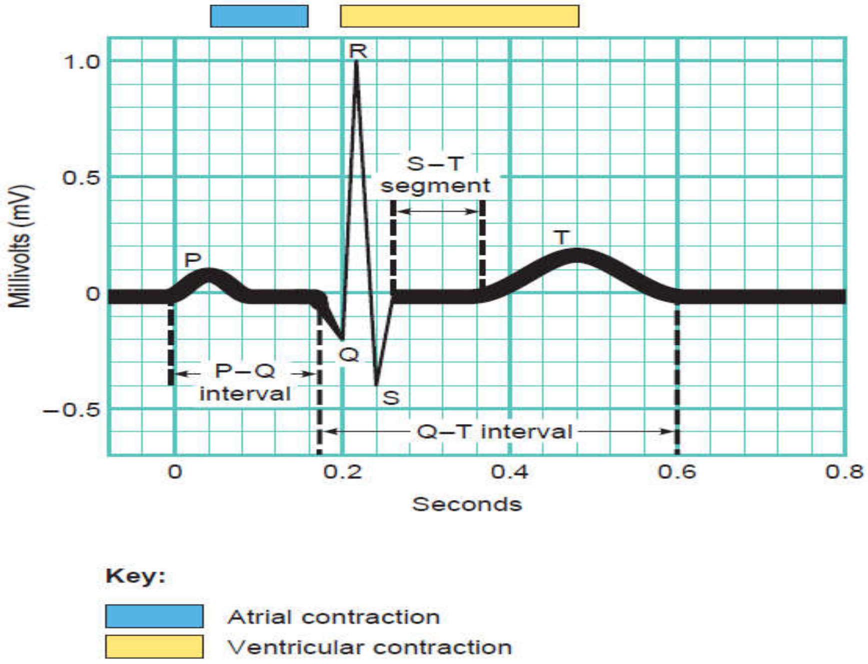


Figure 1.8 Typical ECG Signal (Tortora, G.J. and Derrickson, B.H., 2008)

The second deflection is called the QRS complex, which is the result of rapid ventricular depolarization. This shows action potential spreading through ventricular contractile fibers. The third deflection is the T wave, which tells us that ventricles have started to relax.

The time segment in a typical ECG signals for normal person have been provided in the Table 1.

S.No.	Intervals	Average Duration(s)	Events
1	P-R	0.18	Atrial depolarization and conduction through AV node
2	QRS duration	0.08	Ventricular depolarization and atrial repolarization
3	QT interval	0.4	Ventricular depolarization plus, ventricular repolarization
4	ST interval	0.32	Ventricular repolarization (during T wave)

Table 1. Time duration of different segments of an ECG

1.7 An overview of methods being used for ECG Signal analysis:

Various methods are being used today to identify and analyse ECG Signal such as filtering methods, neural network and wavelet transforms etc. (Cohen (1988)). Automatic ECG beat classification is based on extracting features from the beat. This can be achieved in two ways: direct method (time domain) and transformation method (deals with other domains) (Dokur et al. 2001). In the literature, two very commonly used methods for feature extraction were (i) Fourier transform Dokur et al. 1997, and (ii) wavelet transform I. Jouny et al. 1994. T. Olmez 1997 tried to classify ECG waves using Fourier transform. A brief introduction to some of the techniques are discussed below:

1.7.1 Time Domain Analysis: This belongs to the direct analysis methods. Here, time domain-based features like thickness and elevation of QRS complex, RR interval, QRS complex area, etc. are extracted and feature vectors are created after this transformation.

This technique is simple and is used to extract basic features like the mean, standard deviations, highest peaks, distance between peaks etc. from the signal. They are based on RR intervals, which correspond to the highest peaks in the signal. As the amplitude of the R peaks is the highest, their location is easily found out, and in turn can help to find other peaks also. The location of RR peaks also helps us find average Beats-per – minute (BPM). No changes to the signal are made in the time domain analysis.

1.7.2 Frequency Domain Analysis: The time domain analysis is easy to implement and does not have much computational burden, but a major drawback is that it is susceptible to noise. Also, many features are hidden in frequency domain, which are not easily seen in the time domain. FFT is used to switch from the time domain to frequency domain.

Fourier transformation gives the signal spectrum and range of frequency amplitudes embedded in the signal. But, Fourier transform only delivers the spectral constituents, not their temporal relations. As the morphological characteristics and the frequency values of the ECG signal varies in time, the analysis of the electrocardiogram signal according to time variation is also equally important to properly describe the ECG signal characteristics.

The main frequency regions of the heart are provided in the Table 2.

Frequency Type	Frequency Range(Hz)
Ultra-low frequency (ULF)	$0.0001 \leq ULF < 0.003$
Very low frequency (VLF)	$0.003 \leq VLF < 0.04$
Low frequency (LF)	$0.04 \leq LF < 0.15$
High frequency (HF)	$0.15 \leq HF < 0.4$

Table 2. Frequencies of an ECG Signal (Blinowska, K.J.)

1.7.3 Time-Frequency Domain Analysis (Wavelet Analysis): Time domain and frequency domain do not always give all the information present in the signal when applied independently. When the analysis performed in the time domain, then information in the frequency domain is lost, and vice-versa. So, if the frequency components of the signal are needed then we perform FFT. But the problem with FFT is that the technique fails to give the information regarding the exact location of frequency components in time. The frequency components of the signal vary with time, and thus we need a method to see variations in frequency components with respect to time.

Wavelet transformation can be used to study ECG signals as they are dynamic in nature and change their properties with time. Every wavelet has an individual frequency bandwidth and time duration.

1.7.4 Pattern Recognition using Classifiers (ANN and SVM):

Both Artificial Neural Network (ANN) and Support Vector Machine (SVM) are types of Machine Learning (ML) algorithms which work with data. Data can represent anything, like audio, images, signals etc. The "model" which we get is the result of Machine Learning. These are learning techniques because initially training data is provided to the algorithm for supervised learning.

ML helps us in solving various problems in several medical fields. It helps in the study of the important clinical variables and of their combinations for prognosis, e.g. prediction of disease, extraction of medical knowledge to help in patient management.

ANN: Artificial Neural Network (ANN) is a mathematical algorithm which helps to build transient state thus allowing the machine to function in a more human like manner by mimicking biological neural networks. It tries to process human like a human brain. A huge number of interrelated neurons work in co-ordination to process information and provide meaningful results.

ANN may contain the following 3 layers:

- Input layer – this layer has nodes where raw information is fed into the system.
- Hidden layer – Analysing and processing occurs here by iteratively changing the weights and biases of the connections between the nodes. There might be more than one hidden layers.
- Output layer – This layer gives the end result after analysis and processing.

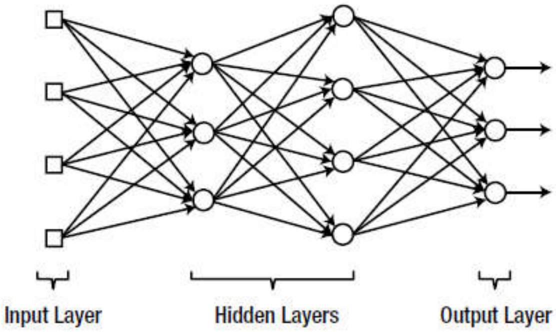


Figure 1.9 A typical neural network (Kim, 2017)

SVM – Support Vector Machine: It is another type of machine learning algorithm in which training data is used to classify data. They analyze the large amount of information to recognize patterns from them.

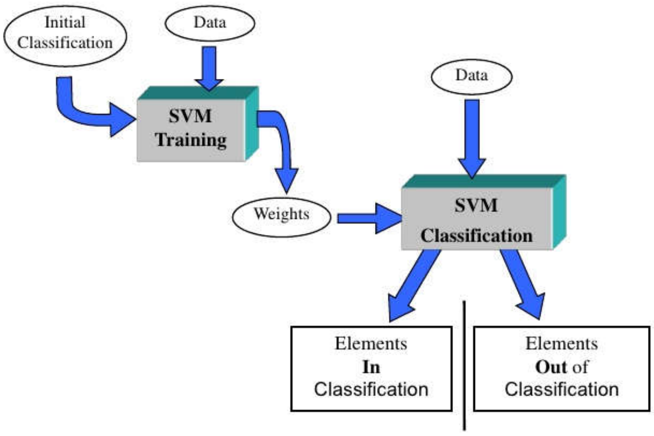


Figure 1.10 Flowchart of SVM process

An SVM tries to generate a hyperplane which helps in data classification.

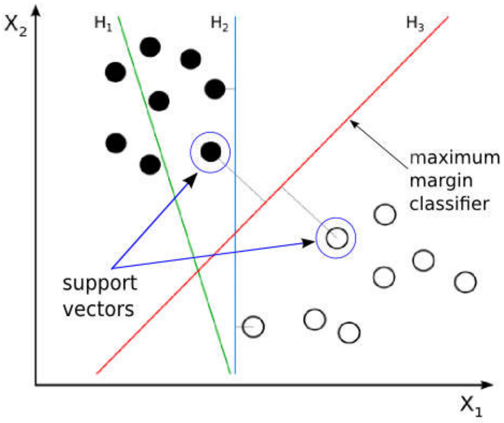


Figure 1.11 Support Vector Machine

SVM maximizes the margin between the data points and the hyperplane, to minimize the error. After training is done, the testing data is input, which is then classified accordingly. SVM provides the largest flexibility, of all the classifiers.

Based on the critical review of existing literature on biomedical signals, ECG signals their classification and feature extraction it was recognized that for accurate diagnosis of heart signals further improvement in the accuracy of the models is required. Therefore, in the present work attempt has been made to achieve the following research objectives.

1.8 Research Objectives

1. To investigate into biomedical signals (specifically ECG signals), using automatic classification by extracting time and frequency-based features.
2. To develop two models based on ANN (Artificial Neural Network) and SVM (Support Vector Machine) for classifying ECG signals.
3. Comparison of developed models and parameter (such as learning rate, batch size, regularisation type etc.) optimisation in order to improve the accuracy of models.

Chapter 2

Literature Review

2.1 Introduction:

ECG signals have been used for a long time to detect the problems in the cardiovascular system. Initially, doctors used visual interpretation to study the signals. Then time domain analysis was used, but it was too simple and did not provide much information. Recently methods like frequency domain analysis, time-frequency domain analysis, machine learning etc. are being used and researched upon. Many techniques have been developed and are still being developed, to automate this process.

QRS detection by using a computer is being researched upon for more than 30 years. The evolution of these algorithms is due to the improvement in computer power in the recent decade. Earlier calculational load used to be huge, which factored in determining the type of algorithm being used to study the signal but now the computational load becomes less and less important. The only exception from this trend is probably the development of QRS detection algorithms for battery-driven devices.

Many new methods for ECG signal analysis have been used previously like, algorithms from ANN (Artificial Neural Networks) (Hu et al. 1993), genetic algorithms (Pol et al. 1995), wavelet transforms, filter banks (Afonso et al. 1999) as well as empirical techniques mostly based on nonlinear transforms (Köhler et al. 2001).

In the following sections literature has been presented based on methods which have been used to study ECG signals in the past and the advantages of machine learning (ANN and SVM) over those conventional techniques.

2.2 Conventional Methods for Heart Signal Analysis:

Previously used ECG classification methods are based on pattern matching and statistical pattern recognition. In contrast to the neural network technologies, these conventional practices used are rarely based on the intuitive reasoning used by cardiologists. Thus, their performances may not be similar to that of the neural technologies.

Usually, a doctor, by simple observation seems to look at the ECG data at hand to evaluate and identify actions based on the structures of the signals. The techniques used to study and interpret an ECG signal is different from that of a cardiologist because they deal with a sequence of samples rather than with a visual representation of a signal.

The techniques that have been used up until now are based upon different approaches like:

1. Numerical algorithms (resulting from signal analysis)

2. Clustering and pattern matching methods (resulting from statistical pattern recognition)

But as these techniques are based upon the mathematics used by engineers and physicists, it becomes difficult for the doctors to fully understand the nuances, and thus find the process cumbersome and confusing.

Usually, these approaches of supervising and diagnosing arrhythmia rely upon the detection of some specific signal features. The approach these techniques take is to convert qualitative diagnosis criteria of the signal. The techniques/methods that have been used to tackle this problem are:

- Auto-correlation
- TSA (both in time and frequency domain)
- wavelet transforms
- features after FFT

Guillén et al.1990 used the Autocorrelation Function (ACF) method based on peak analysis for detecting arrhythmias. Initially, it was assumed that detection of arrhythmias could be achieved via the behavior analysis of some ACF variables such as relative peak amplitude and relation between peak width and period.

Time series ACF is defined as a new time series in which the general term is defined as

$$R(k) = \frac{\text{Autocovariance (k)}}{Variance} \qquad 2.1$$

where k lies between 0 to N-1 (N is the length of the original time series).

The time series, of the signal taken from a patient, is having periodical features, the function peaks appear at the same interval of the signal period and their amplitudes lie exactly in a straight line given by function, R(k) = 1 - k/N. If the signal is showing irregular features, then the ACF peak decreases at a higher rate. Based on these initial assumptions, a set of variables are used to measure the differences between VT (Ventricular Tachycardia) and VF (Ventricular Fibrillation) by analyzing the peaks in ACF:

- Positive peak (maximum value of ACF that lies between two zero crossing)
- The time when positive peaks have occurred
- The time passed between the preceding and the succeeding zeros.
- The delay in positive peak appearance
- The effect of peak value on the level of control.

2.3 Frequency analysis:

Another method used is frequency analysis, which helps us to classify the signal based on the contributions of the frequency components. In a study conducted by R. H. Clayton et al. (1993) frequency-based methods were used to sense Ventricular Fibrillation (VF). It is important to study self-terminating VF as it may provide information regarding the electrophysiological mechanism which helps us discern arrhythmias in an early stage. 10 s of the VF recording is taken. Then the 10 s of the signal is divided into 1 s epoch, where the initial epoch begins at the abnormality first encountered, of each VF. To counter the problem offset might pose in the analysis, the mean of each epoch is taken, and then subtracted from each point of data. The 250-point epoch is then converted to a 1024-point epoch, by adding zeros to either side of the signal. Hanning function is used to window the signal and then FFT (Fast Fourier Transform) is applied, to bring the signal to the frequency domain. Thus, an ECG spectrum is created with a spacing of 0.244 Hz between adjacent frequency components.

Attempt has been made, by the author to focus on dominant frequency peaks and the amplitude of the peaks. The frequency component having the most influence in the signal is called dominant frequency. The peak size could be defined as the fraction of the spectrum in 0.5-40 Hz bandwidth. The effect of the peak could be obtained by summing the square of the amplitude over a fixed peak width consisting of seven components of 1.46 Hz width which are centered around the dominant frequency. To get the peak size the author divides the square of the amplitude in the range 0.5 to 40 Hz. To see how the dominant frequency and peak size are affected in each successive epoch, Wilcoxon’s signed rank test has been used.

As the ECG signals are not stationary, the frequency analysis method fails to provide any useful data. Clayton and Murray (1998) devised five, time-frequency distribution techniques to study the signal. The methods used to generate spectrogram are:

- Sliding Fourier Transform (SP),
- Wigner Distribution (WD),
- Smoothed Wigner Distribution (SWD)
- Two variations of the Choi-Williams Distribution (CWD) given by CW5 and CW6.

Three parameters are extracted by using the above-mentioned algorithms which will help in comparing the performance of TFD (Time Frequency Domain) techniques.

2.4 Wavelet Transform

Wavelet transform (WT) has been used recently to classify signals automatically. WT uses a basis function called mother wavelet. It is a powerful and a new signal analysis technique. Mother wavelet is translated and scaled to give a set of wavelets. A basis function is selected based upon how much it looks like the ECG morphology. As the ECG signals are non-stationary in nature and have a wide bandwidth, this technique has been proven to be very efficient.

In a study by Brohet et. al. 1998, WT was used to detect atrial flutter in ECG signals. Of all the CVD (Cardio Vascular Diseases) encountered, detection of atrial flutter is most difficult. This is so because of the baseline drift present in the signal and complex nature of flutter waves.

As the ECG signals are not continuous when used in the machine/computer, Discrete Wavelet Transform (DWT) was used in this study. DWT uses convolution theorem on the data and two limited series coefficients of which one is real and other is complex. Their amplitude is given by the quadratic summation of the two parts. It is tedious to locate every flutter in the wave, however the parameters like mean heart rate and its variance could be used to identify them. The study concluded that although automation of diagnosis of ECG provided better results, the whole process is very complicated and takes more computational time.

In recent times, as the computers have become very powerful, the computational load is not a big problem, which has intensified research for studying biomedical signals. Neural networks are being used for pattern recognition. The advancement of computer technology in recent years has intensified study in biomedical signals. Therefore, machine learning is being used today to overcome problems faced by using traditional techniques.

Tumer et al. 1998 has been working on a method for automatic diagnosis of continuous waves. In this method, many fuzzy automatons were constructed which can identify an arrhythmia. When the diagnosis is under operation, the sampled measurements are presented to the automaton. This method helps in input processing from several perspectives while also is tolerant of some noise and other ambiguities. However, the automation of ECG signals is still a challenge due to associated noise and baseline wander.

Analysis of a system whose performance is categorized by continuous time and amplitude presents distinct problems such as the following:

1) Signals can be disturbed by noise and localized baseline wander.

2) Good and pathological signals may be problematic to distinguish in computerized systems that get easily muddled trying to manage noise, baseline wander, and local and global feature perturbations.

It is for these reasons that ordered high fuzzy automatons (HFA) is used for automatic diagnosis of ECG signals. HFA is fuzzy automatons that work on signals at the different level of details. Whenever we reach a higher level in the hierarchical result, a more complex and global structure is identified. At the highest level, the fuzzy automation can recognize a string which will represent an arrhythmia. The input given to the HFA is ECG signal in the time domain which is divided into primitive using Adaptive Resonance Theory 2 (ART2). An essential feature of HFA is its non-deterministic operation. This fuzzy automaton will help in the transition from a starting state to next potential state.

The advantage of using the ART2 construction is that this learning is independent of supervision and can identify forms in segments of the input signal. The three main take away from this study were:

1) Decision-theoretic approach is complemented with the Syntactic approach by using ANN (which constructs primitive alphabet).

2) HFA uses- state fuzziness and transition fuzziness. This improves the robustness of the fuzzy state machines with respect to state machines. Fuzziness gives flexibility to a state machine to make multiple transitions concurrently and provides the state machine with the competence of being at multiple states altogether.

3) As the input is synchronized in the current method, signals can be evaluated irrespective of the point at which the operator starts to present the input to the system. This prevents any prior modification of the original signal from a predetermined point.

Hierarchical architecture for intelligent structures is a new way in the AI (artificial intelligence) research which will help in the advance of the next generation of intelligent systems.

In another study done by Safara et al. 2013, wavelet transformation (WT) was used on the signals, to extract orthonormal bases/nodes and use these nodes for feature extraction. Multi-Level Basis Selection (MLBS), is used in this study to preserve the important information, and remove noise or unwanted data, by using some specific exclusion criteria. The exclusion criteria consist of three parameters, i.e. frequency range, noise frequency and energy threshold. The tests provided an accuracy of about 97% in classifying few cardiovascular conditions such as aortic regurgitation, mitral regurgitation, normal heart sounds and aortic stenosis.

In this study, the sounds and vibrations produced by the opening and closing of the heart are used to detect the diseases present. A phonocardiographic (PCG) reading of the heart signal is taken which is simply the sound and vibrational levels of the heart. Then these PCG readings, which represent the heart murmurs are used to classify heart problems. The sounds produced by the heart (valves opening and closing) are the low frequency in nature, whereas the murmurs present in the heart (due to turbulence in blood etc.) are of high frequency in nature.

Thus, the main purpose of this study was co-relate the heart sounds with the existing or forthcoming diseases in the heart. The procedure followed is to decompose the signal by taking its wavelet transform. The orthonormal nodes after WT, are segregated based on two techniques- 1. Best Basis Selection (BBS) and 2. Local Discriminant Basis (LDB). Thus, the nodes having less information are removed. The flowchart below shows the process.

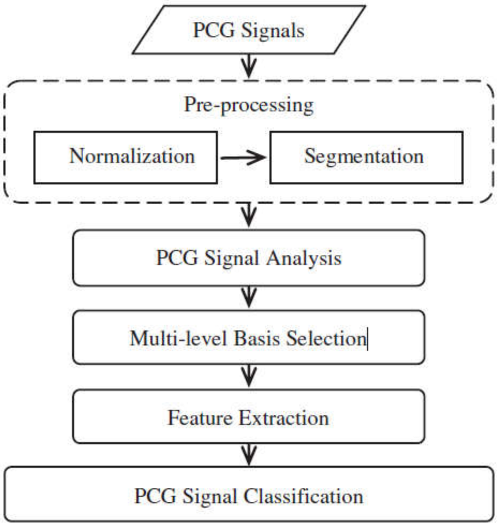


Figure 2.1 Flow chart of Phonocardiographic (PCG) method (Safara 2013)

Coefficients of the apt nodes are used to excerpt features from wavelet decomposition of the PCG signal. Nodes are selected on the foundation of the minimum cost function. Classification is done by using SVM. Kernel functions are used t-map the input feature space to a higher dimension. This is done to find the best hyperplane which provides the largest margin.

In this study, two variables are used to gauge the effectiveness of SVM which are regularization parameter and slack variable. Slack variables (ξ) help in determining the amount of mis classification and regularization parameter (C) controls the hyper-plane margins. In this paper radial basis function (RBF) uses kernel function with ξ=0.001 and C=150.59 heart sounds were studied which had previously been examined by a cardiologist. A feature vector was made for the PCG readings and later SVM was used to classify the feature vector. It was thus concluded that multilevel basis selection (MLBS) was a good basis selection method having an average accuracy of 97%.

In another study by Sabiq and Faziludeen (2013), classification of three types of beats was achieved i.e. bundle branch block, normal sinus rhythm, premature ventricular contraction. QRS detection was done using Pan Tomkins Method, and then wavelet transform (WT) was applied using Daubechies 4 wavelet. Then features such as mean value, standard and variance, approximate coefficient and minimum and maximum coefficient, R-R interval features. After this three SVM's were created to classify the signal, which gave an average accuracy of about 98.5%.

The method for discrete wavelet transform (DWT) is shown below

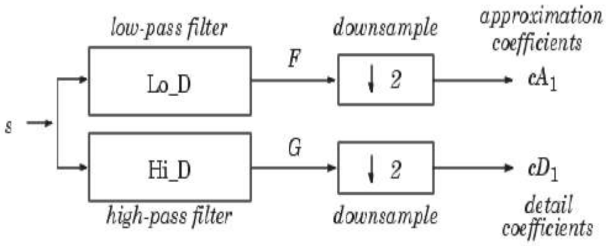


Figure 2.2 Wavelet Decomposition (Sabiq and Faziludeen, 2013)

Thus, the signal is segregated based upon approximate coefficients and detailed coefficients. Order 4 of Daubechies wavelet was used to decompose the signal. As SVM is a binary classifier, thus to use SVM for multi class classification, a One Against One (OAO) or pairwise coupling was used for each pair. The kernel used was the linear kernel.

Anuradha et al. 2008, used ANN to classify signals based on four parameters i.e. Spectral Entropy, Poincare plot geometry, largest Lyapunov exponent and detrended fluctuation analysis. Entropy in spectrum of a signal is a measure of its spectral power distribution. It is given by

$$H = \sum_{f} p_f \log(\frac{1}{p_f}) \qquad 2.2$$

Where p_f is the probability density function value at frequency f. The spectral entropy, H lies from 0-1 and describes the complication of the Heart Rate Variability signal. (Woo et al. 1992).

In Poincare plot, an oscillator model is used to modulate low and high-frequency components of the sinus node into R-R intervals. This helps provide a connection between time domain Poincare geometry analysis and frequency domain spectral analysis (Kaman et al. 1996).

Largest Lyapunov exponent helps in measuring the chaos in the dynamic system by finding out the largest Lyapunov exponent. This method examines for the nearest neighbor of each point in space and assesses their separation with respect to time.

Detrended fluctuation analysis is used to define the scaling properties of short R-R intervals. This method resembles the root mean square variation of detrended and integrated time series.

$$F(n) = \sqrt{1/N \sum_{k=1}^{N} [y(k) - y_n(k)]^2}. \qquad 2.3$$

Where y(k) represents a signal of length n, F(n) represents the integrated time series.

After all the four parameters have been calculated for the complete database, WT was used to decompose the signal. The decomposition was done at different resolutions to accurately detect non-stationary features of ECG signals. Discrete wavelet transform coefficients were obtained which contained information regarding the arrhythmias.

ANN (Artificial Neural Network) was then used to classify the signals with the help of backpropagation algorithm. The parameters of ANN that were used i.e. four layered feedforward network, sigmoid was used as the activation function and the stochastic gradient descent was used to update the weights. The study showed an accuracy of nearly 90% for classification of arrhythmias.

Nurul et al. 2016 implemented a method to extract ECG features based on a system that uses second order differential equation which helps to represent the short-term behavior of ECG signals. This representation depended upon natural frequency (ω), damping coefficient (ξ) and forcing input (μ).

Windowing of the signal was done into 2,3,4,6,8,10 seconds to find the correct window size for the signal. ANOVA and T-test were done to find out the useful features. Also, machine learning methods like ANN and SVM were used with k fold cross-validation to develop an ECG classification system. The conclusion of the study was that 4-second window is the bet windowing length and the two features used were- Natural frequency (ω) and Forcing input (μ) with a detection accuracy of about 95%.

Acharya et al. 2016 focused on developing a computer rated method for ECG signal processing so that myocardial infarction can be automatically detected in clinics. First, the signal was decomposed till four levels. Some features were extracted from this decomposed signal:

- Approximate entropy
- Signal energy
- Fuzzy entropy
- Kolmogorov–Sinai entropy
- Permutation entropy
- Renyi entropy
- Shannon entropy
- Tsallis entropy
- Wavelet entropy
- Fractal dimension
- Kolmogorov complexity
- Largest Lyapunov exponent

The ranking of extracted features is done based on the T value. K nearest neighbor (KNN classifier) was used for the classification.

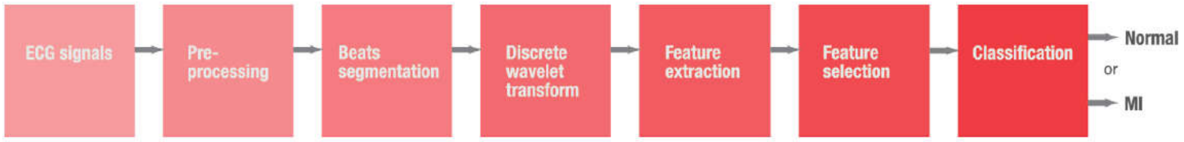


Figure 2.3 Steps followed in using KNN classifiers for feature extraction (Acharya et al. 2016)

The procedure followed is depicted in the above shown figure 2.3. First preprocessing of the signal was done so that the sampling rate of all the signals is same. The signal was thus digitized at 1000 samples/ second. DB6 is used to remove baseline wander and any noise.

R peak detection was then done by Pan Tonkin's algorithm. R peaks are then identified as they are the easiest to locate. Once detected, 250 samples to the left and 400 samples to the right were selected for each EG beat. The next step was to extract features, and this was done by subjecting the signal to four levels of DB6 wavelet function. The twelve features mentioned above were extracted from 8 DWT coefficients. The authors faced problem in the difficulty in finding the features necessary for MI detection. This was solved by using ANOVA and T-test which rank the features according to their significance. KNN classifier was then used to classify the features previously ranked and to find the minimum number of features required for efficient performance and result. This study achieved an accuracy of up to 98%.

2.5 Machine Learning (ANN and SVM)

Patidar et al. 2015 studied that Tunable Q Wavelet Transform (TQWT) was used to detect coronary artery disease (CAD). CAD is the thinning of arteries of the heart which causes a decrease in the amount of nutrients and oxygen to the muscles of the heart. TQWT was used to decompose the signal and then features were extracted from the signal. The decomposed heart signals were then divided into various subbands for better classification. A nonlinear feature called Centered Core entropy (CC) is calculated from the subband and then Principal Component Analysis (PCA) was done on CC which helped in transforming the number of features. The features which were found were then used in automatic diagnosis with help of SVM classification. Morlet wavelet kernel function was used.

Preprocessing of the signal consisted of:

- Removing high-frequency noise
- Power line interference (nearly 50 Hz)
- Removal of baseline wander

To achieve this, bandpass filters with cut off frequency of 0.3 Hz and 50 Hz was used. A notch filter was used to remove interference due to the power line. Pan Tonkin's algorithm was used to detect R peaks in the QRS complex. Next step was TQWT decomposition. This used three variables which can be easily adjusted to facilitate analysis of the signal.

These three parameters were:

- Q wave
- Total over sampling rate (r)
- Number of the level of decomposition (j)

This was followed by feature extraction. Directly heart rate signals were not used as features because it led to more computational load and more space requirement. The features were calculated from Tunable Q Wavelet Transform (TQWT) based decomposition. Feature transformation was done by a technique called PCA which provided more descriptive power to the original feature. A new feature vector was created by the linear grouping of the original features. Principle components have a characteristic that they are orthogonal to each other so they don't contain any unwanted information. Covariance of the raw features was calculated. Eigen values and Eigen vectors are later calculated from the covariance matrix. Arrangement of the vectors is in descending manner of their Eigen values.

Support Vector Machine (SVM) was then used as the classification technique, using least square to convert the problem from quadratic to linear equations, which thus give faster and better results. MATLAB codes and toolbox were used for computation of various parameters. Values of CC were obtained after using third level wavelet coefficients. The testing and training of SVM are done using cross-validation method. The parameters used to help in regularization are tuned, keeping in mind the cross-validation score. In the study, they found out that SVM classifiers performed better than most other classifiers.

In another study done by, Silipo et al. 2005 ANN having different parameters were compared with one another. Another area which was focused upon was to see how much, pre-processing of the signal affected the performance on ANN. The study wanted to compare what benefits ANN had over other techniques or whether those methods can be compared with ANN at all.

The study also says that pre-processing of the signal is important as it affects the quality of classification. Data were taken from three different sources:

- MIT-BIH database
- European Society of Cardiology/ VALE database
- University of Leuven

The results of the study showed that ANN performed much better when it comes to classifying arrhythmias. This can be attributed to the fact that they have better non-linear separation surfaces. This was better than clustering methods which logic rules took more time.

When it came to ischemia detection, ANN did not perform better than other methods. Two ANN approaches were used i.e. static and recurrent, but it was found that static was less time-consuming. Comparison between threshold algorithm and ANN showed that ANN was much better. This was because threshold technique, utilizes the previous QRS segment and thus its quality depends upon that, but in the case of ANN, it is independent of previous QRS detection.

Mporas et al. 2015, worked upon the detection of seizures using EEG and ECG signals. Here time domain, in combination with frequency domain features were used to study and classify signals. Different kinds of signals like ECG, EEG, EMG, etc. were monitored and stored.

The evaluation of seizures is usually done by looking at the EEG signal. But as the EEG signals are erratic in nature, it becomes very time consuming and difficult to study them, especially for long records. Also, there are different types of medical and physiological patterns, which only add to the problem. ECG signals were recorded because it is known that seizures are related to variations in respiratory and cardiovascular systems. Thus, a collective study of ECG and EEG signals could help in automatic diagnostics rather than relying upon visual parameters.

In this study first preprocessing of ECG and EEG signals was done by dividing them into time segments, called epoch and then filtering of these epochs was done. Next stage was feature extraction where a feature vector was extracted from each time segment. During the classification, each time segment and feature vector was labeled as having seizure or non-seizure. During the preprocessing stage, frequency domain, time domain, and time-frequency domain parameters were extracted from the signal. Time domain and frequency domain analysis were performed on EEG signal and time-frequency domain signal was performed on ECG signal. Features in time domain extracted were:

- Minimum value
- Maximum value
- Mean-variance
- Standard deviation
- Percentile
- Mean absolute deviation
- Range
- Skewness
- Kurtosis
- Shannon entropy
- Number of positive and negative peaks
- Zero crossing rate

The features extracted in the frequency domain were:

- Power spectral density
- Autoregressive filter coefficients
- Spectral entropy

The parametrization of ECG signal was done based on the features

- Heart rate absolute value
- Heart rate variability
- Statistical Values
- Mean absolute deviation

Data labeled by medical experts was used for training purposes. The labeling was done on the basis of seizure versus non-seizure. To avoid overlap between testing and training data, tenfold cross-validation method was used. The accuracy of detection of the seizure varied from 77% to 92%.

This variance inaccuracy was attributed to less amount of training data present. The study concluded that EEG signals on their own did not provide good results but in combination with ECG signals they provided good results.

In a paper by Moraes et al. 2013, Artificial Neural Network (ANN) and Support Vector Machine (SVM) techniques have been compared. The authors found out that ANN performed better than SVM in most of the cases except when unbalanced data was involved. It was also found out that to reduce the training time, the number of input nodes has to be reduced. The basic functioning of ANN and SVM is very similar as both give an output based on some linear combination of some input. A bias term is present in both the techniques. The difference in two models occurs in how the solution is obtained.

Support vectors used resultant of optimization and these vectors were always a subset of data in SVM. This was not the case for ANN where the number of hidden nodes is a free parameter that can be altered. In contrast to ANN, SVM had a characteristic of selecting their own model size. Sometimes a large number of vectors were needed to give the output. This resulted in slow time in real life uses.

ANN also faces a problem where the number of hidden nodes must be small to reduce the complexity at the expense of accuracy. The advantage of SVM over ANN is that SVM uses an optimization approach by finding a global minimum and a unique solution. However, in the case of ANN training is done by stochastic gradient method which may not converge to an optimal solution.

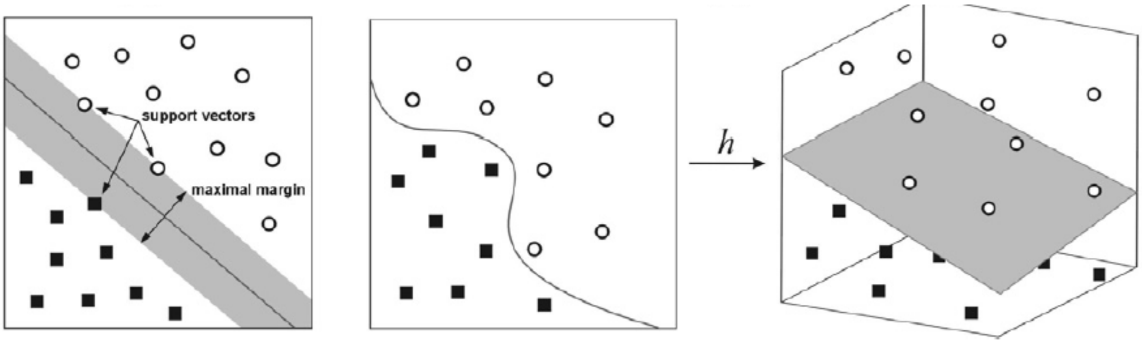


Figure 2.4 Kernel Trick used to convert to a higher space (Moraes et al. 2013)

The focus of this study was also to determine how to get the best classification and what critical parameters in the process. The parameters taken into consideration was the regularization parameter and in the case of ANN was the number of neurons in the hidden nodes. In this study, a radial basis function was used as a non-linear kernel. In the case of ANN, a single hidden layer was used in the feed forward network. The total number of neurons was varied from 15-55 in the hidden layer. This was done to overcome non-convergence.

Another technique used was to randomize initial weights of the node and repeat training. Traditional gradient descent method was not used and instead scale conjugate method was used to improve the speed of convergence of the solution.

Based on the existing literature review it can be concluded that ANN was less biased than SVM. ANN always gave a better result than SVM, but the difference was never more than three percent. It was also seen that ANN was more sensitive to the noisy and unbalanced data. The results indicated that the standard feature selection did well to refine data and thus reduce the calculation load in the training process. Training time of the ANN was poorer than that of SVM, but it had greater accuracy.

Moavenian and Khorrami (2010) compared multi-layered percepton (MLP) of a neural network with an improved version of SVM using Kernel Adatron (KA) technique. Signals were used to train from MIT- BIH arrhythmia databank which helped them to classify six dissimilar types of arrhythmias. In the experiment initially, only one ECG lead signal was used. In the second part of the experiment, another ECG lead was used to test and train the dataset. This was done to find out the influence on testing and training performance and to see which classifier used less time. The results showed that the modified SVM algorithm was faster in training and also showed higher performance but the generalization in MLP was three times better than SVM.

The KA training algorithm in SVM needs less training input and do not use local minimum points. This was the reason they took less time than ANN.

The comparison between the two techniques is done based on three parameters:

- Training performance
- Testing performance
- And training time

In case of ANN three layers of feed-forward neural network were used and training was done with the help of generalized delta learning rule to minimize the cost function. SVM maps the training data to a higher dimensional space. This was done to find the maximum margin hyperplane which is used to separate the data. The procedure to find the hyper plane is to use a quadratic technique which is computationally intensive and complex. The KA algorithm can help in finding a solution more quickly in less number of iterations.

The result of the study was that modified SVM should be used when training time required is the main focus. Otherwise, MLP should be used when performance is concerned. Also, adding the second ECG lead to increase the data set did not yield better results with only showing 33%incrrease in performance in case of ANN and only 7% increase in performance using SVM.

2.6 Summary

Based on the existing literature review it can be said that automatic classification is important for appropriate diagnosis of heart. The earlier researchers attempted to used time-based techniques for signal classification, however these techniques missed the important information and it was realized that frequency analysis is also important to extract useful features. So, frequency based analysis could provide the large set of features which were missed by time domain analysis however it was found that most of the bio-medical signals were non-stationary in nature, thus due to dynamic nature of signals real time analysis was found to be difficult using frequency domain Short Time Fourier Transform (STFT) and further using wavelet technique (WT) which uses variable window size for better resolution. In the recent time the two machine learning techniques are gaining importance such as Support Vector Machine (SVM) and Artificial Neural Network (ANN).

It was found that the accuracy of the various models used to extract information from heart signals was limited to 90-93 %. In order to further improve the accuracy of the models, critical investigation is required in the model development based on various techniques. In view, of this in the present work, attempt has been made to first obtain the peaks using FFT (Fast Fourier Transform) and then these peaks were used as input to Artificial Neural Network (ANN). It is also important to find the input parameters which affect the accuracy of the model. Hence parameter optimization has also been performed in the present work.

Chapter 3

Feature Extraction and R peak Detection of ECG Signal

3.1 Introduction:

ECG signals consist of a lot of information which would be beneficial in studying the signal. The first step is to filter the signals and then find the R peaks in the signal. R peaks are necessary to detect because they also help in the detection of other peaks such as Q peaks, T peaks etc. They also help in finding the R-R intervals of the signal, so that the location of peaks can be known, and later these help in classification of signals, while using ANN and SVM. R-peaks are the first to be detected, because they have the highest amplitudes and are the easiest to detect. Huge amounts of variation in the signal make the detection of QRS segments very complex (Thakor et al. 1984) Thus, it is important to detect R-peaks.

Most of the times, an ECG signal contains noise due to different sources, like, muscle noise, power line noise (around 50Hz), electrode dislocation, baseline wander etc. This makes, the detection of R peaks very difficult (Yeha ad Wang, 2008). Pan and Tompkins (Pan and Tonkins, 1985) algorithm is also a very widely used technique for QRS detection. In this, detection of the QRS complex, was achieved by filtering the signal, then applying a non-linear transformation and then finally, using a decision-based rule algorithm, to detect the R peaks.

A study performed by Pan and Tonkins (1985) showed that by using differential operation, it was possible to find the QRS complex after finding the R-peaks. Another method was used by Afonso et al. (1999), which used filter banks and multi rate signal processing, to detect the QRS complex. Hilbert Transform was used by Beniteza et al. (2001), to identify the R peaks, after differentiating the ECG Signal.

Fast Fourier Transform:

It is known that, Fourier series is a method, to write any function as a sum of sine and cosine functions with increasing frequency. This means that a time varying data can be converted to a frequency domain. This idea forms the basis of digital signal processing(DSP), and is useful because frequency domain reveals much more information, hidden in time domain.

Fourier transform of any function f(x) is given as:

$$F(u) = \int_{\infty}^{-\infty} f(x)e^{-i2\pi ux}dx, \tag{3.1}$$

Where i=√-1 and *u* is the frequency variable. Applying Euler's equation:

$$F(u) = \int_{\infty}^{-\infty} f(x)(\cos 2\pi ux - i\sin 2\pi ux)dx. \tag{3.2}$$

The magnitude of F(u) is defined as the Fourier Spectrum of f(x) and θ(x) is called the phase. The square of $|F(u)|^2$ is denoted by P(u) or the Power Spectrum of f(x). The Fourier spectrum is often plotted against the values of u.

3.2 Discrete Fourier Transform:

It was a method developed to perform Fourier transform on discrete signals, because the signals read by the computer are never continuous. Consider N number of discrete samples of f(x), sampled uniformly:

$$F(u) = \frac{1}{N}\sum_{x=0}^{N-1} f(x)e^{-i2\pi ux/N} \tag{3.3}$$

where, u lies from [0, N-1].

$$f(x) = \sum_{u=0}^{N-1} F(u)e^{i2\pi ux/N} \tag{3.4}$$

where, x lies from [0, N-1].

3.3 Fast Fourier Transform:

Discrete Fourier transform helps in the transformation on the computer, but its computational load is very large. The calculation required for (3) and (4) are proportional to N^2. The FFT algorithm helps decomposes the DFT into the $\log_2 N$ stages.

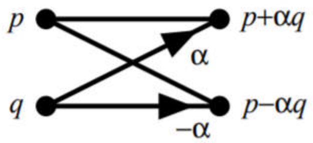


Figure 3.1 Butterfly Diagram

Each of these consist of N/2 butterfly calculations. Each butterfly takes p and q, two complex numbers, and calculates two number, p + αq and p − αq, from them. Where α is a complex number. Below is a diagram of a butterfly operation.

3.4 Methodology:

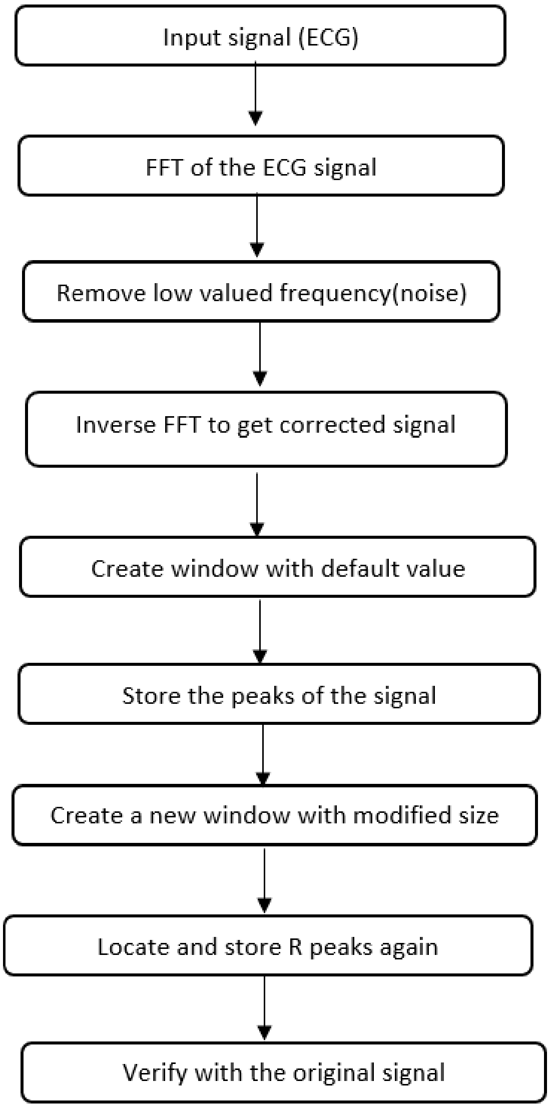


Figure 3.2 Flow chart for R-peak detection

3.5 Filtering:

Filtering refers to remove the power line interference noise from the signal. This also helps to remove baseline wander from the signal. To do this, the signal is first transformed from time domain to frequency domain. Then, the low frequency noise from the signal, up to 50 Hz is set to zero. After this, inverse FFT transform is applied, to retrieve the original signal. The results are shown in Figure 3.3 and 3.4.

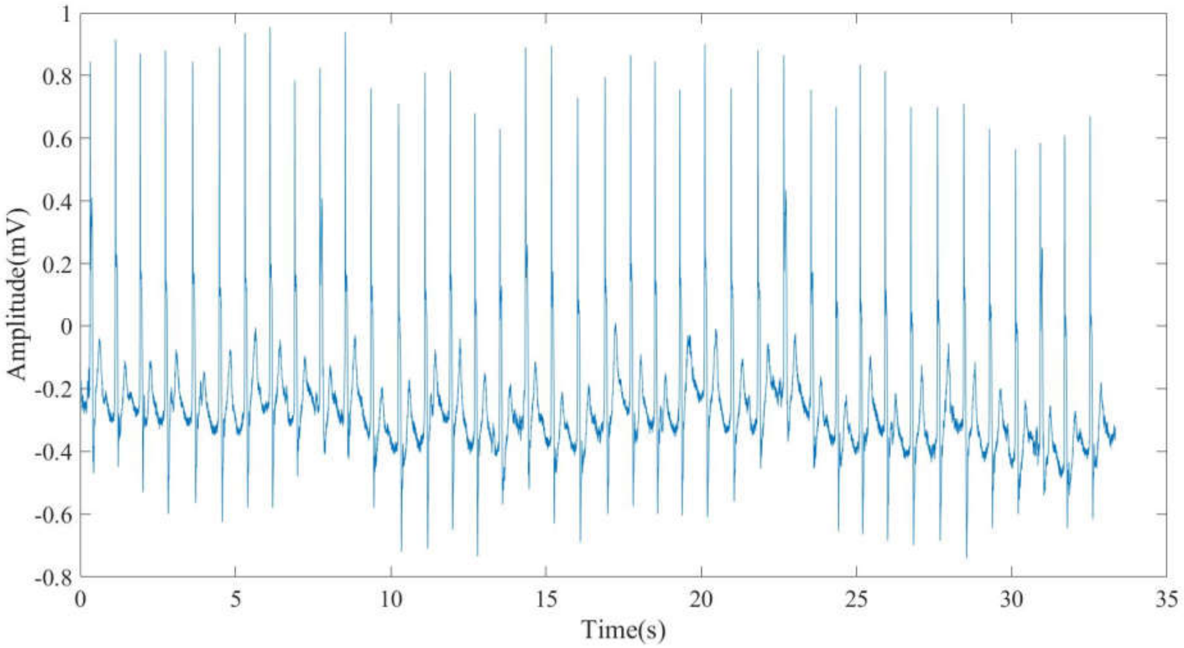


Figure 3.3 Original ECG Signal

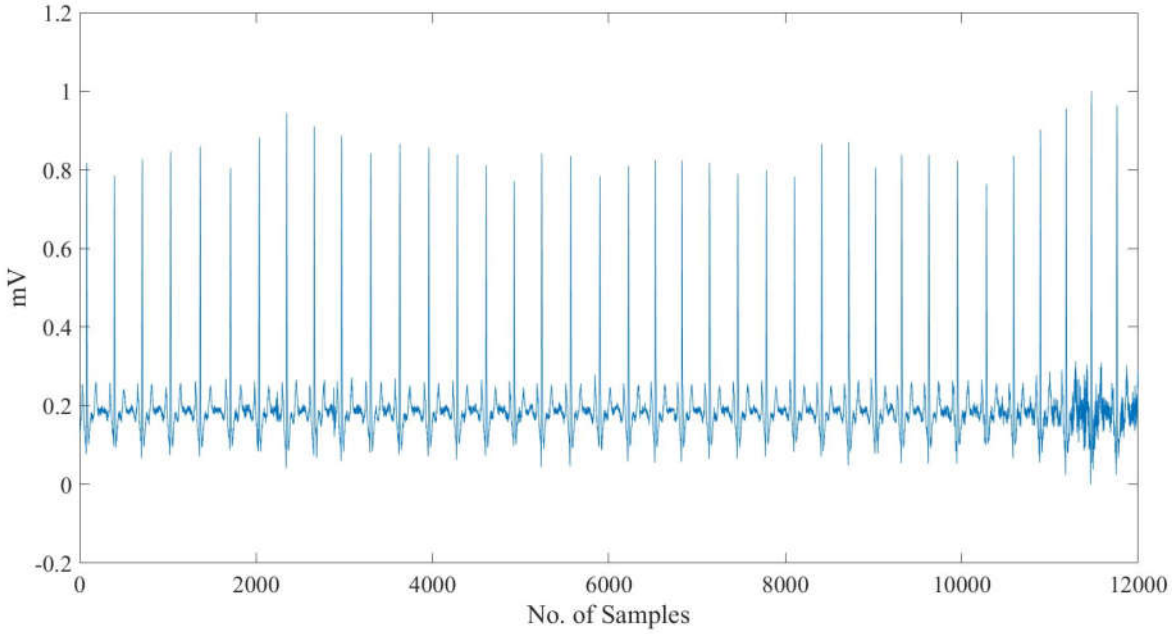


Figure 3.4 FFT Filtered Signal(Baseline wander is removed)

3.6 Windowing:

In this stage, a window is used to find the location of R-peaks. The main function of the window is to divide the signal into small segments, of a defined size. For the first iteration, this window size is set as default window size, provided by the MATLAB function. In that window, the position and the amplitude of the point with the highest peak is stored.

Initial window size is calculated as:

$$\text{Initial size} = \text{Sampling Rate} * \frac{570}{1000}. \qquad 3.5$$

Then, the location of peaks, and their amplitudes are stored. The results are shown below:

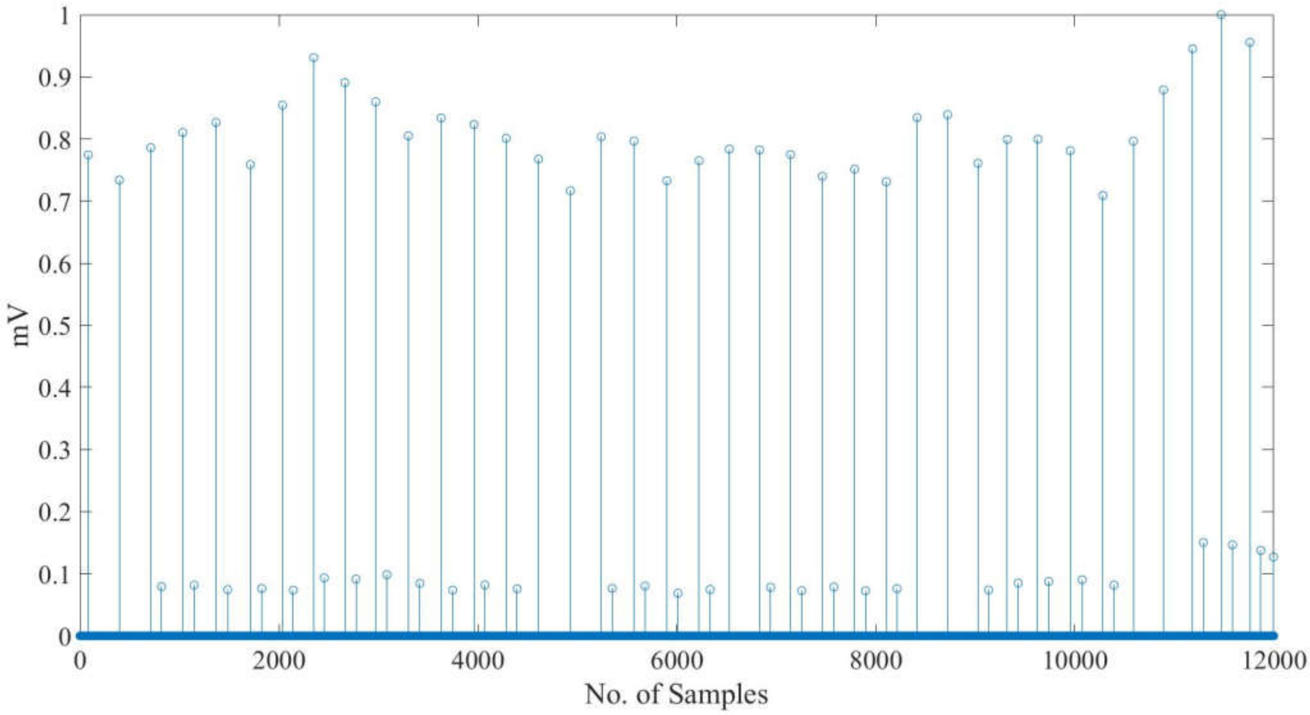


Figure 3.5 Filtered ECG- First Pass

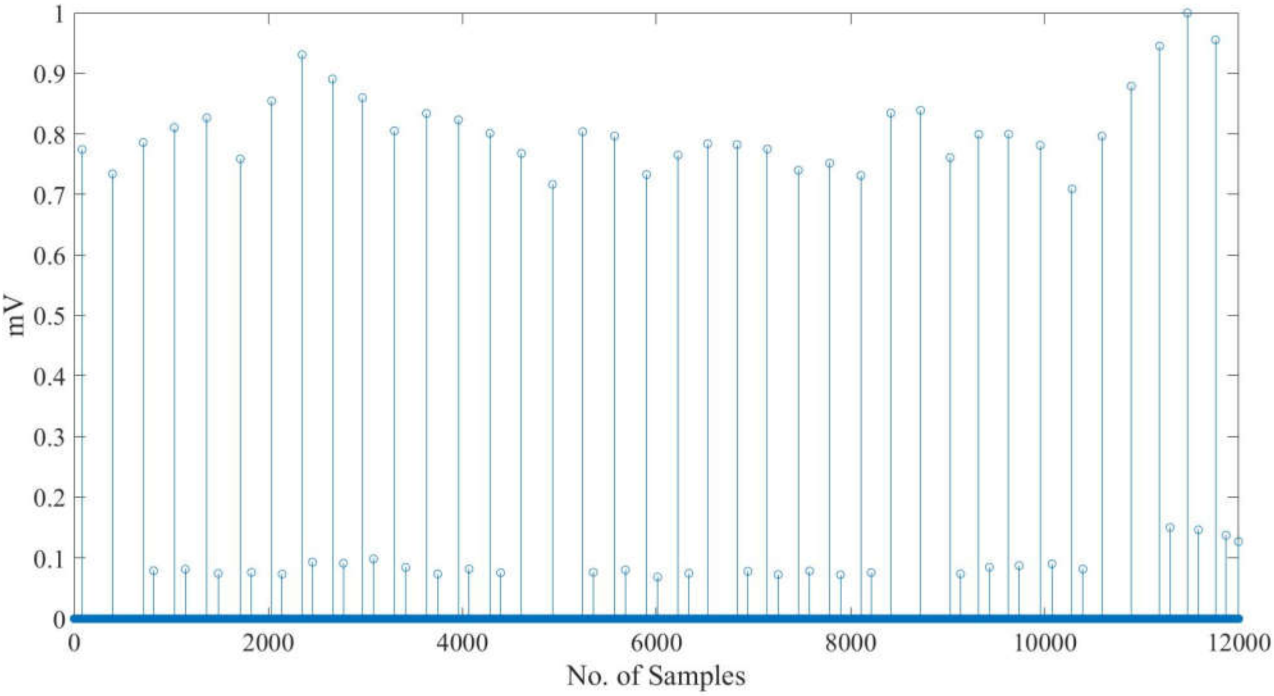


Figure 3.6 Peaks Detected

As we can see in the above diagrams, that most of the R-peaks have been detected, but also some extra peaks are also included, as the window size was not correct.

Next, modified window size is found out by using the information about R-peaks, found in the previous section. To do this, we calculate the minimum distance between two peaks, by using a for loop in MATLAB.

Q-R distance is calculated as:

Q-R distance= 0.04*sampling rate 3.6

New window size= (2*Minimum peak distance) - (Q-R distance) 3.7

The same procedure is followed, where the signal is sent to the window function, with a new window size. The results are as shown below:

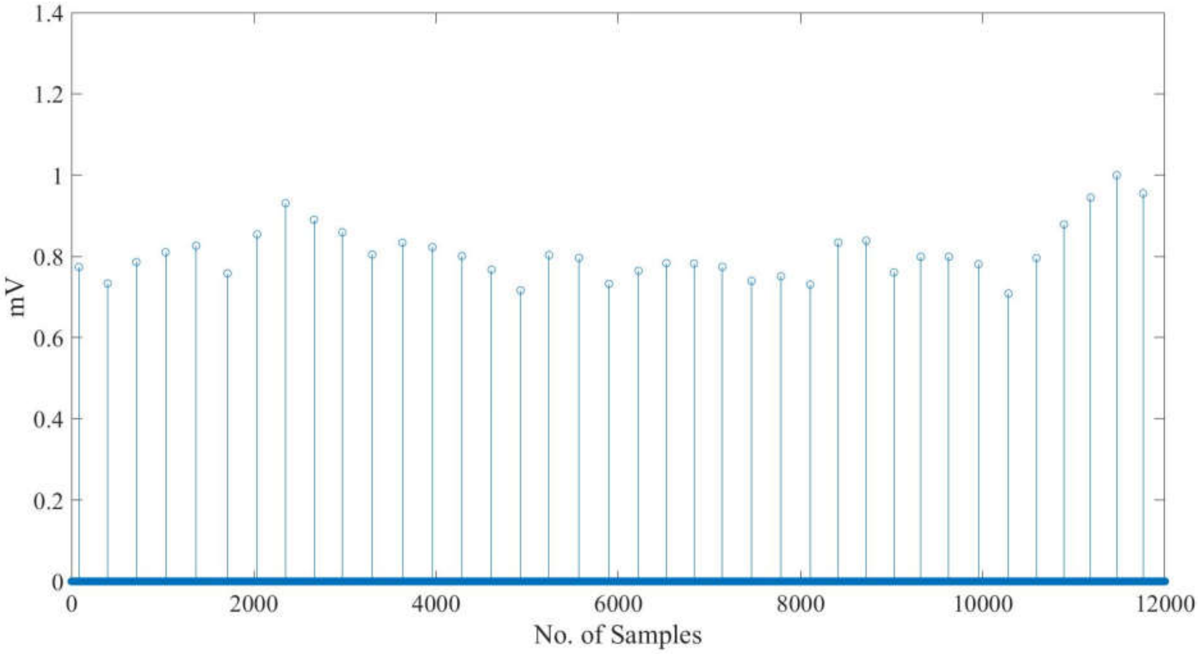


Figure 3.7 Filtered ECG- Second Pass

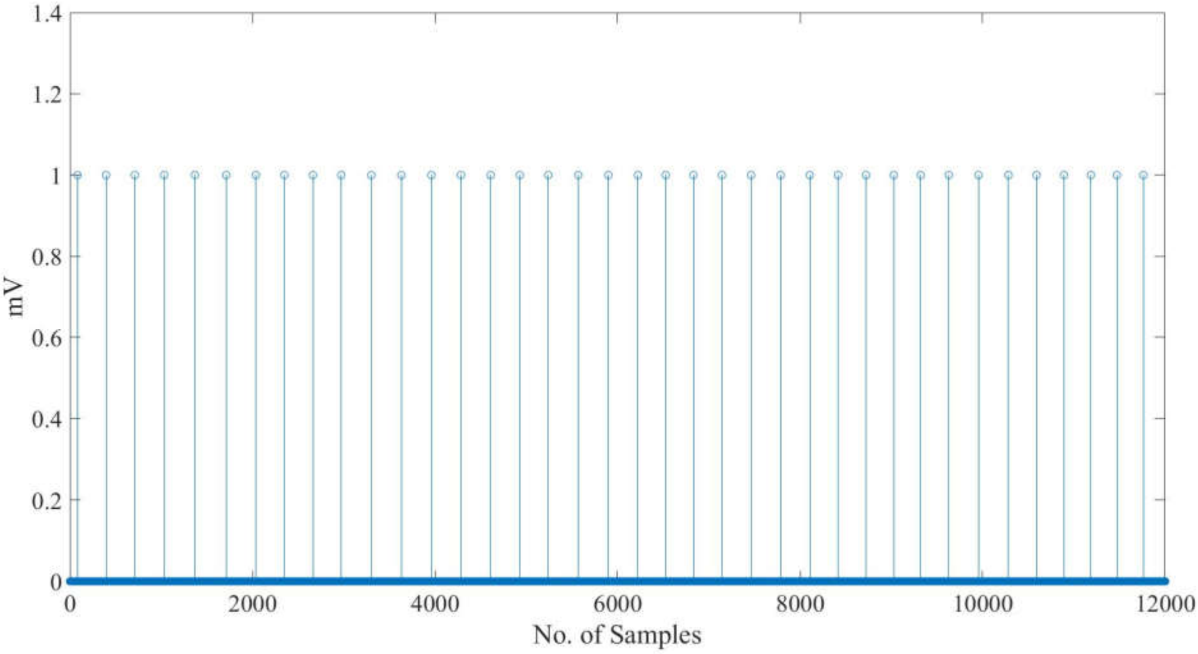


Figure 3.8 Detected Peaks-Finally

Lastly, we verify the detected peaks on the original signal by overlapping the peaks over the original signal as shown in the Figure 3.9.

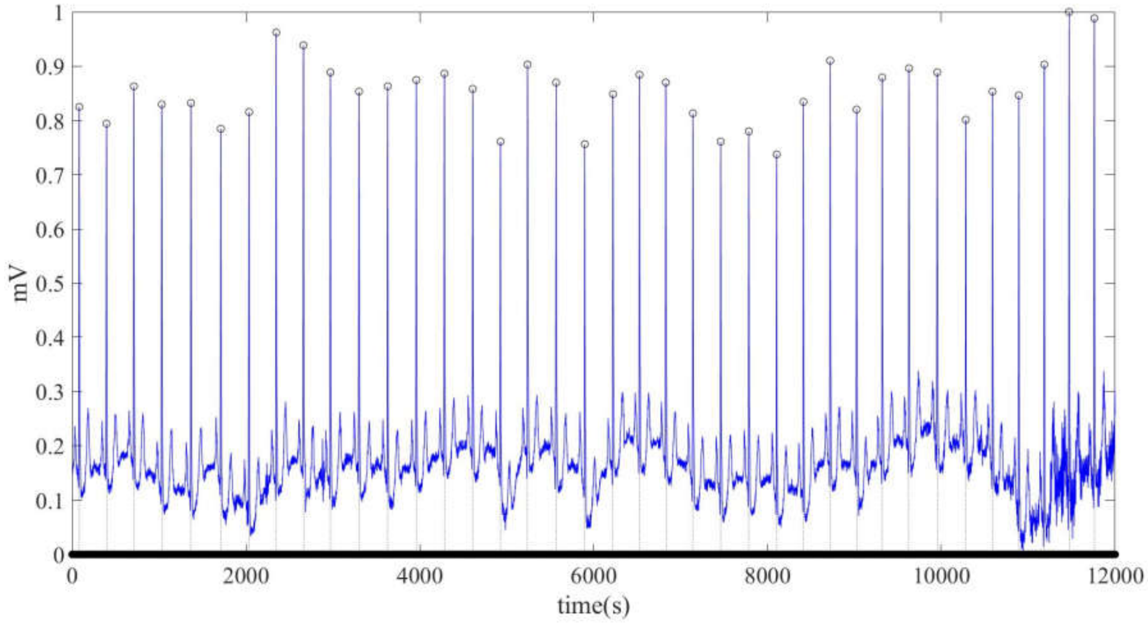


Figure 3.9 Comparitive ECG Peak detection Plot

As can be seen in the Figure 3.9 that almost all the peaks present in the original signal have been detected. But as we can see in figure 3.10, not all R peaks were detected so we did not pursue this algorithm further for pattern classification because it failed for irregular and noisy ECG signals.

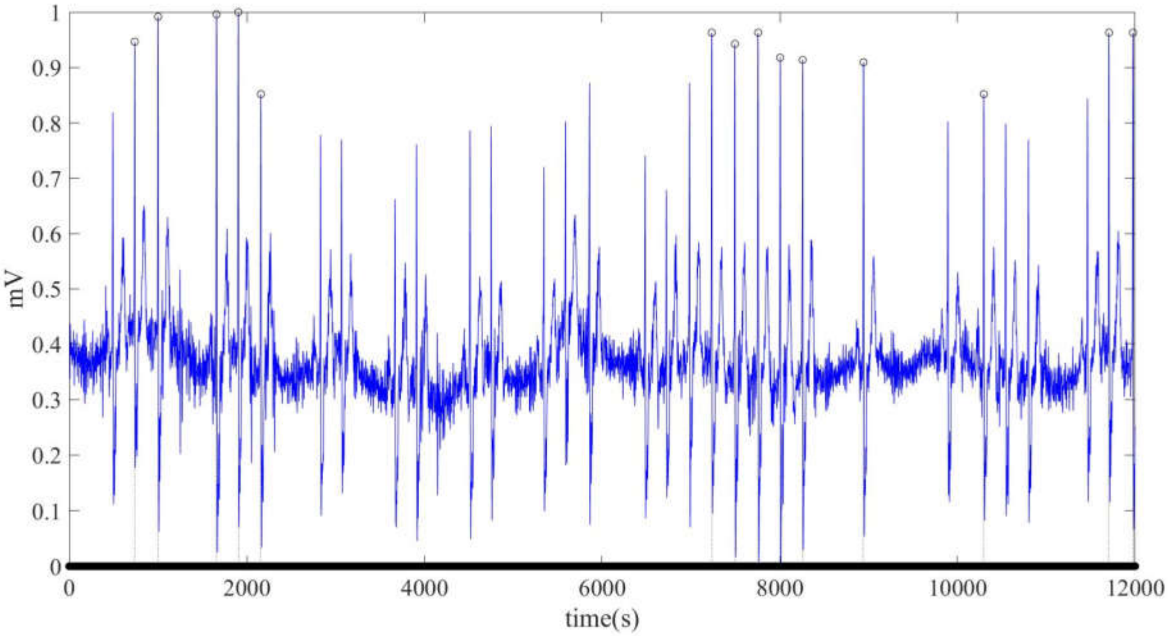


Figure 3.10 Comparitive ECG Peak detection Plot for noisy and irregular signal

Chapter 4

Wavelet Analysis

Conventional methods of ECG signal analysis consider the signal to be stationary in nature, but that is not the case (Sedjic et al. 2009). Most of the biomedical signals, are in general, non-stationary in nature, that is, they dynamically change their behaviour over time (Mendez et al. 2009). Thus, we cannot simply analyse the ECG signal, only in the time domain or in the frequency domain alone (Sedjic et al. 2009). Time domain provides us information regarding the amplitudes and location of signal peaks but is susceptible to noise and baseline drift. It also does not provide any frequency content present in the signal. Frequency domain provides information on the spectral contents of the signal, but it is silent on where these frequencies are located on timeline. Hence techniques which counter these problems were developed, and these were called Time-Frequency Analysis techniques.

Time-Frequency Analysis (TFA), attempts to provide a spectral picture of the signal, as a function of time. Afonso et al. 1995 used a technique to classify arrythmia that can differentiate shockable heart rhythms from non-shockable beats. This technique worked on TFA, specifically STFT to calculate the energy distribution of the signal. Features were then taken out and used to classify the signals. The main problem encountered was that the resolution of STFT was not good, or sudden variations in the signal could not be tracked. In an attempt to keep the time window short, the frequency resolution suffered and vice-versa.

Wavelet analysis was found to be an alternative technique, as it has good time-frequency localisation property Dokur et al. (1999). A comparison between discrete wavelet transform (DWT) and Fourier transform was performed, to classify beats and the former was found to be better. Christov et al. (1992) used two TFA methods for classification of heart signals. Matching Pursuits (MP) method and QRS pattern detection method. The former was used to extract time frequency-based features, while the latter was used to collect morphological features. They found that MP method was best suited for passed and normal beats, while bizarre variations could be recognized by morphological analysis.

4.1 Introduction:

To extract data from the raw signal transformations are applied. Almost all the signals are in time domain, which do not provide the complete information. To a large extent, useful information is hidden in the frequency domain (which provides information regarding all frequencies present in the signal). For this Fourier transform is used which gives a frequency-amplitude representation of the signal. Since the ECG signals are non-stationary in nature and frequencies might change with time, therefore, either time or frequency information can be accessed at one time, not both.

Wavelet analysis is a multi-resolution analysis (MRA) method. This means that one can analyse the signal at different frequencies with different resolution. Here the spectral components are not resolved equally. MRA attempts to generate a decent time resolution and poor frequency resolution at high frequencies and vice versa. This technique is intuitive because mostly high frequency components are present for a short duration and vice versa.

Continuous Wavelet Transform(CWT) of a signal x(t) is given as:

$$CWT(a,b) = \int_{-\infty}^{\infty} x(t)\frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right)dt \qquad 4.1$$

Ψ- wavelet function

a- scaling parameter

b- shifting parameter

Here, the resultant is a function of two parameters a and b also known as scaling and translation parameters, respectively. Ψ(t), is also denoted as the mother wavelet. Wavelet is a small wave of a limited amount of length. It is called a wave because the function is oscillatory in nature. Mother wavelet refers to the main, original function which can then later be scaled and translated to give, other functions.

Scale: High scaling parameter means that we get a non-detailed global view of the signal, as is done with low frequencies, while to get a detailed view, we use a low scale, like for higher frequencies. Practically, low scale or high frequencies do not occur for the entire duration of the signal, while the high scale or low frequency is present for the entire duration. Basically, scaling is used to dilate or compress the signal. Larger scale will stretch a signal out (a>1) and vice versa.

The mother wavelet is selected, as a sample, for the windowing procedure. All the windows are scaled and shifted variations of the mother wavelet. Once the mother wavelet is selected, calculation stars with a=1 and the transform is calculated for a less than and greater than 1.

Wavelet is placed at the beginning of the signal starting at t=0. For a=1, eq. 4.1 is calculated. Then b is changed, and this wavelet is made to go over the entire signal. Next, the value of a is increased by some amount and thus the whole process is repeated. By shifting the wavelet in time axis, the signal becomes localized in time whereas by varying the value of a, the signal becomes localized in frequency.

If the signal contains spectral component that look like the value of a, the product of the wavelet and the ECG signal at the site where this spectral element exists will give a large value.

4.2 Discrete Wavelet Transform:

Implementation of continuous wavelet transform on the computers is possible, it provides us with a lot of redundant information and takes a lot of time. DWT is much easier to process. Here, a time-scale depiction of the signal is obtained with the help of digital filtering techniques. CWT is a kind of correlation amongst a wavelet of different scales and the signal with the scale (or the frequency) being used to measure the amount of resemblance.

CWT is calculated by varying the scale of the analysis window via varying the window in time, multiplying by the original signal, and then integrating for all times axis. In the DWT, filters of variable cut-off frequencies will be used to analyse the data at varying scales. The signal is made to go through a series of high pass filters to work on the high frequencies, and through low pass filters to analyse the lower frequencies.

The resolution of the signal, which tells us the amount of detail information present in the signal, can be varied by filtering, and the scale can be changed by up-sampling and down-sampling (subsampling) operations.

Subsampling of a signal by a factor n lessens the number of models in the signal n times. Up-sampling means to increase the sampling rate of the signal, via addition of new samples to the signal.

$$DWT(j,k) = \frac{1}{\sqrt{|2^j|}} \int_{-\infty}^{\infty} x(t)\psi\left(\frac{t-2^j k}{2^j}\right)dt, \qquad j,k \in Z \qquad 4.2$$

Where $a=2^j$ and $b=k*2^j$

This method is called the decomposition of the signal as shown below.

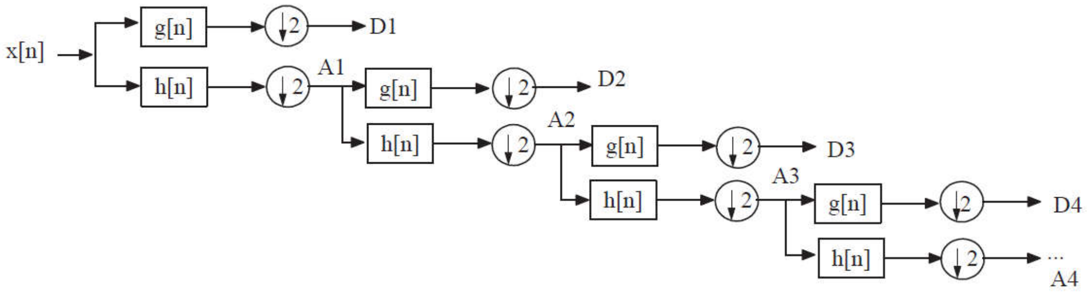


Figure 4.1 Sub-band decomposition by Wavelet Transform

The signal x[n] is made to go through two filters g [] and h []. The first is g [] which is a high pass filter and is called the discrete mother wavelet. The second one is h [], is a low pass filter, the mirror of first filter. The signal is down-sampled by 2. The output of the high pass filter gives us the detailed co-efficient D1 while from the low pass filter we get A1 or the approximation coefficients. A1 is further decomposed and this goes on. Half band low pass filter has an impulse response h[n]. Thus, here convolution of the signal takes place with the impulse response of the filter.

$$x[n] * h[n] = \sum_{k=-\infty}^{\infty} x[k] \cdot h[n-k] \qquad 4.3$$

Low pass filter removes all the frequencies, which lie above half of the highest frequency of the signal. When the signal has been passed through half band lowpass filter, half of the samples are removed according to the Nyquist's rule, as the signal will now have the highest frequency. Overlooking every other sample present will subsample the signal by two, and thus, the signal will contain only half the number of points. The signal scale has now doubled.

Note that the lowpass filtering eliminates the high frequency data but leaves the scale unaffected. Only the subsampling procedure varies the scale. Resolution, on the other hand, is connected to the amount of data in the signal, and therefore, it is affected by the filtering processes. Half band lowpass filtering eliminates one half of the present frequencies, which leads to losing other half of the data. Thus, the resolution is split fifty-fifty after the filtering. The lowpass filtering halves the resolution but does not change the scale.

This can be shown as:

$$y[n] = \sum_{k=-\infty}^{\infty} h[k] \cdot x[2n-k] \qquad 4.4$$

DWT examines the signal at varying frequency bands which show different resolution after decomposition of the signal into a rough estimate and detail data. Two sets of functions- 1. The scaling functions and 2. The wavelet functions, are used, which correspond to low-pass and high-pass filters, respectively. The decomposed signals are obtained by high pass and lowpass filtering of the time domain signal.

One level of decomposition can be shown as follows:

$$y_{high}[k] = \sum_{n} x[n] \cdot g[2k - n] \qquad 4.5$$

$$y_{low}[k] = \sum_{n} x[n] \cdot h[2k - n] \qquad 4.6$$

Approximation co-efficient in wavelets refer to the smoothed signal after all the lowpass filtering. Details co-efficient are the remaining noise after all the high pass filtering. A1 designates the approximation coefficients and D1 designates the details coefficients. These coefficients can be broken down (decomposed) into additional coefficients in higher level systems.

Detail coefficients show orthogonality, which is they are independent of one other and can be added together in any random sequence.

4.3 Choice of Mother Wavelet:

Many different types of mother Wavelets are present. We can choose them based on the need and the type of applications. Simplest of them all is the Haar wavelet. Daubechies are more complex than Haar, but give better results

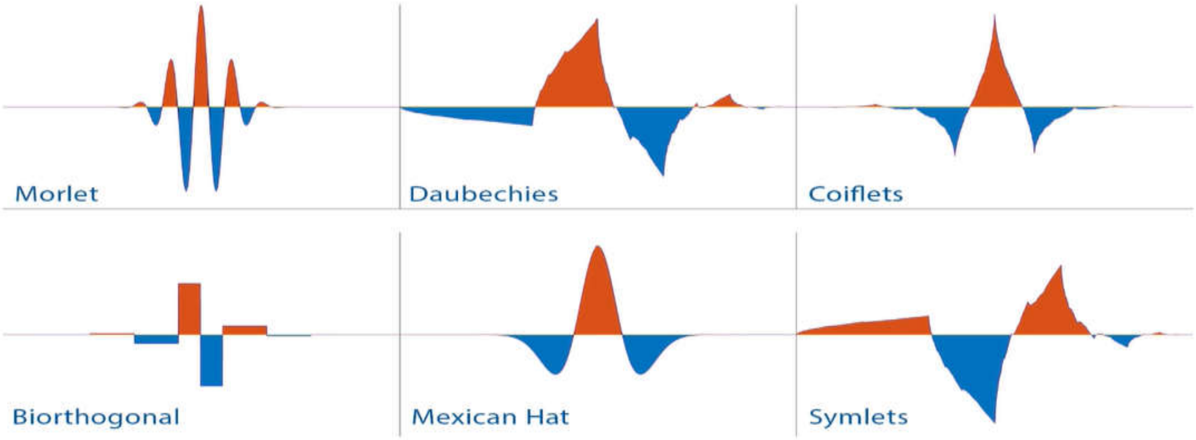


Figure 4.2 Different types of Mother Wavelet

Daubechies wavelets are known to have structure like QRS complexes and they have energy bands of low frequencies. Thus, it is probable that detail coefficients give better resolution after decomposition of the signal in time domain.

4.4 **Methodology**:

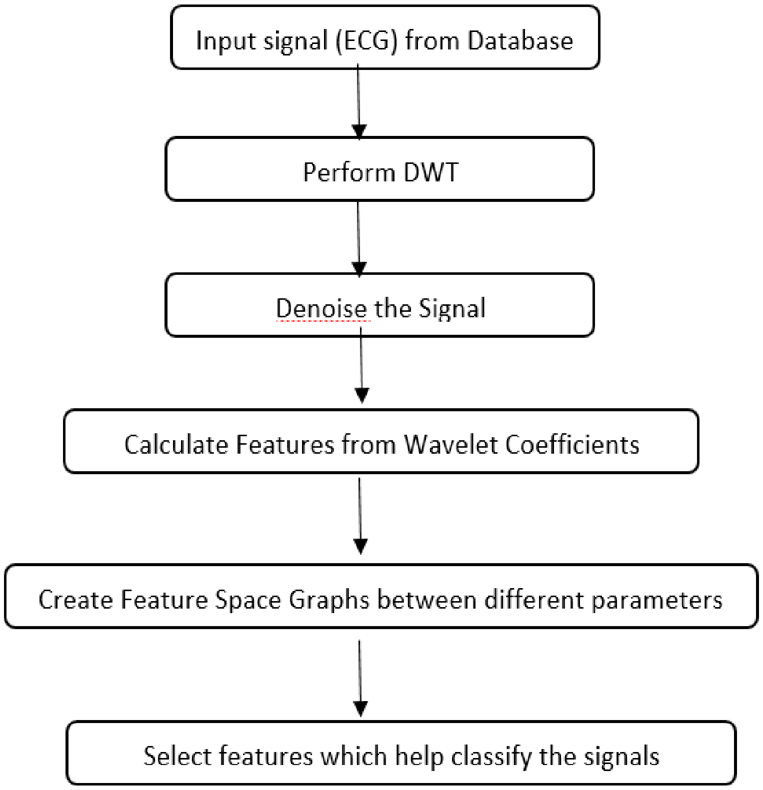


Figure 4.3 Methodology for Wavelet Transform

4.5 **Database**:

PTB database is provided by, the National Metrology Institute (NMI) of Germany, for research purposes. The ECGs were taken from healthy individuals and patients with dissimilar heart ailments. The ECGs were obtained by using a PTB prototype (non-commercial) recorder with the following features:

- 16 input channels
- Input voltage: ±16 mV
- Resolution: 16-bit with 0.5 µV/LSB
- Bandwidth: 0 - 1 kHz

The database consists of 549 records taken from 290 subjects (ages 17 to 87, mean 57.2; 209 men, mean age 55.5, and 81 women, mean age 61.6). Each subject is represented by one to five records. Each record comprises 15 concurrently measured signals: 12 conventional leads (i, ii, iii, v1, v2, v3, v4, v5, v6, avl, avr, avf) along with the 3 Frank lead (vx, vy, vz). Each signal is digitized at one thousand samples per second having a 16-bit resolution over a range of ± 16.384 mV.

4.6 Wavelet Analysis:

Wavelet analysis allows real time study of a signal in time and frequency domains. Daubechies 6 (db 6) was used for the multiresolution study of the electrocardiogram signals of a group of healthy and non-healthy people. In the initial results in the study, it was seen that the signal broke down at D7 and D8 levels when there were prominent variations in the two groups. Therefore, features like mean, Shannon entropy, basic statistical parameters, log energy, and energy density were stored upto D8 level and the different features were shown in feature space, to analyse whether the same could be put to use for signal classification.

4.7 Procedure:

First the signals are downloaded from the PTB database and segregated if they belong to healthy control or unhealthy person. Then the signal was decomposed using DWT using Daubechies 6 (db6) and the coefficients are shown below:

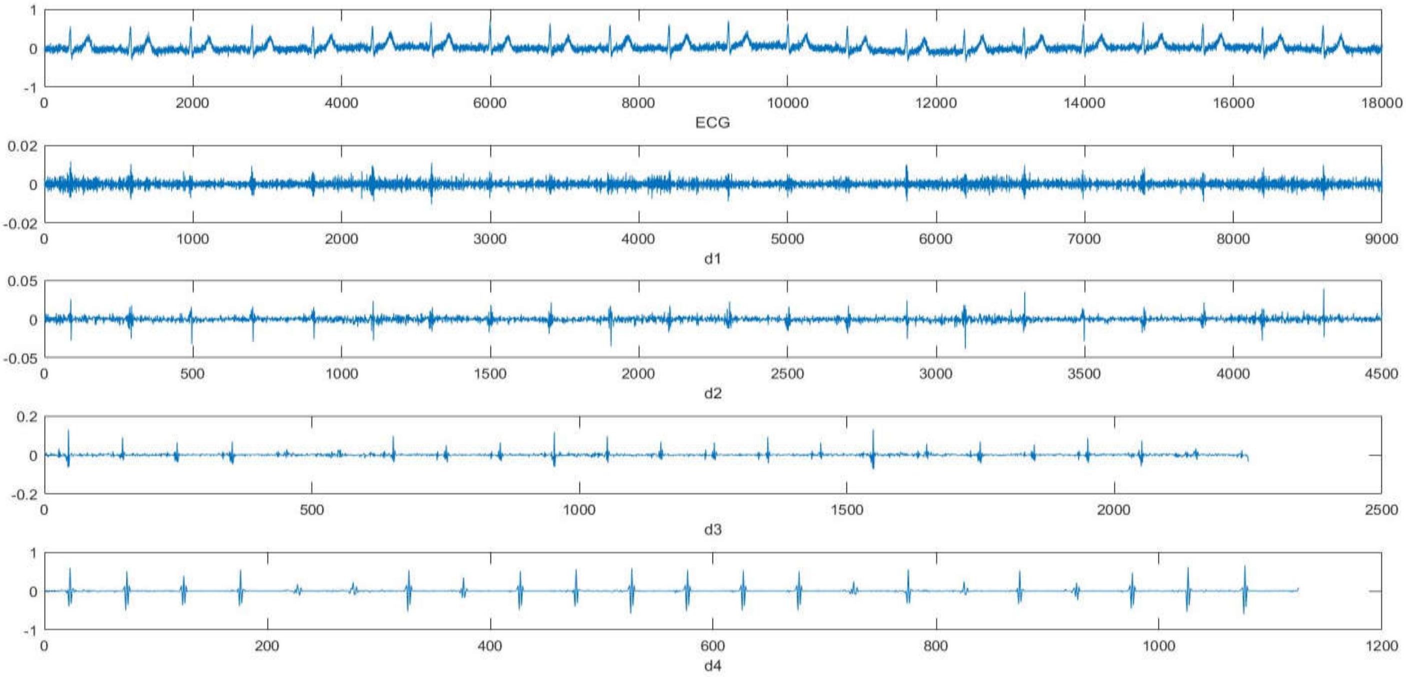


Figure 4.4 Wavelet Coefficients After Decomposition

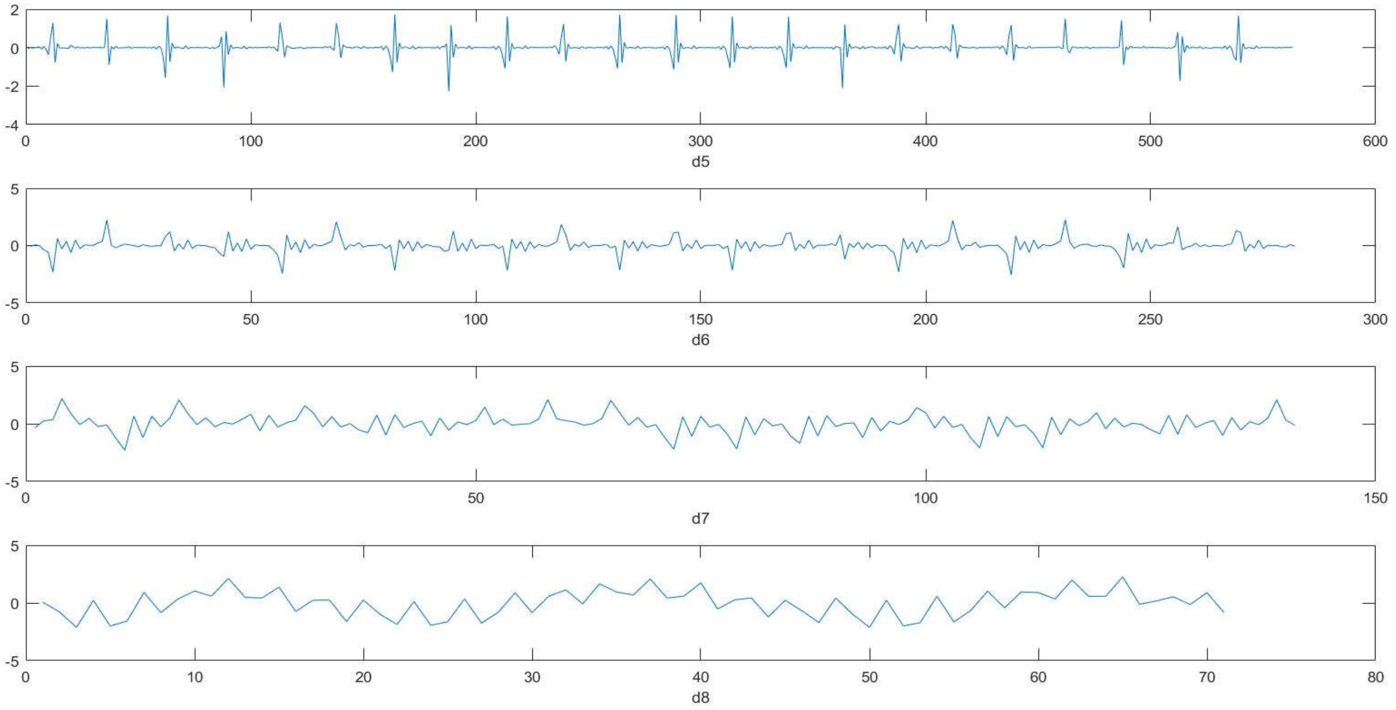


Figure 4.5 Wavelet Coefficients After Decomposition

The signal was analysed at higher levels of decomposition, so that low frequency components could be detected in the signal. The Figure 4.4 and 4.5 show that at lower levels, high-frequency components are distinct while at higher levels, low frequency components become more distinct. In the graphs we saw that d5 and d6 gave us best signal details. Values from the d5 and d6 levels specify their resemblance with db6 wavelet scaling function. Hence, d5 and d6 coefficients were recognized for the recognition and extraction of features like Shannon entropy, covariance, etc. for ECG signal classification as shown in Figure 4.7. Before feature extraction, wavelet has been used for de-noising of the signal as shown in Figure 4.6.

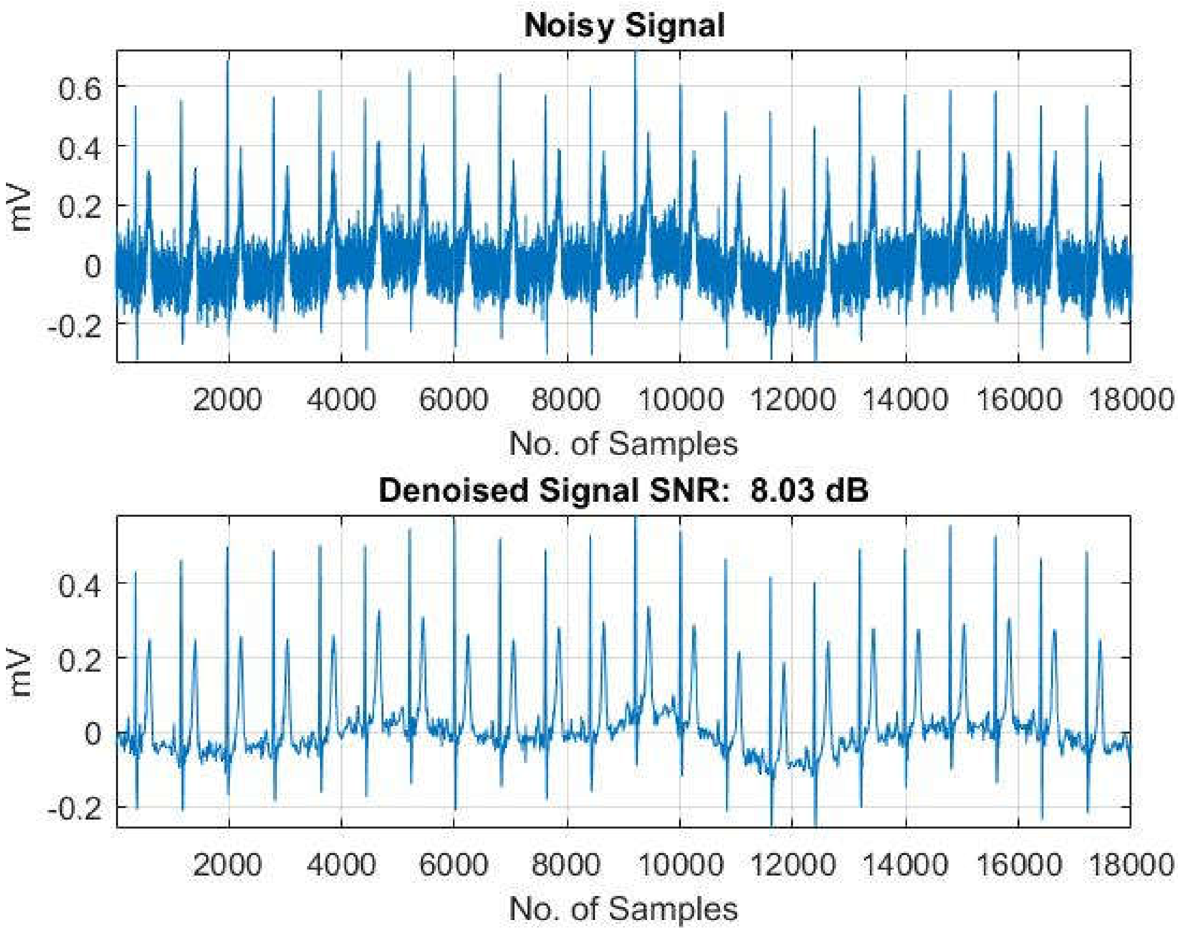


Figure 4.6 Signal with Noise and After Denoising Using Wavelet Transform

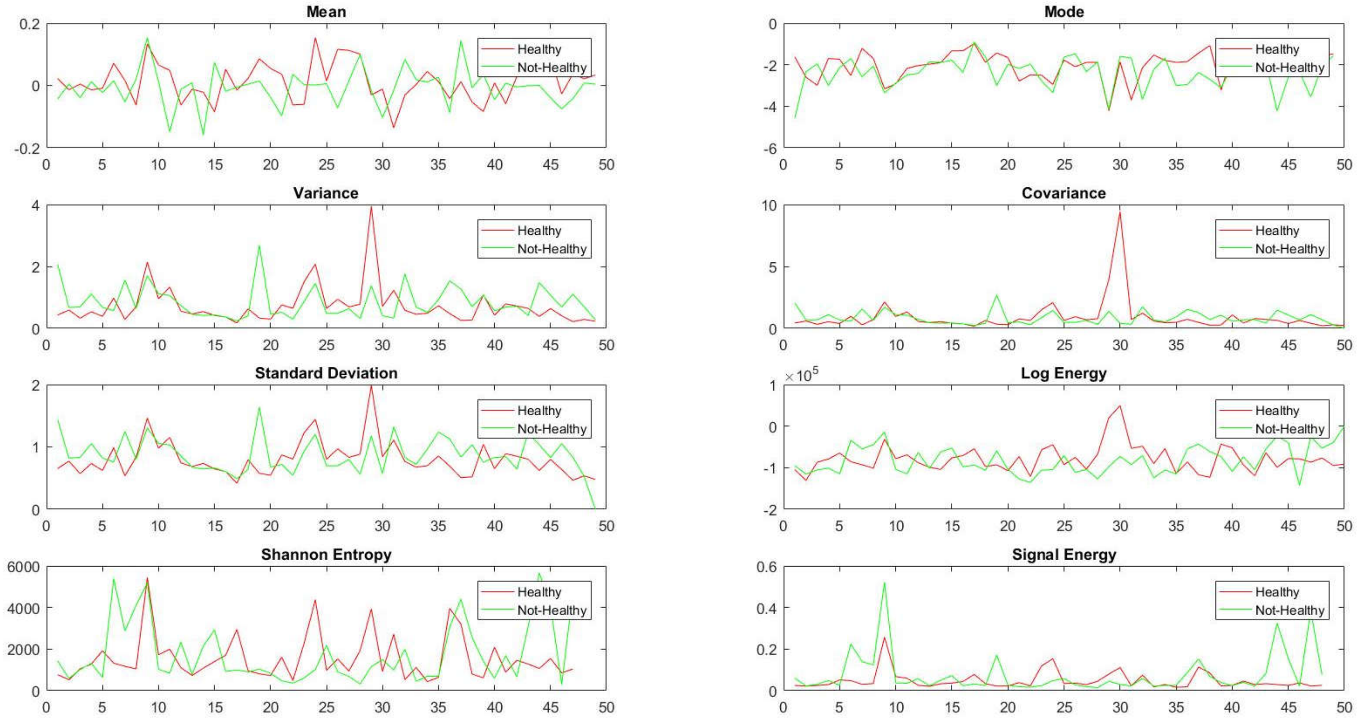


Figure 4.7 Features Extracted For Healthy And Non Healthy Patients

Feature space plot between different features was constructed using the extracted features as shown in Figure 4.8 and Figure 4.9. Out of various combinations of plots only two sets have been shown.

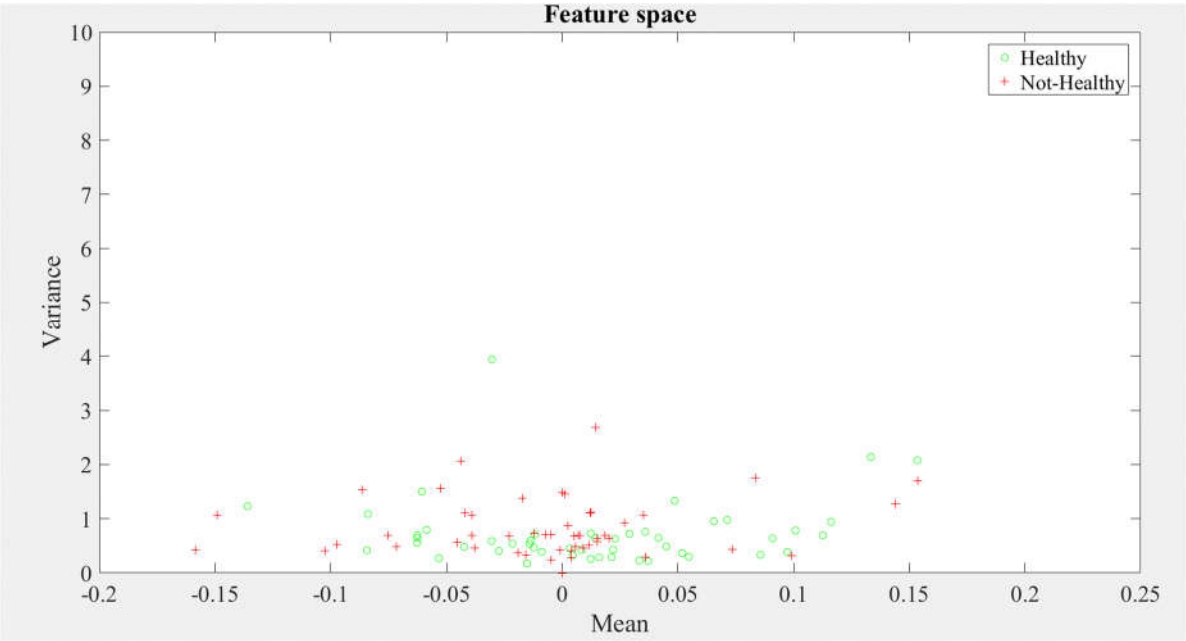


Figure 4.8 Feature Space of Mean vs Variance

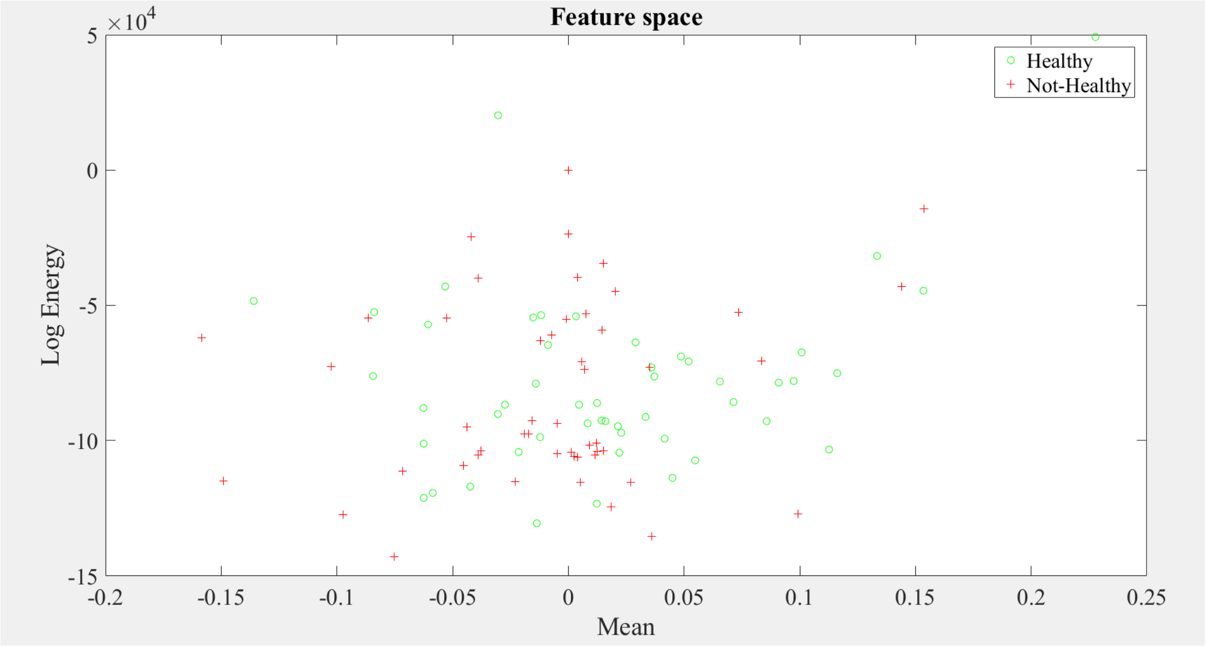


Figure 4.9 Feature Space of Mean vs Log Energy

As shown in the Figure 4.9, clear demarcation could be seen in the data for healthy and non-healthy signals. Hence these two parameters i.e. log Energy and mean could be promising for classification between healthy and non- healthy signals.

4.8 Results and Discussion:

The signal was analysed using wavelet transform (Daubechies wavelet) which helped in extracting artefacts from the signal. Mean and log energy entropy gave the best results and were plotted in the feature space to see if classification can be done by any classifier. The results indicate that the features were overlapping with each other for most cases and thus no classifier may be used for signal classification, except in the case of log energy and mean. Thus, they may be used in SVM and ANN for classification.

Chapter 5

Classification of ECG Signals (ANN and SVM)

5.1 Introduction:

Automatic detection of diseases in the signal is very essential in today's time, to prevent heart from deteriorating. Real time analysis of the signals is a challenging task for the doctors due to certain difficulties such as large number of patients to be diagnosed and lack of expertise of the doctor etc. This coupled with the increasing number of heart problems, in today's time, due to various factors, has initiated the need of research in this field. Emphasis lies on to automatically detect any problem in the signal, so that the involvement of the doctor can be minimised, and also aid the doctors in detection.

Many types of computer based digital processing techniques have been implemented in the past, with considerable amount of success. But for the past decade, with the advent of artificial intelligence and machine learning, pattern recognition has gained significant importance for research.

In a study done by Inan et al. 2006 features derived from wavelet transform and time domain were used to classify beats of a dataset by using neural networks. It was found out that the fourth scale of a dyadic wavelet transform with a quadratic spline wavelet together with the pre-/post RR-interval ratio is effective for distinguishing normal and premature ventricular contraction (PVC) from other beats. In another study done by Kong et al. (2007), a block based NN approach is shown for ECG classification. Modular blocks of NN, which is made up of 2-D array and is internally configurable, is applied using reconfigurable digital setup.

Weights of the connections and the type of network is optimized by using stochastic gradient method, with the rate of change directly proportional to the effect they had in the previous iteration. Another type of machine learning, support vector machine SVM, was used by Linh et al. 2004. They used Hermite polynomials and statistical based features to classify signals. Hu et al. (1997), used a local classifier along with a global classifier using a mixture of experts (MOE) approach. In a study performed by Chazal et al. 2004, linear discriminants were used to classify heart beats. Systems changes by training first with a local classifier and then uses a global classifier.

5.2 Theoretical Background-ANN:

Machine Learning is derived from Artificial Intelligence. It may indicate any form of technology that includes some intelligent aspects rather than pinpoint a specific technology field. In contrast, machine learning (ML) is a specific field of study. In other words, one can use ML to point to a specific group of artificial intelligence. ML includes many tools as well. Machine learning is a modelling technique that involves data. It is used where physical laws fail to give a good enough model. As the name implies, some initial data is fed into the system, so that it can understand the pattern. This is called training data. A problem that this method faces is that, during training, unbiased data should be used, finding which is a problem. Thus, the method, which helps us in making the performance of machine learning consistent, regardless of the training data, is called generalisation. Generalization fails, when the system is designed to over fit the data (Kim, 2017), as shown in the Figure 5.1 and 5.2.

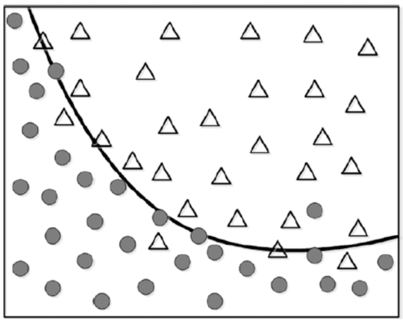


Figure 5.1 Non Overfitting Classification (Kim, 2017)

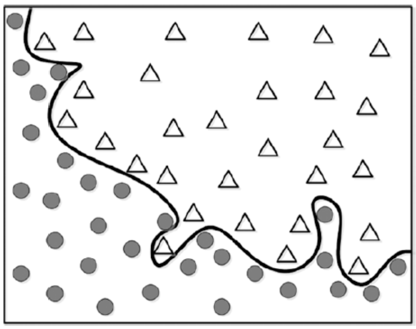


Figure 5.2 Overfitting Classification (Kim, 2017)

Regularization is a method employed, to keep the model as simple as possible. This avoids overfitting at a small cost of performance. The neural network attempts to operate like the nervous system of humans. The brain has connections of numerous neurons and the neural network is built with connections of nodes, which are entities that are analogous to the neurons of the brain. It mimics the neuron inter-linking, which is the most important mechanism of the brain.

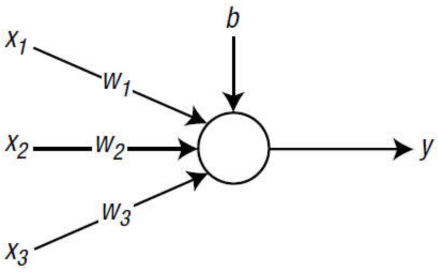


Figure 5.3 Node receiving three inputs (Kim, 2017)

The circle and arrow shown in the Figure 5.3 denote the node and signal flow, respectively. x1, x2, and x3 are the input signals. w1, w2, and w3 are the weights for the respective signals. B is the bias, another factor linked with the storage of data suggesting that information of the neural network is stored as bias and weights.

The input signal from the outside is multiplied by the weight in advance (before it reaches the node). Once the weighted signals are collected at the node, these values are added to be the weighted sum.

The weighted sum of this example is calculated as follows:

$$v = (w_1 \times x_1) + (w_2 \times x_2) + (w_3 \times x_3) + b \qquad 5.1$$

This equation indicates that the signal with a greater weight has a greater effect.

$$v = wx + b \qquad 5.2$$

Where w is a row vector of w1, w2, w3 and x is a cloumn vector of x1,x2,x3.

The node enters the weighted sum into the activation function and yields its output. The activation function determines the behavior of the node.

$$y = \varphi(v) \qquad 5.3$$

where, *$\varphi(v)$* is the activation function.

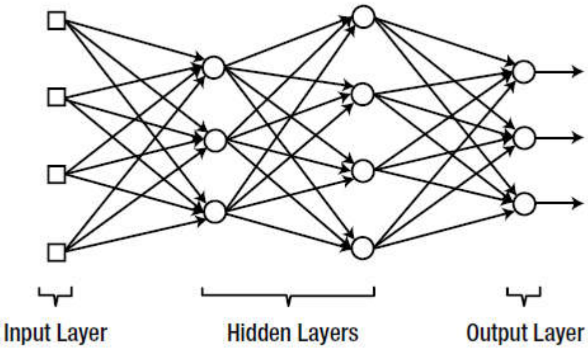


Figure 5.4 Structure of Nodes (Kim, 2017)

Supervised Learning of a Neural Network (Kim, 2017) refers to the use of training dataset. The basic algorithm is that, the weights of the nodes are initially randomised, with proper values. Then training data, which consists of inputs and the correct outputs, is given to the network. An error is calculated based on the outputs, present in the training data. Weights are adjusted again, to minimise error or also known as cost function.

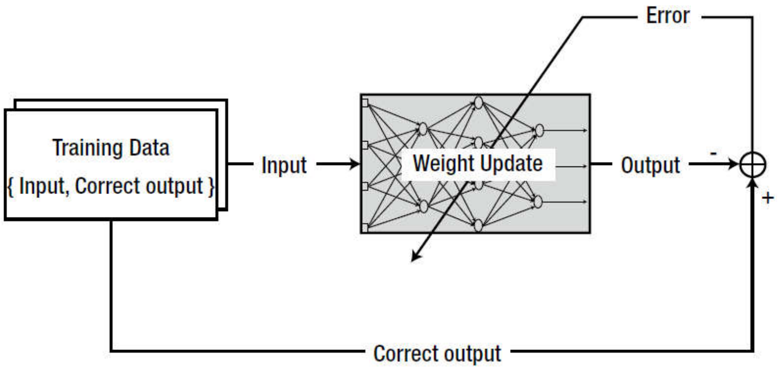


Figure 5.5 Supervised Learning (Kim, 2017)

Delta Rule: As we have seen, to train the neural network, weights have to be changed in each iteration. The rule which governs this weight change is called the Delta Rule.

$$(\text{error})e_i = d_i - y_i \qquad 5.4$$

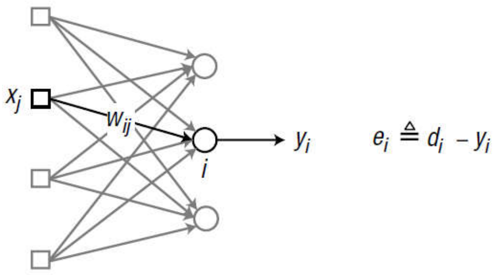


Figure 5.6 Single Layer Neural Network(Kim, 2017)

The rule states that the change in the weights, for a given node, is directly proportional to the input to the node x_j as well as the error produced by the node e_i.

$$w_{ij}(new) = w_{ij}(old) + \alpha\, \delta_i\, x_j \qquad 5.5$$

Where $\delta_i = \varphi'(v_i)\, e_i$

Here x_j-output of the j node

e_i-error of node i

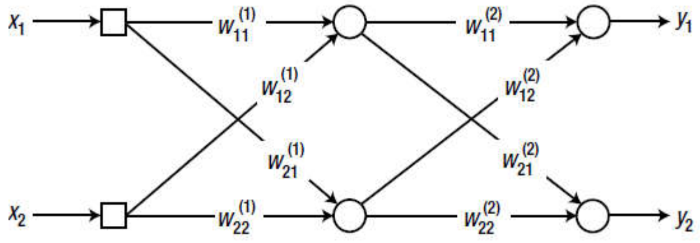
w_{ij}- weight between the output node i and input node j

α- Learning rate $(0 < a < 1)$

v_i- weighted sum of the output node i

The value of learning rate will help decide the change in weights per iteration/epoch. If the value is very large, then the solution may fail to converge, while if it is too low then, the calculations approach the solution very slowly.

Multi-Layer Neural Network (Kim, 2017)**:** Delta rule as discussed, does not work for multilayer neurons, because the error could not be initially defined for the hidden layers. This was because, the training data does not provide any value for the hidden nodes. To solve this problem, Back Propagation Algorithm was developed. In the back-propagation algorithm, the output error starts from the output layer and moves backward until it reaches the right next hidden layer to the input layer.



.

Figure 5.7 Neural network that consists of two nodes for the input and output (Kim, 2017)

The weighted sum of the hidden node is given as

$$\begin{bmatrix} v_1^{(1)} \\ v_2^{(1)} \end{bmatrix} = \begin{bmatrix} w_{11}^{(1)} & w_{12}^{(1)} \\ w_{21}^{(1)} & w_{22}^{(1)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad 5.6$$

$$\triangleq W_1 x$$

The output is calculated by putting this weighted sum into activation function where $y_1^{(1)}$ and $y_2^{(1)}$ are outputs from the hidden nodes

$$\begin{bmatrix} y_1^{(1)} \\ y_2^{(1)} \end{bmatrix} = \begin{bmatrix} \varphi\left(v_1^{(1)}\right) \\ \varphi\left(v_2^{(1)}\right) \end{bmatrix} \qquad 5.7$$

Weighted sum of the output node is given as

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} w_{11}^{(2)} & w_{12}^{(2)} \\ w_{21}^{(2)} & w_{22}^{(2)} \end{bmatrix} \begin{bmatrix} y_1^{(1)} \\ y_2^{(1)} \end{bmatrix} \qquad 5.8$$

$$\triangleq W_2 y^{(1)} \qquad 5.9$$

Thus, the output of the neural network after putting the result of equation 5.9 into the activation function is:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \varphi(v_1) \\ \varphi(v_2) \end{bmatrix} \qquad 5.10$$

Now calculating delta for each node as

$e_1 = d_1 - y_1$

$\delta_1 = \varphi'(v_1) e_1$ 5.11

$e_2 = d_2 - y_2$ 5.12

$\delta_2 = \varphi'(v_2) e_2$

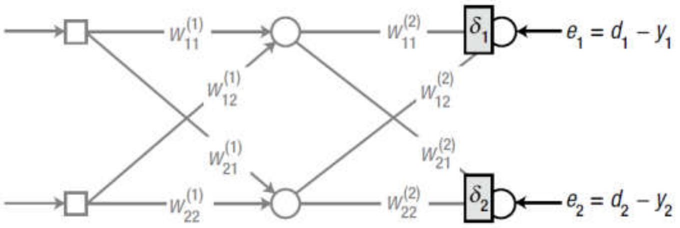


Figure 5.8 Train the neural network using the back-propagation algorithm (Kim, 2017)

where y_i is the output from the output node, d_i is the correct output from the training data, and v_i is the weighted sum of the corresponding node.

Next calculating delta for the hidden nodes, as shown

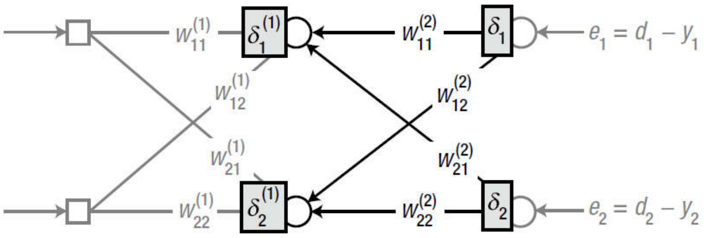


Figure 5.9 Proceed leftward to the hidden nodes and calculate the delta (Kim, 2017)

In the back-propagation algorithm, the error of the node is defined as the weighted sum of the back-propagated deltas from the output layer as

$$\begin{aligned} e_1^{(1)} &= w_{11}^{(2)}\delta_1 + w_{21}^{(2)}\delta_2 \\ \delta_1^{(1)} &= \varphi'\left(v_1^{(1)}\right)e_1^{(1)} \end{aligned} \qquad 5.13$$

$$\begin{aligned} e_2^{(1)} &= w_{12}^{(2)}\delta_1 + w_{22}^{(2)}\delta_2 \\ \delta_2^{(1)} &= \varphi'\left(v_2^{(1)}\right)e_2^{(1)} \end{aligned} \qquad 5.14$$

$$\begin{bmatrix} e_1^{(1)} \\ e_2^{(1)} \end{bmatrix} = \begin{bmatrix} w_{11}^{(2)} & w_{21}^{(2)} \\ w_{12}^{(2)} & w_{22}^{(2)} \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix} \qquad 5.15$$

$$\begin{bmatrix} e_1^{(1)} \\ e_2^{(1)} \end{bmatrix} = W_2^T \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix} \qquad 5.16$$

Once all the deltas have been calculated, we train the neural network. We use the following equation to adjust the weights of the respective layers.

$$\begin{aligned} \Delta w_{ij} &= \alpha \delta_i x_j \\ w_{ij} &\leftarrow w_{ij} + \Delta w_{ij} \end{aligned} \qquad 5.17$$

Where x_j the input signal to the corresponding weights.

5.3 Theoretical Background-SVM:

Support Vector Machine (SVM) (Haykin, 2009) is a supervised learning algorithm which tries to analyse data and classify it. It was studied by Vapnik et al. 1992, who used the Kernel Trick, thus allowing SVM to classify non-separable data.

SVM employs a hyperplane or as shown in the figure, a line, for 2-D cases which tries to find out on which side of the plane the data falls. This hyperplane initially is formed from the training dataset the user provides.

The next step after finding the hyperplane, is to find the optimal hyperplane. An optimal hyperplane is the plane, which has the widest margins, and this is the objective of SVM, so that it will generalize better with new data.

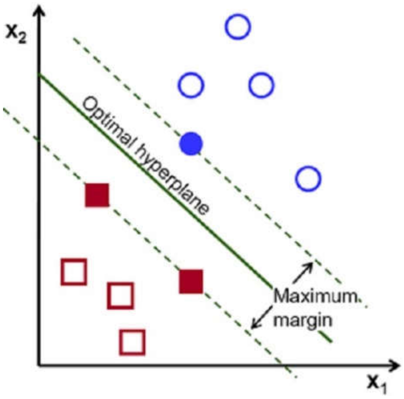


Figure 5.10 Classification with a Liner Hyperplane (AI Trends, 2018)

Inner-product kernel between a support vector x_i and a vector x, taken from the input space of data, is a concept, integral to SVM. The support vectors are a subset of all dataset, chosen by a learning algorithm.

Take a training dataset $(x_i, d_i)^N_{i=1}$ where x_i refers to the input pattern and d_i is the output for i^{th} value. The equation of the hyperplane is

$$\mathbf{w}^T\mathbf{x} + b = 0 \qquad 5.18$$

x is input vector, w is adjustable weight vector, b is a bias.

Let w_o and b_o be the optimal values of weight and bias, then the equation of the optimal hyperplane is

SVM will minimise the cost function and try to find the optimal hyperplane.

The cost function(J(w)) is defined $\mathbf{w}_o^T\mathbf{x} + b_o = 0$ as: 5.19

$$J(w) = \frac{1}{2}w^T w = \frac{1}{2}\|w\|^2 \qquad 5.20$$

Which is subjected to the constraints $d_i \left[w^T x_i + b\right] \geq 1 \qquad i = 1, ..., N.$ 5.21

This process, to optimise the constraint is called primal problem. Thus, here we must hypothesize that the cost function is a function which is convex in nature and that we have linear constraints. The constraint optimization is tackled by using Lagrange multiplier.

The Langrangian function is $J(w, b, \alpha) = \frac{1}{2} w^T w - \sum_{i=1}^{N} \alpha_i \left[d_i (w^T x_i + b) - 1\right]$ 5.22

α_i-Lagrange multipliers (auxiliary non-negative variables)

The Lagrange function is minimized based on weight(w) and bias(b), but maximised w.r.t α_i. as shown:

$$\begin{aligned} &\text{Condition 1: } \frac{\partial J(w,b,\alpha)}{\partial w} = 0 \\ &\text{Condition 2: } \frac{\partial J(w,b,\alpha)}{\partial b} = 0 \end{aligned} \qquad 5.23$$

When condition 1 is applied, we get α_i $w = \sum_{i=1}^{N} \alpha_i d_i x_i$ 5.24

When condition 2 is applied, we get $\sum_{i=1}^{N} \alpha_i d_i = 0$ 5.25

Another tool that SVM uses is the concept of duality. In dual problem, Lagrange multipliers are given as the optimal solution. Duality theorem states that:

1. Optimal solution of the two problems is the same as the solution of the primal problem.

2. a_0 is the optimal solution of dual problem and w_0 is that of the primal problem.

Rewriting the Lagrange multiplier function as:

$$J(w, b, \alpha) = \frac{1}{2} w^T w - \sum_{i=1}^{N} \alpha_i d_i w^T x_i - b \sum_{i=1}^{N} \alpha_i d_i + \sum_{i=1}^{N} \alpha_i \qquad 5.26$$

Using equation 5.24. And applying it in the above equation 5.26, we get

$$w^T w = \sum \alpha_i d_i w^T x_i = \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j d_i d_j x_i^T x_j \qquad 5.27$$

We know from the duality problem J (w,b,α)=Q(α), is

$$Q(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j d_i d_j x_i^T x_j \qquad 5.28$$

Now, the Lagrange multipliers, are used so that we can find out the dual problem solution as:

$$\max\{Q(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i (d_i d_j x_i^T x_j) \alpha_j\} \qquad 5.29$$

Subject to the constraints: $\alpha_i \geq 0, \quad \sum_{i=1}^{N} \alpha_i d_i = 0$ 5.30

If training data is not disconnected by a hyperplane, then a slack variable ξ_i is used, where $\xi_i > 0$, i=1…, N, that signifies the linearity is not present.

$$d_i \left[w^T x_i + b \right] \geq 1 - \xi_i \qquad 5.31$$

$$\xi_i \geq 0, \qquad i = 1, \ldots\ldots, N. \qquad 5.32$$

For $0 < \xi_i < 1$, data lies inside the separation region.

Cost function can be minimised as

$$\min_{w,\xi} P_1(w, \xi) = \frac{1}{2} w^T w + C \sum_{i=1}^{N} \xi_i \qquad 5.33$$

Where C is the regularization parameter.

5.4 Methodology:

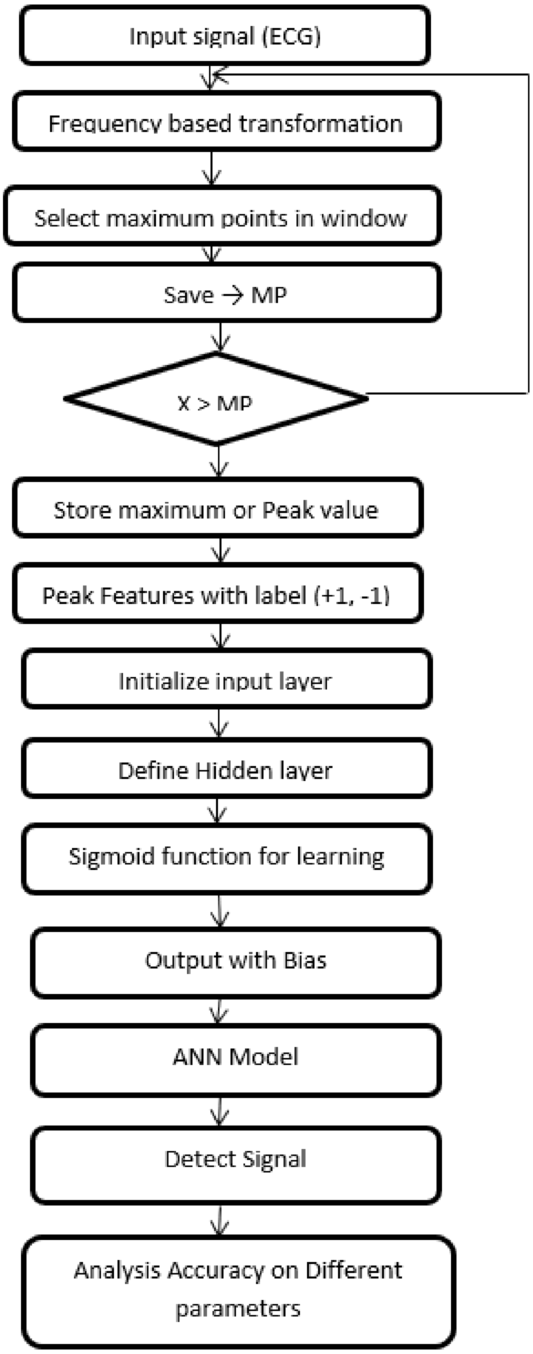


Figure 5.11 Flowchart of ANN working

Algorithm

Step 1: Input the ECG signal.

Step 2: Perform frequency-based transformation.

Step 3: Select the maximum points in the window and save them.

Step 4: Check that x> MP.

Step 5: Store the maximum or peak value.

Step 6: Select the peak features with label {+1, -1}

Step 7: Initialize input layer

$$u_j^1 = f\left[\sum_{i=1}^{N} w_{i,j}\, u_{i,m}\right], 1 \leq j \leq N \qquad 5.33$$

Give the output of input layer as an input in hidden layer

$$u_j^2 = f\left[\sum_{i=1}^{N} w_{j.k}^{\,i} u_j^i\right], 1 \leq k \leq H \qquad 5.34$$

Output of hidden layer is input to the output layer.

$$y_m = f\left[\sum_{K=1}^{H} w_{k,m}^{\,2} u_k^2\right], 1 \leq m \leq M \qquad 5.35$$

Step 8: Detect the signal

Step 9: Analyze the accuracy of different parameters.

5.5 Procedure:

At first the signal (as shown in Figure 5.12) is loaded for which classification is to be performed. Then the location of the peaks and their amplitudes is stored in a file. The window size is taken as ‘6’ samples.

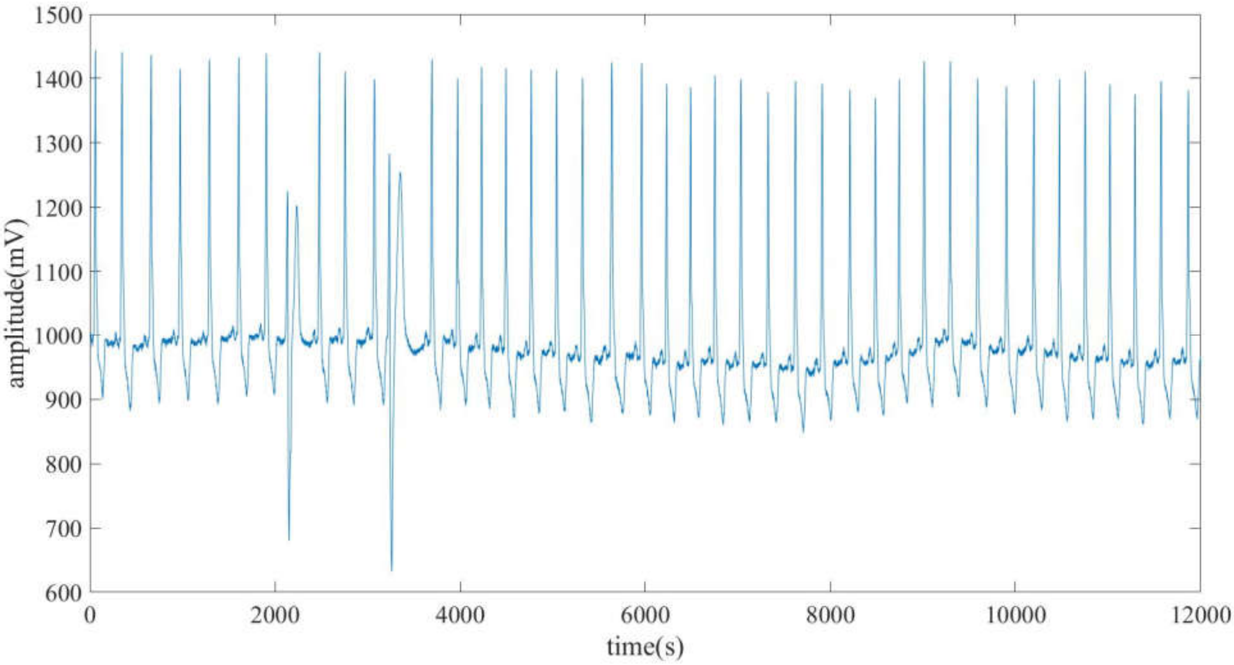


Figure 5.12 Original ECG Signal

Then the period of the signal is calculated, and the value is stored. Then the frequency analysis is performed on the signal as shown in Figure 5.13.

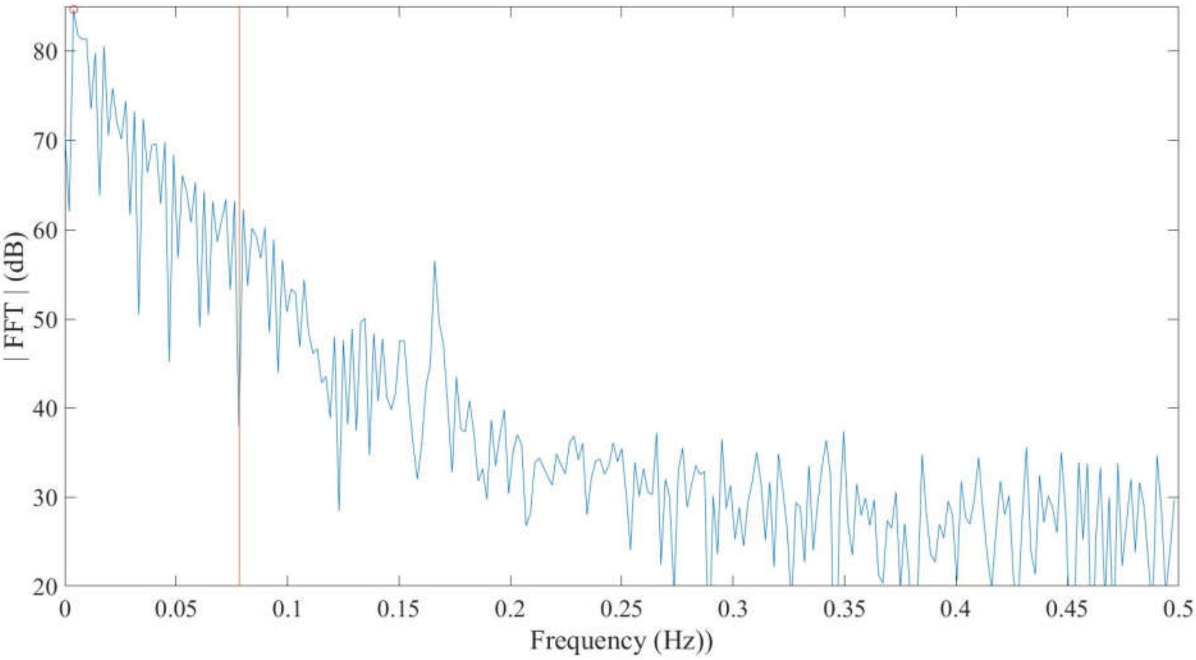


Figure 5.13 Signal After FFT

In order to obtain the unbiased data set to be used as training set, the data is split and shuffled into two equal parts. This is done so that no bias occurs in the dataset. One such dataset obtained is used for training and the other for testing. Neural network is created, by defining certain parameters as following:

1. Learning Rate
2. Regularization Type
3. Activation Function
4. Batch Size
5. Number of nodes in each layer

After the training and testing is over, the testing and training accuracies could be calculated, along with the cost function.

$$\text{Training Accuracy} = 100*\frac{Correct\ Trained\ Points}{Total\ No.of\ points} \qquad 5.36$$

$$\text{Testing Accuracy} = 100*\frac{Correct\ Tested\ Points}{Total\ No.of\ points} \qquad 5.37$$

The cost function, which provide the error, can also be determined for every iteration as shown in Figure 5.14.

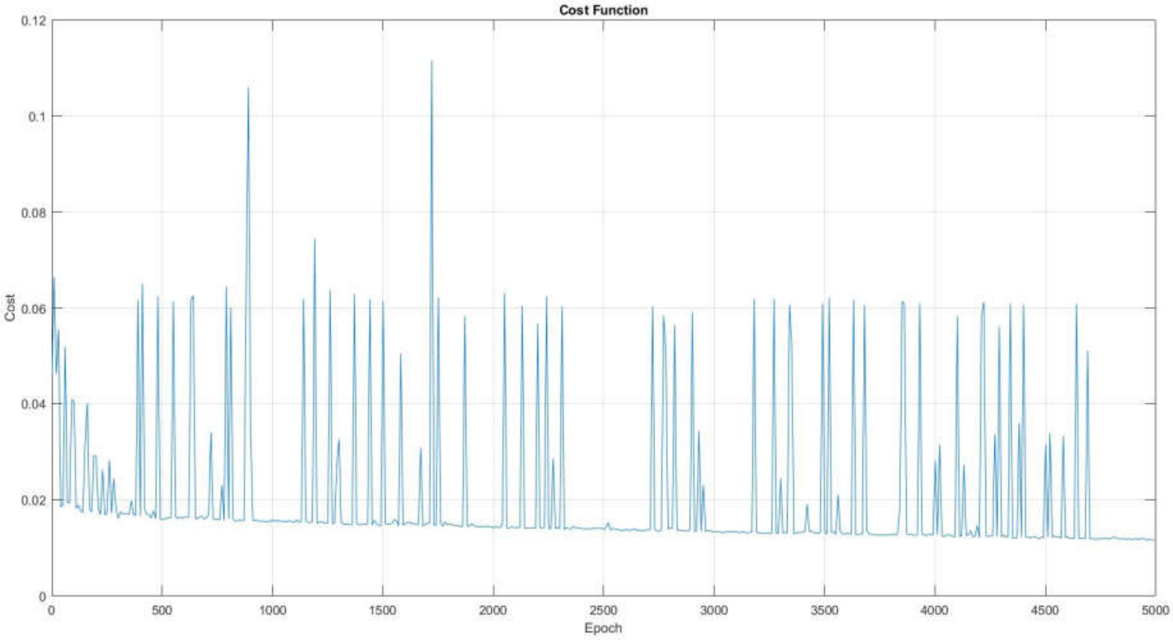


Figure 5.14 Cost Function vs Iteration

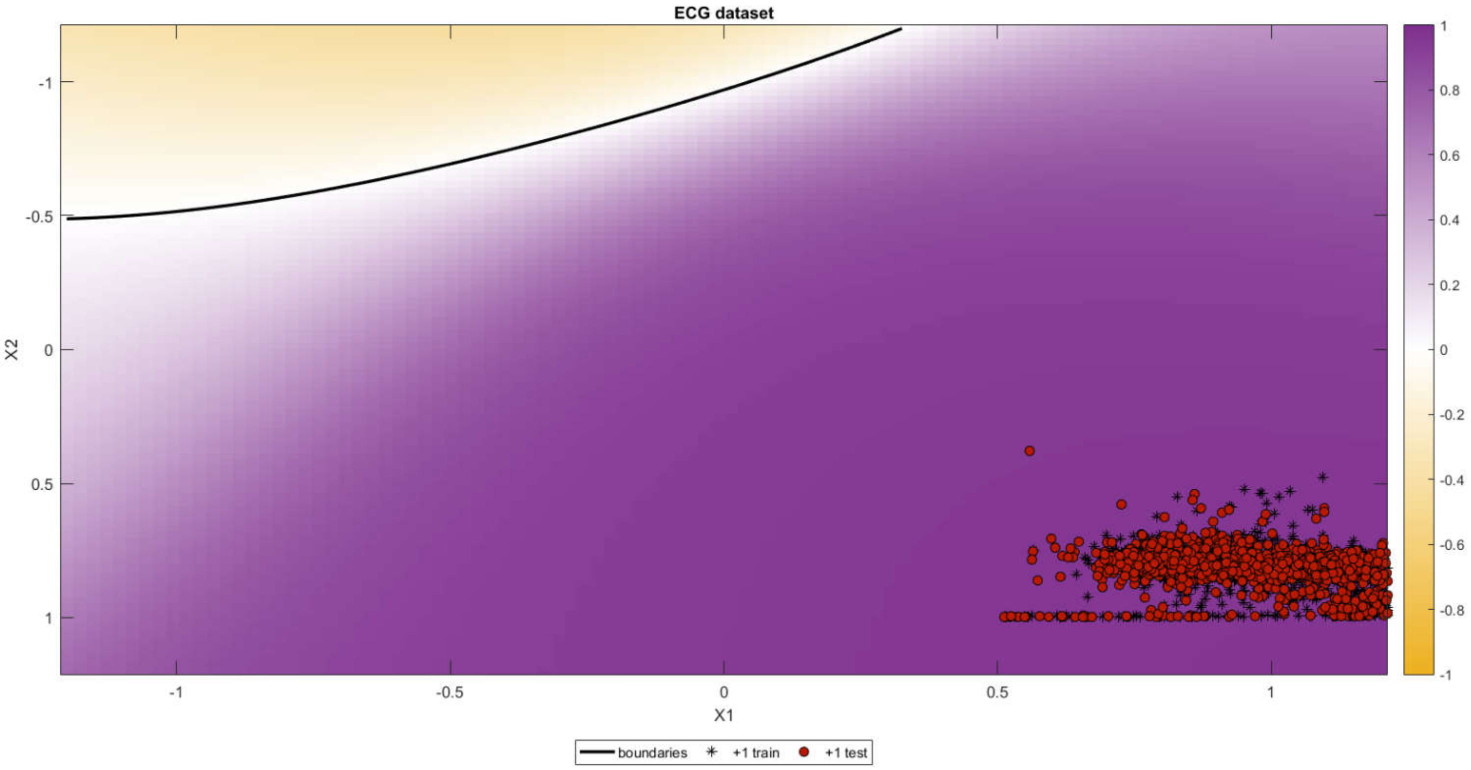


Figure 5.15 Training and Testing Dataset

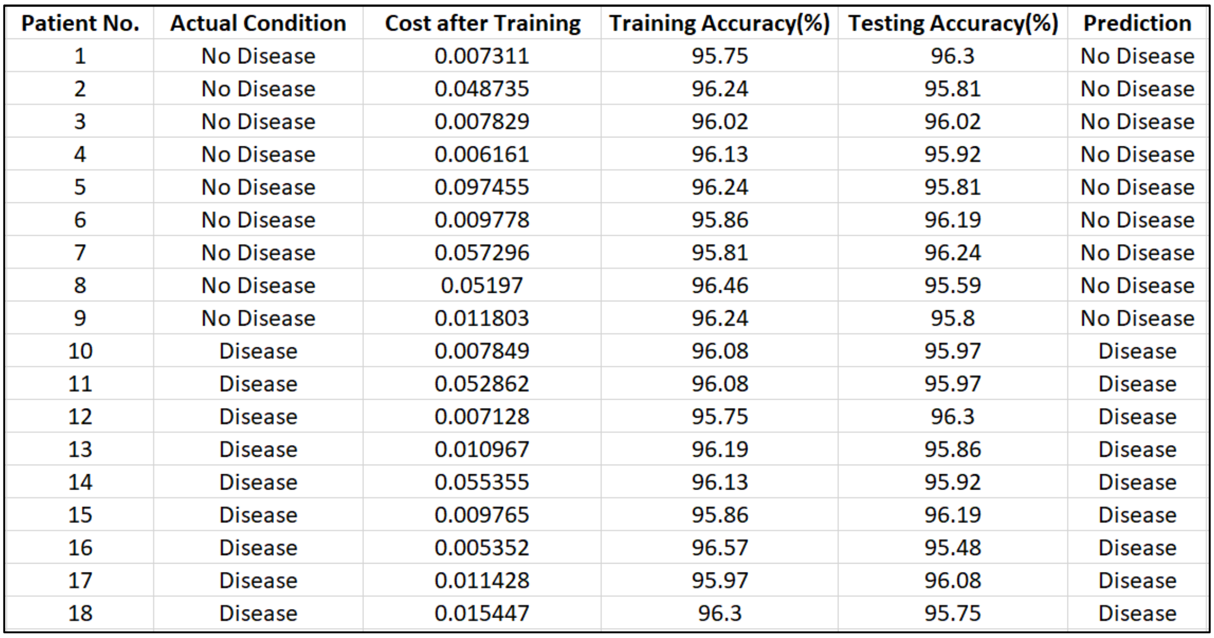
As the neural network has been created, now the file is sent into the network, to be classified and the result is obtained as ‘Disease’ or ‘Not Disease. Figure 5.15 shows the signal classification and dataset used for training and testing.

As can be seen in the Figure 5.15, the dataset used for training is highly overlapping and therefore is difficult to classify. The black curved line shown in the graph classifies the data into two categories (i.e. ‘disease’ and ‘non-disease’). The yellow region represents the data with disease while the magenta portion shows the non-disease region.

5.6 Results and Discussion:

In this section SVM model was created and accuracy of the model was found to be 50%, so the model was not further used for classification.

The ANN network was tested upon 18 signals downloaded from the MIT-BIH database, of which first 9 were non-disease signals, and the next nine were disease signal. The results are shown in the Table 3.



Patient No.	Actual Condition	Cost after Training	Training Accuracy(%)	Testing Accuracy(%)	Prediction
1	No Disease	0.007311	95.75	96.3	No Disease
2	No Disease	0.048735	96.24	95.81	No Disease
3	No Disease	0.007829	96.02	96.02	No Disease
4	No Disease	0.006161	96.13	95.92	No Disease
5	No Disease	0.097455	96.24	95.81	No Disease
6	No Disease	0.009778	95.86	96.19	No Disease
7	No Disease	0.057296	95.81	96.24	No Disease
8	No Disease	0.05197	96.46	95.59	No Disease
9	No Disease	0.011803	96.24	95.8	No Disease
10	Disease	0.007849	96.08	95.97	Disease
11	Disease	0.052862	96.08	95.97	Disease
12	Disease	0.007128	95.75	96.3	Disease
13	Disease	0.010967	96.19	95.86	Disease
14	Disease	0.055355	96.13	95.92	Disease
15	Disease	0.009765	95.86	96.19	Disease
16	Disease	0.005352	96.57	95.48	Disease
17	Disease	0.011428	95.97	96.08	Disease
18	Disease	0.015447	96.3	95.75	Disease

Table 3 Result After ANN Classification

As, can be seen in the Table 3 that the code developed was able to predict the disease from the non-disease signal with 100% accuracy in all the 18 signals. This might vary when the size of the sample is increased.

In order to understand the effect of various parameters (which were used to define the properties of the ANN model) on the model accuracy, the parameters were varied for the same signal.

1. Learning Rate(α): Value of α lies between 0 and 1. It signifies how much weight is changed per iteration. If this value is higher than what is wanted, then the output fails to converge, whereas if it is too low, calculations might take a lot of time.

Figure 5.16 Cost Function vs Learning Rate

As can be seen in the Figure 5.16, as the learning rate increases, final cost initially decreases, then increases slightly and further decreases. It was found that the lowest value of cost function came around 0.4 learning rate, after which the value of cost function again increased. This is in line with Kim (2017). which states that either too low or too high value of learning rate decreases the efficiency of the algorithm.

Figure 5.17 shows the effect of learning rate on model accuracy.

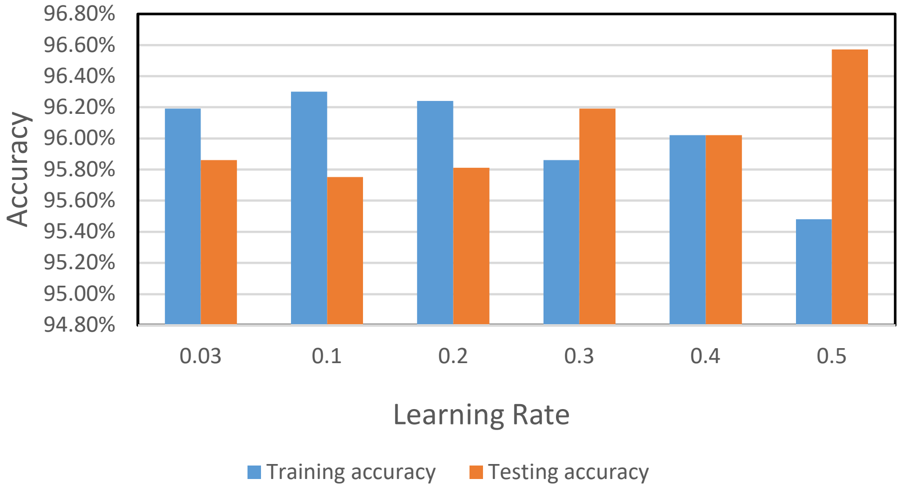


Figure 5.17 Training and Testing Accuracy vs Learning Rate

Figure 5.17 shows that, as the learning rate increases, the training accuracy with respect to the testing accuracy, decreases. Around the learning rate 0.4, the testing accuracy and training accuracy are about the same, and after that testing accuracy increases.

2. Activation Function(φ): It determines the behaviour of the nodes. The main use of an activation function is to convert an input signal of a node into an output signal, which would be the input to the next layer. Therefore, summation of the products of entered data(X) and weights (W) are input to an activation function φ (x) to get the resultant of that layer and forward it as an input to the next layer. The algorithm was tested on 4 functions:

1.Linear: When we do not use any kind of activation function, then the output shows a linear nature with respect to the input. A linear function is a polynomial of one degree. They are easy to solve but are restricted in their complexity and thus have less power to learn complex functional mappings from data.

2.Sigmoid: It is an S shaped curve, ranging from 0 to 1. It is very easy to apply and understand.

$$f(x) = \frac{1}{1+exp(-x)}$$

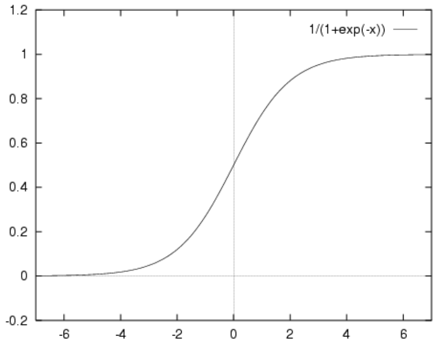


Figure 5.18 Sigmoid Function

3.Tanh: It is a function which is centered around zero as it lies between -1 to 1. This provides for a better optimisation and thus is considered much better than sigmoid.

$$f(x) = \frac{1-\exp(-2x)}{1+\exp(-2x)}$$

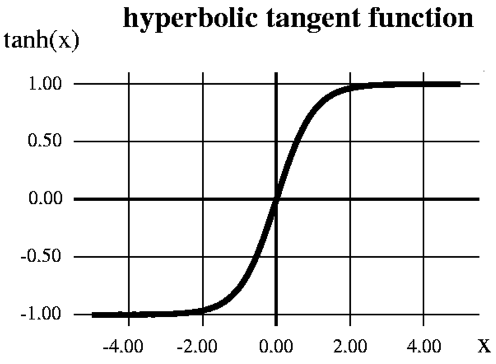


Figure 5.19 tanH Function

4. RELU: Rectified Linear Units is nowadays very popular activation function. It was recently proved that it had 6 times improvement in convergence from Tanh function. It is known to be very efficient and simple.

R(x) = 0 if x < 0,

R(x) = x if x >= 0

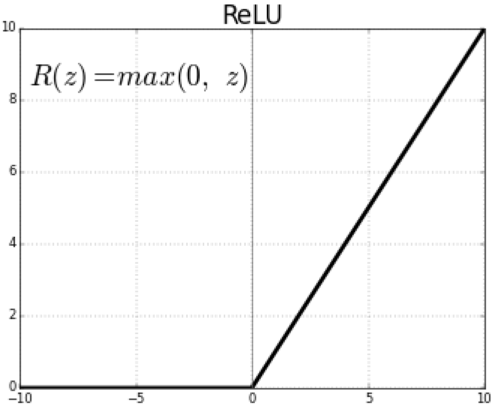


Figure 5.20 Relu Function

Figure 5.21 shows the comparison of training and testing accuracy obtained using different activation functions.

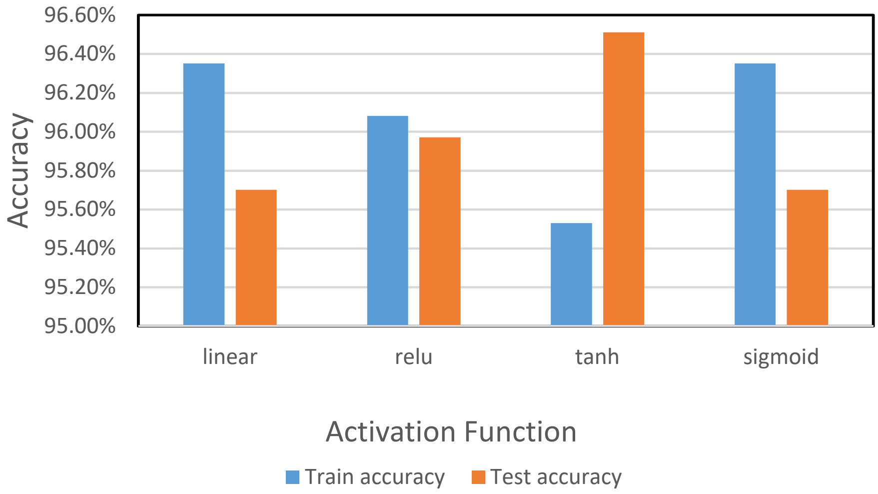


Figure 5.21 Activation Function vs Training And Testing Accuracy

As can be seen in the Figure 5.21, out of the 4 functions used, ReLU function shows the best results, considering there was not much gap between training and testing accuracy. For others, either the testing accuracy was low, or the training accuracy was low.

Figure 5.22 shows the variation of cost function versus activation function.

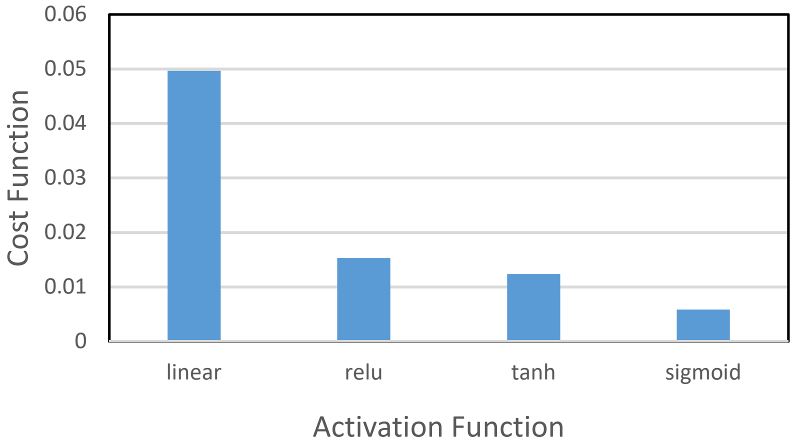


Figure 5.22 Final cost after training vs activation function

As shown in the Figure 5.22, representing cost function versus activation function, one can see that linear function has the most error, as it showed the highest value of cost amongst all functions. This behaviour can be attributed to the fact that linear model is too simple and thus does not perform well. Sigmoid function is easy to implement and understand, but of late, is not being used much, as found in the literature. This was because, it showed vanishing gradient problem, also it was not centered around zero. Moreover, they show slow convergence rates.

Tanh function solved all other problems of sigmoid function, but still showed vanishing gradient problem. Of late, ReLU function has been the most used and in the present work best results were found using ReLU. The only problem associated with ReLU function is that it can only be used in the hidden layers.

3. Batch Size: The Stochastic Gradient Descent (SGD) measures the error for every training data and adjusts the weights. In batch method, a collection of datapoints is taken and used for training, with average weight update. This is done because it requires less memory and the network trains faster with mini- batches.

Figure 5.23 shows the variation of cost function with batch size.

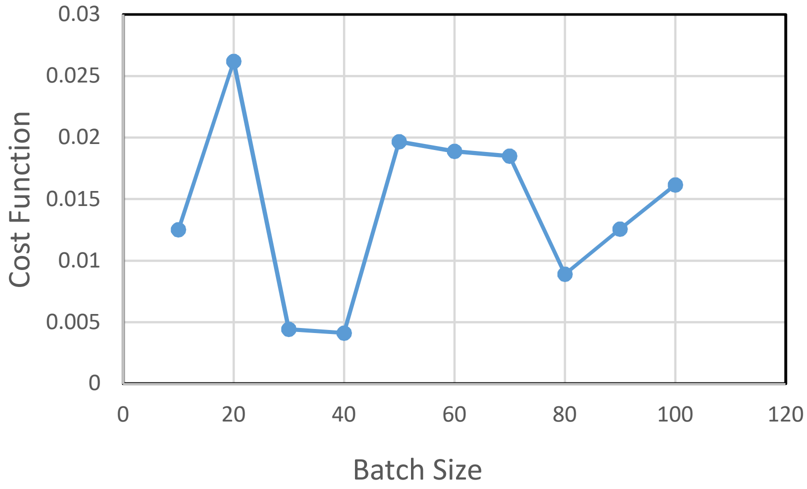


Figure 5.23 Cost Function Vs Batch Size

As can be seen in Figure 5.23, a batch size of 30-40 provides the lowest value of cost function. As we go on increasing the batch size, then the cost function increases. In a study done by Keskar et al. 2016, it has been seen that in practice when a larger batch size was used there is a noteworthy decline in the quality of the model, as seen by its ability to generalize. From the present results it can be seen that this trend continues.

Figure 5.24 shows the effect of batch size on testing and training accuracy.

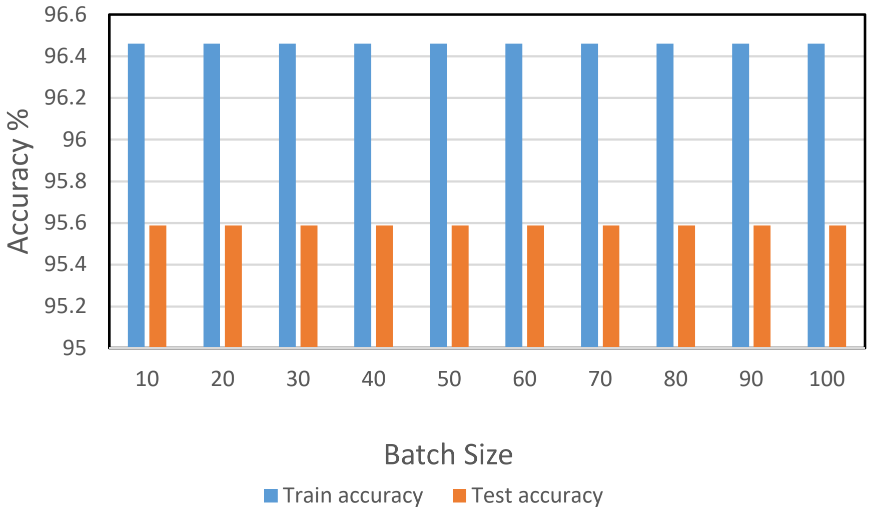


Figure 5.24 Batch Size vs training and test accuracy

As shown in the Figure 5.24, the batch size did not have any affect on the training and testing accuracy. This could be due to the fact that increasing the batch size does not show much effect on the model.

4. Regularization Type: In machine learning, the objective is to focus on the pattern and ignore the noise. If this is not done, then one faces a problem called overfitting. In order to improve the classification accuracy of the model one would like to increase the number of variables. However, at a certain point of time, adding more variables, does not lead to better accuracy, but only tries to fit the noise. A method to decrease overfitting is then to artificially punish higher degree polynomials. This guarantees that a higher degree polynomial is used only if it decreases the error meaningfully related to a simpler model, to overcome the penalty. This is known as regularization. In the present work, two types of regularizations has been used:

1. L1-Type: - Also known as the least absolute deviations (LAD). In this the sum of the differences (S) between the target value (Y_i) and the estimated values ($f(x_i)$), is minimised:

$$S = \sum_{i=1}^{n} |y_i - f(x_i)|. \tag{5.38}$$

2. L2-Type: -Also called the least squares error (LSE). In this, sum of the squares of the differences (S) between the estimated values ($f(x_i)$ and target value (Y_i), is minimised:

$$S = \sum_{i=1}^{n} (y_i - f(x_i))^2 \tag{5.39}$$

Figure 5.25 shows the variation of final cost after training w.r.to type of regularisation.

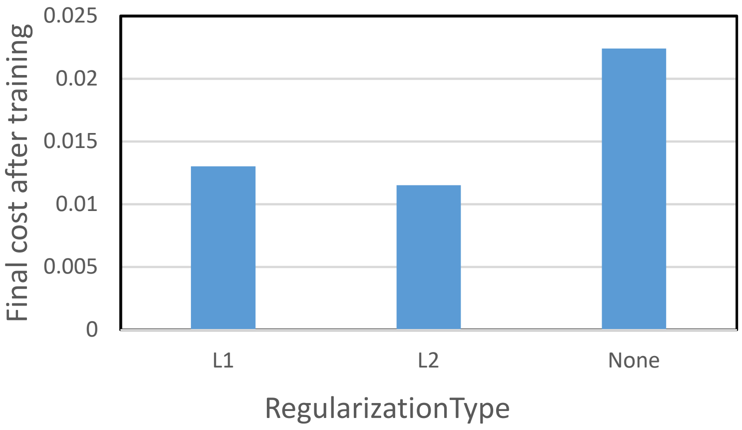


Figure 5.25 Final Cost After Training Vs Regularization Time

As can be seen in the figure 5.25, that L2 regularization showed the best results, as it had least value of cost, while when we had no regularization, the cost function was high. Also, the testing and training accuracy of L2 regularization was the best of the three.

Figure 5.26 shows the variation of training and testing accuracy w.r.to type of regularisation.

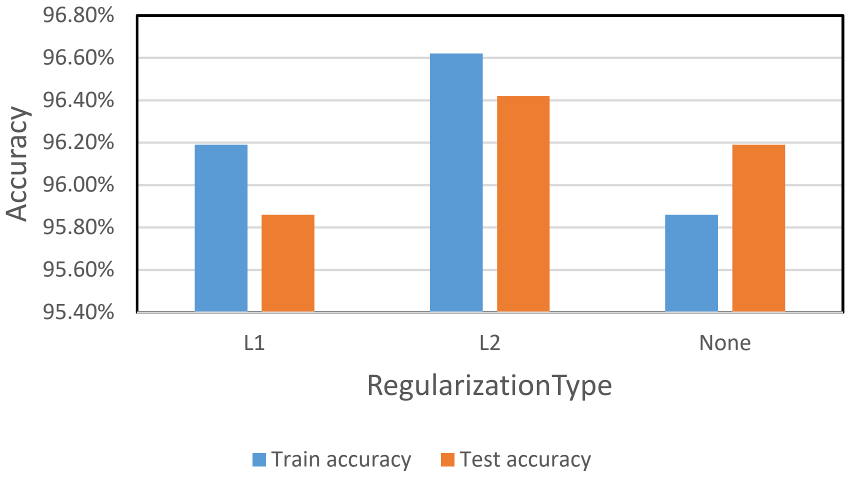


Figure 5.26 Accuracy vs Regularization type

As can be seen in the Figure 5.26 that L2 type regularisation provides the maximum accuracy. This can be attributed to the fact that L1, though known to be robust, could be unstable and might provide multiple solutions. Since the L2 regularization is dealing with square function, which means it is more sensitive, and thus adjusts the model to minimise the error.

Chapter 6

Conclusion and Future Scope of Work

6.1 Conclusion:

In the present work attempt has been made to study biomedical signals and to investigate how automatic detection can be improved. It was recognized that the ECG signals were highly non-stationary in nature. Earlier techniques based on time domain feature extraction, proved to be very simple and could not be successful in feature extraction. Frequency domain analysis could provide better results than time domain, but this method too was not very successful due to non-stationary nature of the signals. The latest techniques being researched is wavelet analysis which belongs to the time-frequency domain feature class. However, the accuracy of the models obtained using these techniques is 90-93 % and it was recognized that in order to improve the accuracy of diagnosis, certain modification in the existing models or the development of new models is required.

In the present work, at first signal features were extracted based on time domain. Then FFT was applied for R-peak detection. This was necessary because R-peaks are easy to detect in and might help in finding the Q peaks, S peaks and T peaks in the signal. It was found that the with the algorithm developed, almost all the R-peaks were detected when the signal was less noisy and not irregular, but it failed for highly noisy and irregular signals.

Next, the signals were analyzed using Wavelet transforms in order to investigate whether the feature extracted from the signal, using wavelet transforms, were able to help in classification of the signals or not. It was found using the feature space diagram developed that the log-energy versus mean, showed real promise in helping to classify the signals. Rest of the features showed a highly overlapping data.

Finally, the features extracted using FFT and Wavelet transform, were used as an input in SVM and ANN. The developed model based on SVM could provide only 50% accuracy. This was because the training and testing dataset used was highly overlapped. Using the Kernel function. did the accuracy reach 50%, or else it might have been worse. Finally, a model was developed using ANN which provided around 96% accuracy. Therefore, ANN model was used to classify the signals. We tested the algorithm on 18 signals, 9 of which were disease and 9 of which were non-disease. The algorithm was able to correctly classify the signals, in almost all the cases.

6.2 Future scope of work

Further scope of work will comprise of:

1. To create a better feature extraction algorithm which will improve the classification result, so that it may work in real time.
2. To investigate the classification accuracy by making use of different classifier so that they can further classify different types of arrythmias.
3. To alter the structure of the network according to cost function of multilayer neural network to achieve better accuracy.
4. Hybrid classifiers can also be used to classify signals. They are different from single classifiers (neural networks, k-Nearest neighbor and Naïve Bayes). They try to find features extracted using different techniques (time based, frequency based, time-frequency based) and try to co-relate them. This might lead to an increase in accuracy.

REFERENCES

A. Cohen, Biomedical Signal Processing. CRC Press, Boca Raton, FL, 1988.

Abdul-Kadir, N., Mat Safri, N. and Othman, M. (2016). Dynamic ECG features for atrial fibrillation recognition. Computer Methods and Programs in Biomedicine, 136, pp.143-150.

Acharya, U.R., Fujita, H., Sudarshan, V.K., Oh, S.L., Adam, M., Koh, J.E., Tan, J.H., Ghista, D.N., Martis, R.J., Chua, C.K. and Poo, C.K., 2016. Automated detection and localization of myocardial infarction using electrocardiogram: a comparative study of different leads. Knowledge-Based Systems, 99, pp.146-156.

AI Trends. (2018). Support Vector Machines (SVM) for AI Self-Driving Cars - AI Trends. [online] Available at: https://aitrends.com/ai-insider/support-vector-machines-svm-ai-self-driving-cars/

B.-U. Köhler, C. Hennig, and R. Orglmeister, “QRS detection using zero crossing counts,” submitted for publication, 2001.

Barrett, K., Barman, S., Boitano, S. and Brooks, H. (2015). Ganong's Review of Medical Physiology 25th Edition. New York: McGraw-Hill Medical Publishing Division.

Brohet, C., Derwael, C. and Fesler, R. (n.d.). Automated ECG diagnosis of atrial flutter by means of wavelet transform. Computers in Cardiology 1994.

Christov, I., Dotsinsky, I. and Daskalov, I. (1992). High-pass filtering of ECG signals using QRS elimination. Medical & Biological Engineering & Computing, 30(2), pp.253-256.

Clayton, R. and Murray, A. (1998). Comparison of techniques for time–frequency analysis of the ECG during human ventricular fibrillation. IEE Proceedings - Science, Measurement and Technology, 145(6), pp.301-306.

Clayton, R., Murray, A., Higham, P. and Campbell, R. (1993). Self-terminating ventricular tachyarrhythmias—a diagnostic dilemma?. The Lancet, 341(8837), pp.93-95.

D. A. Coast, R. M. Stern, G. G. Vano and S. A. Biller, "An approach to cardiac arrhythmia analysis using Hidden Markov Model, " IEEE Trans. Biomed. Eng. ,vol. 37, no. 9,Sep. 1990.

D. B. Geselowitz, _On the theory of the electrocardiogram,_ Proceedings of the IEEE, vol. 77, pp. 857_876, 1989.

D. Beniteza, P.A. Gaydeckia, A. Zaidib, and A.P. Fitzpatrick, “The use of the Hilbert transform in ECG signal analysis,” Computers in Biology and Medicine, vol. 31, pp.399–406, 2001.

Faziludeen, S. and Sabiq, P. ECG beat classification using wavelets and SVM. 2013 IEEE conference on information and communication technologies.

Guillén, S., Arredondo, M., Martin, G. and Ferrero Corral, J. (1990). Ventricular fibrillation detection by autocorrelation function peak analysis. Journal of Electrocardiology, 22, pp.253-262.

Haykin, S. (2009). Neural networks and learning machines. New York: Prentice Hall.

I. Jouny, P. Hamilton, and M. Kanapathipillai, “Adaptive wavelet representation and classification of ECG signals,” in Proc. 16th Annual Int. Conf. IEEE-EMBS, 1994, vol. 2, pp. 1310-1311.

I. R. Legarreta, P. Addison, N. Grubb, G. Clegg, C. Robertson, K. Fox, and J. Watson, R-wave detection using continuous wavelet modulus maxima,_ Computers in Cardiology, 2003, 2003.

Inan, O.T., Giovangrandi, L. and Kovacs, G.T., 2006. Robust neural-network-based classification of premature ventricular contractions using wavelet transform and timing interval features. IEEE transactions on Biomedical Engineering, 53(12), pp.2507-2515.

J. Pan, W. J. Tompkins, “A real time QRS detection algorithm,” IEEE Trans. Biomed. Eng., vol. 32, pp. 230–236, 1985.

Kaman. P.W., Krum. H. and Tonkin. A. M. 1996. Poincare plot of heart rate variability allows quantitative display of parasympathetic nervous activity. Clin. Sci. 91: 201-208.

Kim, P. (2017). MATLAB deep learning. New York (NY): Apress.

Mendez, M.O., Bianchi, A.M., Matteucci, M., Cerutti, S. and Penzel, T., 2009. Sleep apnea screening by autoregressive models from a single ECG lead. IEEE Transactions on Biomedical Engineering, 56(12), pp.2838-2850.

Mendis, S., Puska, P. and Norrving, B. (2011). Global atlas on cardiovascular disease prevention and control. Geneva: World Health Organization in collaboration with the World Heart Federation and the World Stroke Organization.

MIT-BIH Database distribution, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA American National Standard for Ambulatory Electrocardiographs, publication ANSI/AAMI EC38-1994, Association for the Advancement of Medical Instrumentation, 1994.

Moavenian, M. and Khorrami, H. (2010). A qualitative comparison of Artificial Neural Networks and Support Vector Machines in ECG arrhythmias classification. Expert Systems with Applications, 37(4), pp.3088-3093.

Moraes, R., Valiati, J. and Gavião Neto, W. (2013). Document-level sentiment classification: An empirical comparison between SVM and ANN. Expert Systems with Applications, 40(2), pp.621-633.

Mporas, I., Tsirka, V., Zacharaki, E., Koutroumanidis, M., Richardson, M. and Megalooikonomou, V. (2015). Seizure detection using EEG and ECG signals for computer-based monitoring, analysis and management of epileptic patients. Expert Systems with Applications, 42(6), pp.3227-3233.

N.V. Thakor, J.G. Webster and W.J.Thompkins , "Estimation of QRS complex power spectra for design of a QRS filter," IEEE Trans. Biomed. Eng., vol. 31, pp. 702–705, 1984.

P. de Chazal, R.B. Reilly, "A patient-adapting heartbeat classifier using ECG morphology and heartbeat interval feature," IEEE Trans. Biomed. Eng. vol. 53, pp. 2535-2543, 2006.

P.de Chazal, M.O. Duyer, and R.B. Reilly, "Automatic classification of heartbeat using ECG morphology and heart beat interval features," IEEE Trans. Biomed. Eng. vol. 51, pp. 1196-1206, 2004.

Patidar, S., Pachori, R. and Rajendra Acharya, U. (2015). Automated diagnosis of coronary artery disease using tunable-Q wavelet transform applied on heart rate signals. Knowledge-Based Systems, 82, pp.1-10.

R. Poli, S. Cagnoni, and G. Valli, "Genetic design of optimum linear and nonlinear QRS detectors," IEEE Trans. Biomed. Eng., vol. 42, pp. 1137-1141, 1995.

R. Tung and P. Zimetbaum, _Use of the Electrocardiogram in Acute Myocardial Infarction, in Cardiac Intensive Care, 2010, pp. 106_109.

S. Barro, R. Ruiz, D. Cabello, and J. Mira, "Algorithmic sequential decision-making in the frequency domain for life threatening ventricular arrhythmias and imitative artefacts: a diagnostic system," J. Biomed. Eng., vol. 11, pp. 320–328, 1989.

S. Osowski, T. H. Linh, and T. Markiewicz, "Support vector machine based expert system for reliable heart beat recognition," IEEE Trans. Biomed. Eng., vol. 51, no. 4, pp. 582–589, Apr. 2004.

Safara, F., Doraisamy, S., Azman, A., Jantan, A. and Abdullah Ramaiah, A. (2013). Multi-level basis selection of wavelet packet decomposition tree for heart sound classification. Computers in Biology and Medicine, 43(10), pp.1407-1414.

Sejdić, E., Djurović, I. and Jiang, J., 2009. Time--frequency feature representation using energy concentration: An overview of recent advances. Digital Signal Processing, 19(1), pp.153-183.

Silipo, R. and Marchesi, C. (1998). Artificial neural networks for automatic ECG analysis. IEEE Transactions on Signal Processing, 46(5), pp.1417-1425.

Simon Haykin, Neural networks a comprehensive foundation, Pearson Prentice Hall, pp.178-330.

T. Ince, S. Kiranyaz, and M. Gabbouj, "A generic and robust system for automated patient-specific classification of ECG signals," IEEE Trans. Biomed. Eng. vol. 56, pp. 1415-1426, 2009.

T. Olmez, "Classification of ECG waveforms by using RCE neural network and genetic algorithms," Electronics Letters, vol. 33, no. 18, 1997.

Tumer, M., Belfore, L. and Ropella, K. (n.d.). Applying hierarchical fuzzy automatons to automatic diagnosis. 1998 Conference of the North American Fuzzy Information Processing Society - NAFIPS (Cat. No.98TH8353).

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

V. X. Afonso and W. J. Tompkins, "Detecting ventricular fibrillation," IEEE Eng. Med. and Biol., pp. 152–159, Mar. 1995.

V.X. Afonso, W.J. Tompkins, T.Q. Nguyen, and S. Luo, "ECG beat detection using filter banks," IEEE Trans. Biomed. Eng., vol. 46, pp. 192-202, 1999.

W. Jiang and S. G. Kong, "Block-based neural networks for personalized ECG signal classification," IEEE Trans. Neural Network, vol. 18, no. 6, pp. 1750–1761, Nov. 2007.

WHO (2002). World Health Report 2002. Geneva: World Health Organization.

Woo. M.A, Stevenson W.G, Moser, D. K. Trelease. R.B. and Harper. R. H. 1992. Patterns of beat-to-beat heart rate variability in advanced heart failure. Am. Heart J. 123: 704-710.

X. Afonso, W.J. Tompkins, T. Nguyen, S. Luo, “ECG beat detection using filter banks,” IEEE Trans. Biomed. Eng., vol. 46, pp. 230-236, 1999.

Y. Hu, S. Pal Reddy, and W. J. Tompkins, “A patient-adaptable ECG beat classifier using a mixture of experts approach,” IEEE Trans. Biomed. Eng., vol. 44, no. 9, pp. 891–900, Sep. 1997.

Y.C. Yeha, and W. J. Wang, “QRS complexes detection for ECG signals The Difference Operation Method (DOM),” Computer methods and programs in biomedicine, vol. 9, pp. 245–254, 2008.

Y.H. Hu, W.J. Tompkins, J.L. Urrusti, and V.X. Afonso, “Applications of artificial neural networks for ECG signal detection and classification,” J. Electrocardiology, vol. 26 (Suppl.), pp. 66-73, 1993.

Z. Dokur and T. Olmez, “ECG beat classification by a novel hybrid neural network,” Computer Methods and Programs in Biomedicine, vol. 66, pp. 167–181, 2001.

Z. Dokur, T. O. lmez, and E. Yazgan, “Comparison of discrete wavelet and Fourier transforms for ECG beat classification,” Electron. Lett., vol. 35, no. 18, pp. 1502–1504, 1999.

File

ORIGINALITY REPORT

15%	7%	11%	7%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Phil Kim. "MATLAB Deep Learning", Springer Nature, 2017 Publication	2%
2	cseweb.ucsd.edu Internet Source	1%
3	Pal, K., R. Goel, B. Champaty, S. Samantray, and D. N. Tibarewala. "Heart Rate Variability and Wavelet-based Studies on ECG Signals from Smokers and Non-smokers", Journal of The Institution of Engineers (India) Series B, 2013. Publication	1%
4	Vasantha Kalyani David. "Wavelet Transforms and MicroARTMAP", Studies in Computational Intelligence, 2009 Publication	1%
5	www.scribd.com Internet Source	1%
6	www.vutbr.cz Internet Source	1%