

CARDIAC DISEASE ANALYSIS AND DETERMINATION USING DISCRETE WAVELET-BASED ANN

Bavithra Devi.K¹, Muthurani K², Dr.Sivasangari.R³, Dr.Durai Anand.T⁴

^{1,2,3}Department of Electrical and Electronics Engineering, AAA College of Engineering and Technology, Tamilnadu, India

⁴Department of Zoology, VHNSN College (Autonomous), TamilNadu, India

Abstract: *Diagnosis of heart disease is complex. ECG plays an important role in the analysis and diagnosis of cardiac disease. Normally ECG signals are affected by different noises and analysis of those signal is also a tedious process. The main objective of the paper is to de-noise and analyze the ECG signal using DWT (Discrete Wavelet Transform) technique. We use the Daubechies wavelet analysis (level 5) technique in DWT for better performance. Then we obtain the decomposed wave. By analyzing the parameters of the decomposed ECG signal we find the disease present in the heart of the patient by applying the decomposed parameters to ANN (Artificial Neural Network). Thus, we can identify the cardiac disease. This method is very effective because the accuracy of the result is high. Also, the training and testing of the network take very less time. Hence, it is more effective for users.*

Keywords: *ECG, Cardiac disease, Discrete wavelet transform (DWT), Daubechies wavelet analysis, Artificial neural network (ANN)*

I. INTRODUCTION

In recent times, computer-assisted ECG interpretation plays an important role in the automatic diagnosis of heart abnormalities. ECG is the recording of the electrical activity of the heart, and generation of signals. The classification of ECG performance strongly depends on the characterization power of the extracted features from the ECG data and the design of the classifier [3]. The paper describes preprocessing, processing, Feature Extraction, and Classification of ECG signal. An ECG is a linear graphical recording of the electrical impulses that are generated in the heart during the cardiac cycle. The electrical impulses are measured by the electrodes that are placed on the skin. Electrodes which are placed on different sides of the heart measure the activity of different parts of the heart muscle. The ECG displays the voltage between pairs of these electrodes and the muscle activity that they measure. This indicates the overall rhythm of the heart and abnormalities that are present in different parts of the heart muscle [10].

In ECG terms, a lead is a combination of electrodes that form an imaginary line in the body from where the electrical signals are measured [10].

In a 12 lead ECG, three groups of leads can be used, each looking at different aspects of the heart:

- Bipolar limb leads
- Unipolar limb leads
- Unipolar precordial (chest) leads.

Each lead records the electrical signals from the heart of the patient, combination of recording electrodes which are placed at specific points on the patient's body [The bipolar limb leads are known as the lead I, lead II and lead III. They are also known as standard leads [10]. They are placed on each of the patient's arms and legs. These bipolar leads view the frontal plane of the heart from these two points. Like the bipolar leads, unipolar limb leads also record the electrical activity along the heart's frontal plane, but from a different angle [10]. The unipolar precordial leads these leads are placed directly on the chest and view the heart's electrical activity in the horizontal plane.

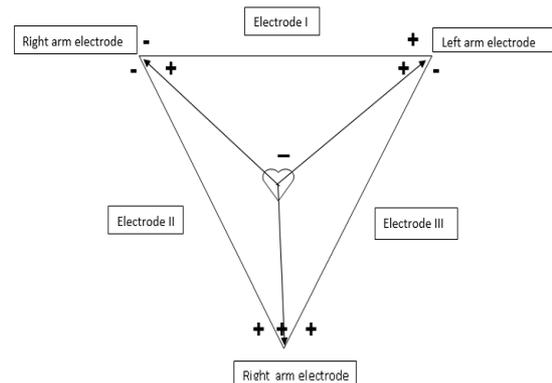


Fig. 1 Placement of ECG electrodes.

II. ECG INTERPRETATION

Now, let us look about the parameters that are to be noticed in an ECG waveform. The following wave shows the normal ECG waveform with the parameters marked in it.

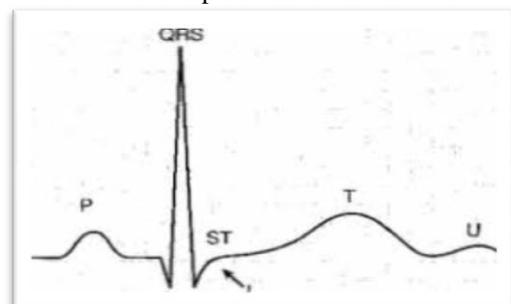


Fig. 2 ECG signal

The P wave is positive in most leads, which means it's above the baseline, on the ECG signal. The PR interval starts from the beginning of the P wave and ends at the beginning of the QRS complex. The QRS complex on an ECG represents the electrical activity associated with the activation of the heart's ventricles, it may have three components:

- The Q Wave- first negative deflection on the complex.
- The R Wave- first positive deflection of the complex
- The S Wave- first negative deflection after the R wave
- The T wave occurs in the same direction as that of the QRS complex [10].

The below-given figure shows the ECG waveform generated using the Matlab code for Normal condition.

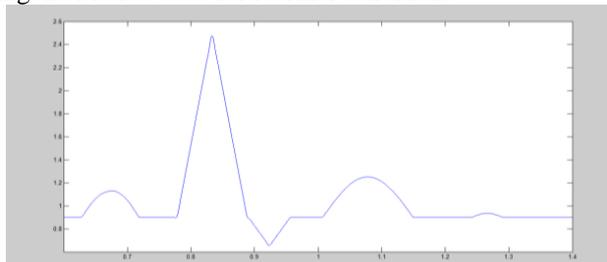


Fig. 3 Normal ECG waveform generated using MATLAB
 Table 1 shows the origin, amplitude, duration of the wave parameters of ECG signal [11].

Table 1. Description of wave parameters.

Parameters and Origin	Amplitude (mV)	Duration(Sec)
P wave Atrial Depolarisation or contraction	0.25	0.12 to 0.22 (P-R interval)
R wave(QRS complex) Depolarisation of atria and depolarisation of the ventricles	1.60	0.07 to 0.1
T wave Ventricular repolarisation(relaxation of myocardium)	0.1 to 0.5	0.05 to 0.15 (S-T interval)
ST interval Ventricular contraction		
U wave Slow repolarisation of the intraventricular	<0.1	0.2(T-U interval)

III. DIFFERENT CARDIAC DISEASES AND ECG SIGNALS GENERATION

There are many types of cardiac diseases. In this paper, there are six diseases analyzed. The diseases are First-degree atrial ventricular block, bundle block, myocardial infarction, coronary insufficiency, ventricular fibrillation, and atrial fibrillation. The following are the ECG signals for the above-mentioned types of diseases.

First Degree AV Block

Here PQ segment is prolonged conduction time(i.e)greater than 0.22 sec[11].

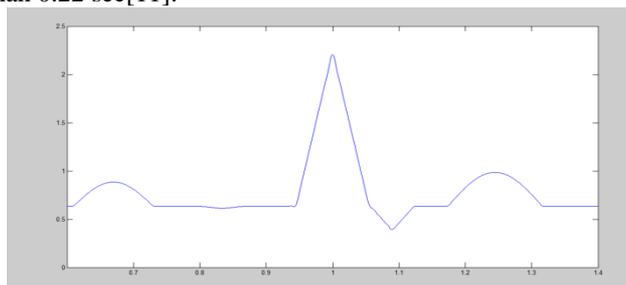


Fig. 4 Waveform generated in MATLAB for First AV block

Bundle Block

Here QRS complex is widened (i.e)QRS interval is greater than 0.1 sec [11].

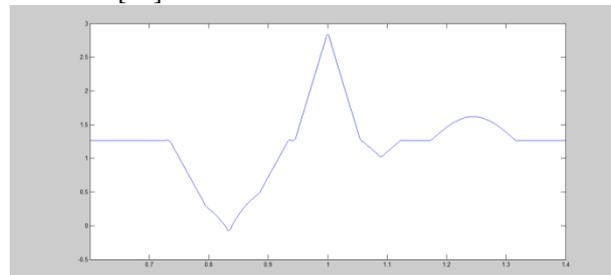


Fig 5 Waveform generated in MATLAB for Bundle block

Myocardial Infarction

Here ST segment is elevated[11].

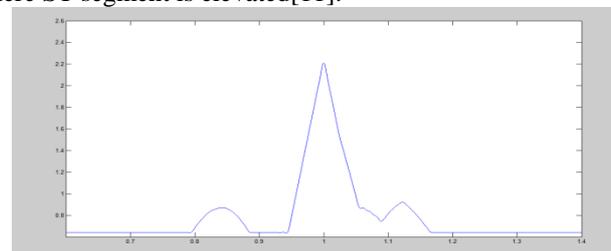


Fig 6 Waveform generated in MATLAB for Myocardial Infarction

Coronary Insufficiency

Here the ST segment is depressed and negative T wave is present[11].

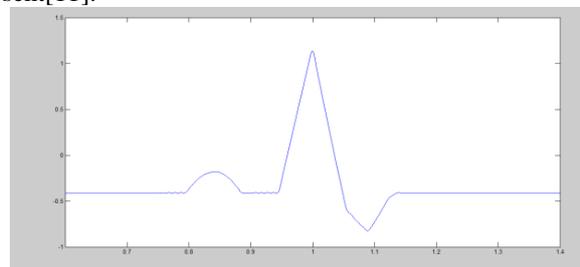


Fig. 7 Waveform generated in MATLAB for Coronary Insufficiency

Ventricular Fibrillation

Due to the fast beating rate of the ventricles.No pumping of the blood to different parts of the body.Here there is a train of pulses instead of PQRST waves[11].

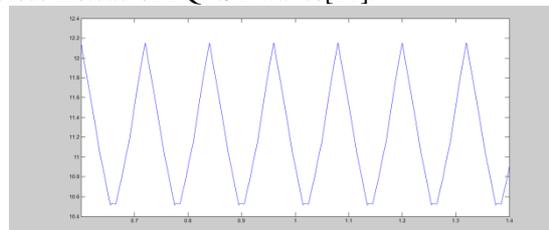


Fig. 8 Waveform generated in MATLAB for Ventricular Fibrillation

Atrial Fibrillation

Due to the fast beating rate (300-500 beats/min) of the atrium. Here ventricles beat very slowly[11].

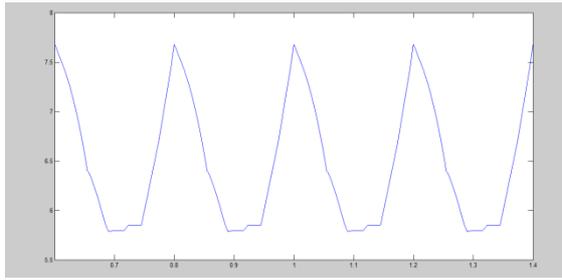


Fig. 9 Waveform generated in MATLAB for Atrial Fibrillation

IV. METHODOLOGY

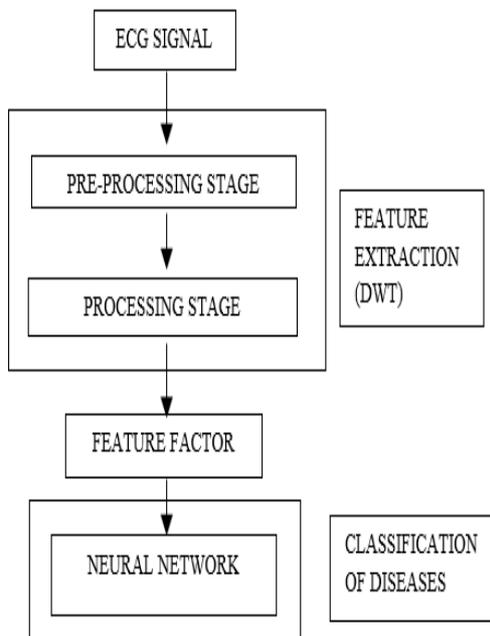


Fig. 10 Flow model of analysis of ECG signal

The above flow chart deals with the flow of identification of the cardiac disease from the generated ECG signal. The generated ECG signal from MATLAB is pre-processed and then processed using the discrete waveform transformation. Here we use the Daubechies Wavelet analysis technique (level 5) for DWT. Now the parameters are extracted from the analyzed waveforms. These are then fed in to the Artificial Neural Network (ANN). ANN is trained and tested, which results in the proper classification of diseases like first degree AV block, coronary insufficiency, myocardial infarction, bundle block, atrial fibrillation, and ventricular fibrillation.

V. DISCRETE WAVELET TRANSFORM

The discrete wavelet transform is performed by repeated filtering of the input signal by using two filters they are a low pass filter (LPF) and a high pass filter (HPF). These filters decompose the signal into different scales. The output coefficient that is gained from the low pass filter is the approximation coefficient.

The scaling function output is usually in the form of:

$$\Phi(t) = 2^{-h} \sum_{q \in \mathbb{Z}} \Phi(2t - q) M^q = 0$$

The output of the high pass filter is the detailed co-efficient.

The wavelet function output is in the form of:

$$w(t) = 2^{-g} \sum_{q \in \mathbb{Z}} M^q \Phi(2t - q)$$

The approximation coefficient is usually divided into new approximation and detailed coefficients. By choosing the mother wavelet the coefficient of such filter banks are calculated. This decomposition process is repeated until the required frequency response is achieved from the given input signals [7].

VI. DAUBECHIES WAVELET

The Daubechies wavelets, a type of DWT, supports the work of Ingrid Daubechies, which belongs to the family of orthogonal wavelets which defines the discrete wavelet transform and it is characterized by a maximum number of vanishing moments for some given data. For each type of wavelet of this class, there is a scaling function called the *father wavelet*, which generates an orthogonal multiresolution analysis of the Daubechies wavelet transforms, they are defined in the same way as that of Haar wavelet transform by computing running averages and their differences via scalar products with scaling signals and wavelets. The only difference between them exists in how these scaling signals and wavelets are defined accordingly [7].

Original signal = $a_1 + d_1 + d_2 + d_3 + d_4 + d_5$

The Daubechies D5 transform has five wavelet and scaling function co-efficients. The scaling function co-efficients are:

- $\{h_0 = 0.0033303573,$
- $h_1 = -0.125807520,$
- $h_2 = 0.0062414902,$
- $h_3 = 0.0775714938,$
- $h_4 = -0.0322448696,$
- $h_5 = 0.2422948871,$
- $h_6 = 0.1384281459,$
- $h_7 = 0.7243085284,$
- $h_8 = 0.6038292698$
- $h_9 = 0.1601023981\}$

The wavelet function co-efficient values are:

- $\{g_0 = -h_9 ; g_1 = -h_8 ; g_2 = h_7 ; g_3 = -h_6 ; g_4 = h_5 ; g_5 = h_4 ; g_6 = -h_3 ; g_7 = h_2 ; g_8 = h_1 ; g_9 = -h_0\}$

Daubechies D scaling function:

$$a_i = h_{0s_{2i}} + h_{1s_{2i+1}} + h_{2s_{2i+2}} + h_{3s_{2i+3}} + h_{4s_{2i+4}} + h_{5s_{2i+5}} + h_{6s_{2i+6}} + h_{7s_{2i+7}} + h_{8s_{2i+8}} + h_{9s_{2i+9}}$$

$$a[i] = h_{0s[2i]} + h_{1s[2i+1]} + h_{2s[2i+2]} + h_{3s[2i+3]} + h_{3s[2i+4]} + h_{3s[2i+5]} + h_{3s[2i+6]} + h_{3s[2i+7]} + h_{3s[2i+8]} + h_{3s[2i+9]}$$

Daubechies D5 Wavelet function:

$$c_i = g_{0s_{2i}} + g_{1s_{2i+1}} + g_{2s_{2i+2}} + g_{3s_{2i+3}} + g_{4s_{2i+4}} + g_{5s_{2i+5}} + g_{6s_{2i+6}} + g_{7s_{2i+7}} + g_{8s_{2i+8}} + g_{9s_{2i+9}}$$

$$c[i] = g_{0s[2i]} + g_{1s[2i+1]} + g_{2s[2i+2]} + g_{3s[2i+3]} + g_{3s[2i+4]} + g_{3s[2i+5]} + g_{3s[2i+6]} + g_{3s[2i+7]} + g_{3s[2i+8]} + g_{3s[2i+9]}$$

VII. DB ANALYSIS OF ECG SIGNAL

The following figure shows the steps that are followed for the analysis of the ECG waveform.

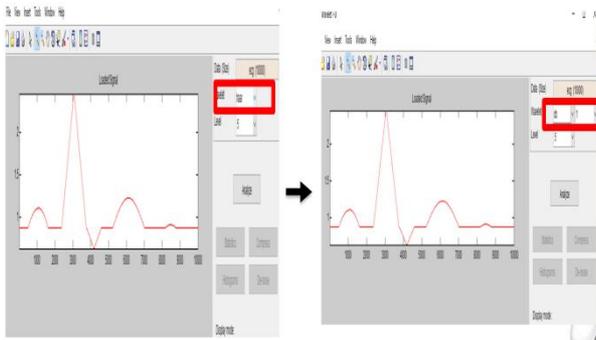


Fig. 11 Selection of DB analysis in MATLAB

The below figure shows the decomposition of the ECG signal at various levels. Here we use level 5 Daubechies wavelet analysis. Here S denotes the signal, a_5 denotes the approximation of the signal, d_1 d_2 d_3 d_4 d_5 denotes the decomposition of the noise level of the waveform in 5 levels.

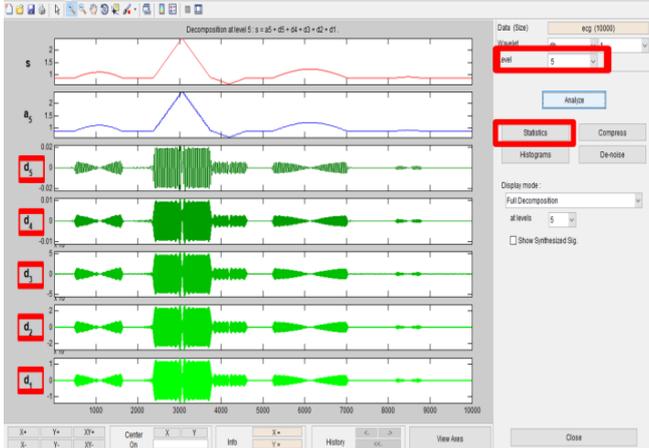


Fig 12 Denoising of ECG signal

The following figure shows the approximated waveform with the analyzed parameters displayed in it.

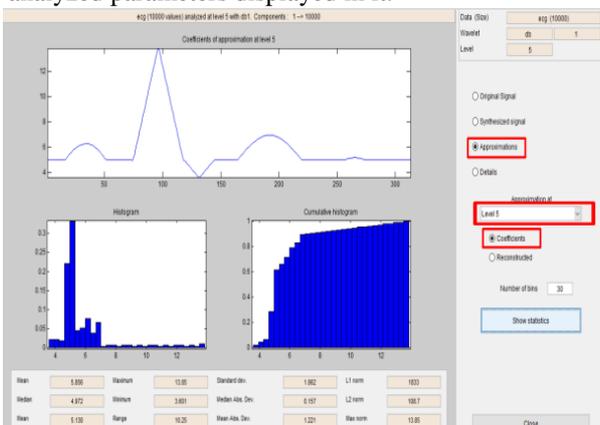


Fig. 13 Approximation of ECG signal

The parameters are extracted from the analyzed waveform. The parameters that are concentrated in the paper are mean, median, maximum, minimum, range, standard deviation.

The following figure shows the extracted parameters from the decomposed signal.

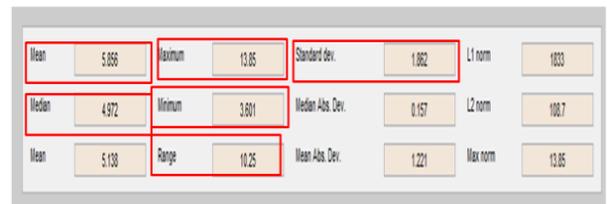


Fig. 14 Derivation of wave parameters in MATLAB

The following table shows the data obtained from the decomposed signal

Table 2: Extracted parameters for each disease

S.No	Types of Cardiac Disease	No. of waveforms generated	No. of parameters for each wave	Total parameters
1	Normal ECG waveform	30	6	30*6=180
2	First degree AV Block	30	6	30*6=180
3	Bundle block	30	6	30*6=180
4	Myocardial Infarction	30	6	30*6=180
5	Coronary insufficiency	30	6	30*6=180
6	Ventricular fibrillation	30	6	30*6=180
7	Atrial fibrillation	30	6	30*6=180
Total		210 waves	-	1260

VIII. ARTIFICIAL NEURAL NETWORK (ANN)

The Neural networks are been seen as an explosion of interest over the last few years and are being successfully applied across the world. NN can handle errors better than traditional computers programs (imagine a scenario where one faulty statement in program extraordinary range of problem domains), in areas as diverse as finance, medicine, engineering, geology, and physics. If problem structure is well analyzed, traditional computers could still outperform NN but in cases where the problem has not been analyzed in details, NN could be used to learn from a large set of examples could halt everything while NN can handle errors in a much better manner. The following figure shows a simple model of neural network working [8].The major advantages of ANN are as follows.

1. Adaptive learning: It is the ability to learn to do a task using the given training data.
2. Self-Organization: ANN can easily organize itself using the training data.
3. Real-Time Operation: Computations may be carried out in a parallel manner.

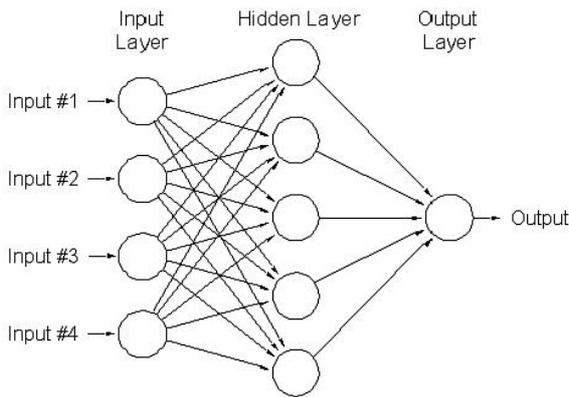


Fig. 15 Simple presentation of ANN

IX. TESTING AND TRAINING OF ANN

After the extraction of the parameters from the decomposed ECG signal, they are fed into the Artificial Neural Network. In the ANN the data is used to train the neural network and test it. The usage of extracted data is as follows.

- Total sets =210
- No of sets used for training of ANN=140
- No of sets used for testing of ANN =70

The ANN is trained and tested in different layers like an input layer, hidden layer, and the output layer. Here we use the number of layers as follows.

Input layer =6
 Output layer =1
 Hidden layer =1250

The following figure shows the usage of layers in the training and testing of ANN.

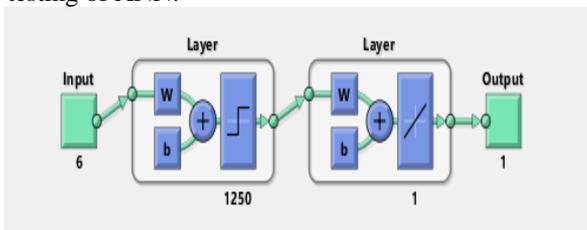


Fig. 16 Layers in ANN during training in MATLAB

The accuracy of training is 98%.The time elapsed is 86.327643 seconds.

X. FEEDFORWARD BACK PROPAGATION NETWORK

The most popular algorithm of ANN is the backpropagation algorithm. Rojas [2005] said that the BP algorithm can be broken down into four main major steps. The network weights are chosen randomly, the backpropagation algorithm is used to compute the required corrections. BP algorithm can be decomposed as follows:

- i) feed-forward computation
- ii) backpropagation to the output layer
- iii) backpropagation to the hidden layer
- iv) weight updates

The algorithm stops when the value of the error function has become negligible [8].

- *Feed-Forward Computation* Feed forward computation or forward pass is a two-step process.

The first part is to get the values from the hidden layer nodes and the second part is using those values derived from the hidden layer to compute value or values for the output layer [8].

- *Back Propagation to the Output Layer:*Once when an error is known, it will be used for backward propagation and weights adjustment. It is a two-step process. Firstly the error gets propagated from the output layer to the hidden layer. Here learning rate and momentum are brought together [8].
- *Back Propagation to the Hidden Layer:*Now the errors are propagated from the hidden layer down to the input layer. It becomes a bit more complicated than propagating error from output to the hidden layer. Once the error of hidden layer nodes is known, then weights between input and the hidden layer can be known easily [8].
- *Weight Updates:*The most important thing is not to update any weights until all errors have been properly calculated. It is easy to forget them and if the new weights are used while calculating errors, results would not be proper [8].

There are many learning functions to train ANN. Here we use the function 'learngdm' to train ANN. The features of the function are as follows.

Purpose

Gradient descent with momentum weight and bias learning function

Syntax

```
[dw,ls]=learngdm(w,p,z,n,a,t,e,gw,ga,d,lp,ls)
[Db,ls]=learngdm(b,ones(1,q),z,n,a,t,e,gw,ga,d,lp,ls)
info = learngdm(code)
```

Description

learngdm is the gradient descent with momentum weight and bias learning

Algorithm

learngdm calculates the weight change dw for a given neuron from the neuron's input P and error E, the weight (or bias) W, learning rate LR, and momentum constant MC, according to gradient descent with momentum: $Dw = mc * dw_{prev} + (1 - mc) * lr * gw$

The previous weight change dwprev is stored and read from the learning state LS [9].

XI. RESULTS AND DISCUSSION

The following are the performance, training state and regression characteristics which are obtained after the training and testing of ANN

Performance characteristics

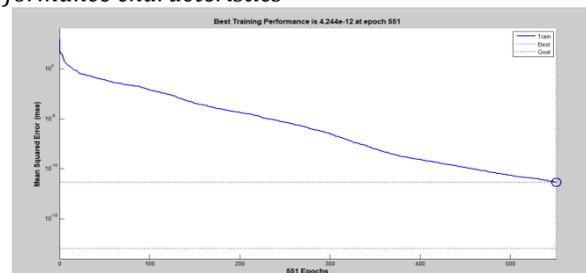


Fig 17 Performance characteristics obtained in MATLAB

Training state characteristics

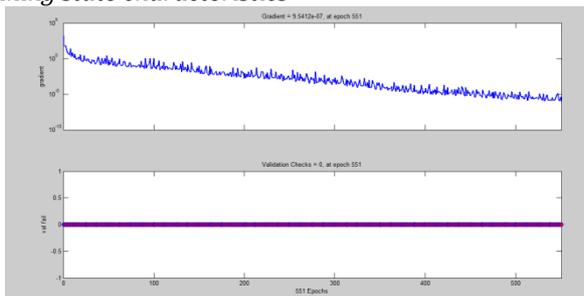


Fig 18 Training state characteristics obtained in MATLAB

Regression characteristics

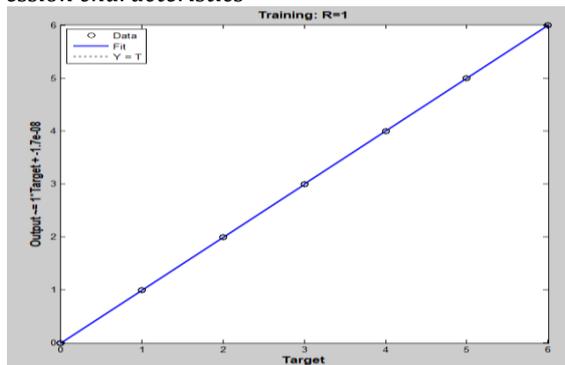


Fig. 19 Regression characteristics obtained in MATLAB

The following figure shows the output obtained after test data is applied to the trained ANN.

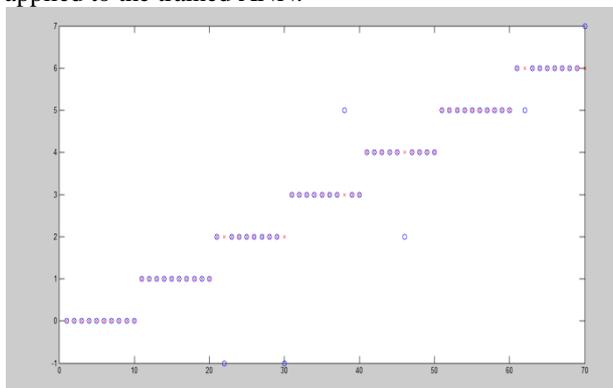


Fig 20. Test Data Results

From the above graph, we infer that there are totally 70 data tested. Also, we can clearly view that out of 70 data tested, 64 data output is accurate. Thus the accuracy of the ANN output is 91.4%.

We can improve the accuracy by further continuing the process in ANFIS.

REFERENCES

[1] S.M.Jadhav, S.L.Nalbalwar, A.A.Ghatol, "Artificial Neural Network Models based Cardiac Arrhythmia Disease Diagnosis from ECG signal Data". *International Journal of Computer Applications* Volume 44, Issue 15, PP-0975-8887, Year 2012.

[2] A.Dallali, A.Kachouri, M.Samet, "Integration Of HRV, WT And Neural Networks For ECG Arrhythmia Classification". *ARNP Journal of Engineering and Applied Sciences* Volume 6, Issue 5, PP-1819-6608, Year 2011.

[3] S.P.Kulkarni, Dr. K.V. Kulhalli, "DWT and ANN Based Heart Arrhythmia Disease Diagnosis from MIT-BIH ECG Signal Data", Volume 3, Issue 1

[4] A.Ghaffari, A.Jalali, C. Nataraj, "Heart Arrhythmia Detection Using Continuous Wavelet Transform and Principle Component Analysis With Neural Network Classifier". *Computing in cardiology* Volume 37, PP-669-672, Year 2010.

[5] Mohit Kumar, Ram Bilas Pachori, U.Rajendra Acharya, "Characterization of coronary artery disease using flexible analytic wavelet transform applied on ECG signals". *Biomedical signal processing and control* Volume 31, year 2017.

[6] Siva A, Hari Sundar M, Siddharth S, Nithin M and Rajesh CB, Classification of Arrhythmia using Wavelet Transform and Neural Network Model", April 4, 2018

[7] Mangala Gowri S G, Dr.Cyiril Prasanna Raj P, "EEG Feature Extraction using Daubechies Wavelet and Classification using Neural Network". *International Journal of Pure and Applied Mathematics*, Jan 18, 2018

[8] Cilimkovic M (2015) Neural networks and back propagation algorithm. Institute of Technology Blanchardstown <http://dataminingmasters.com/uploads/studentProjects/NeuralNetworks.pdf>

[9] Howard Demuth, Mark Beale, Martin Hagan, "Neural Network Toolbox" http://cda.psych.uiuc.edu/matlab_pdf/nnet.pdf.

[10] "Basic ECG Interpretation Learning Package", Bendigo health, 2009, <http://www.bendigohealth.org.au/Content/Docs/ECG%20learning%20package.pdf>.

[11] "Biomedical Instrumentation", Dr. M. Arumugam, Anuradha Publication, 1994, Edition 2.