



Genetic Evolution Based Cardiac Arrhythmia Classifier

¹N.Ikram, ²K.Ramya, ³K.S.Vairavel

¹Assistant Professor, Department of EIE

Bannari Amman Institute of Tech, sathyamangalam, Tamilnadu, India

²Assistant Professor, Department of EIE

Bannari Amman Institute of Tech, sathyamangalam, Tamilnadu, India

³Assistant Professor, Department of EIE

Bannari Amman Institute of Tech, sathyamangalam, Tamilnadu, India

¹ikram@bitsathy.ac.in, ²ramyak@bitsathy.ac.in, ³vairavelks@bitsathy.ac.in.

Abstract— Some of the major diseases that have a high impact on the society are Cardio Vascular Diseases (CVDs). An important category of CVDs are the Cardiac Arrhythmias. Investigation of such signals have drastically helped cardiac experts to identify health related problems more effectively. This paper proposed that the performance of ECG signal can be analyzed using Genetic evolutions (Genetic Algorithm and Genetic Programming) and Artificial Neural Network (ANN). The ECG signal must be clearly represented and filtered to remove all noise and artifacts from the signal. DWT is a powerful tool for ECG signal enhancement, different wavelet transforms used to enhance the ECG signal and choose the best wavelet from analyzing performance measures. This enhanced signal helpful for classification task. Various parameters are extracted from denoised signal. Irrelevant or noisy features unnecessarily increase the complexity of the problem and can degrade classification performance. GP and GA is used only for best feature selection from a large features data set. The algorithm is used to effectively select a smaller subset of features that together form a genetically fit family for classification tasks using ANN. The databases are extracted from MIT-BIH Database (Massachusetts Institute of technology/ Beth Israel A. Hospital).

Index Terms—ECG, MIT-BIH, wavelet transform, ANN

I. INTRODUCTION

ECG signal indicates the electrical activity of the human heart. It offers cardiologists with helpful information regarding the rhythm and functioning of the heart. Medical attention is required for patients with abnormal morphology and heart rate in ECG signals since these abnormal cardiac rhythms may lead to life threatening situations. ECG signals are acquired by placing electrodes across the thorax or chest of human body for

a limited time and the electrical recordings are visualized using an external device. The electrodes pick the electrical signals generated by the polarization and depolarization of heart muscle. The flat line (isoelectric line) indicates the absence of

electrical activity while a deviation from the isoelectric line indicates the electrical action of heart muscles. ECG signal used to find regular and irregular cardiac cycles of patients.

The signal consists of baseline wander, power line interference and high frequency noise. It must be clearly represents by removing noise from the ECG signal.

II. MATERIALS

All the experiments in this work are carried out by considering public database namely MIT-BIH Arrhythmia which is available freely on Physionet. The database includes different types of ECG recordings. In this work, data with a broad range of normal and pathological ECG recordings are considered to evaluate the algorithm performance. Classes taken for this work are Normal, Premature Ventricular Contraction (PVC), Bundle Branch Block (BBB) and Atrial premature Contraction (APC).

A. MIT-BIH Arrhythmia Database (MITBIHAR)

The MIT-BIH Arrhythmia Database contains 48 records which are recorded for 30 minutes. Approximately 60% of these recordings are obtained from inpatients. In most records, one signal is a modified limb lead II (MLII), obtained by placing the electrodes on the chest. While the other is usually a modified lead V1. Normal QRS complex wave is essential in the first part of signal. The lead axis for the second signal may be almost orthogonal to the mean cardiac electrical axis. The recordings are digitized at the rate of 360 Samples per second per channel with the resolution of 11 bit over the range 10 mV.

III. METHODOLOGY

The Proposed method aims to classify the beats into normal, premature ventricular contraction, bundle branch block, atria premature contraction for a patient record. In this work Pan-Tompkins's algorithm used to find peaks of every beats. The next step is to train the ANN system by deriving ECG feature vectors from the signal sample beats. Feature selection done by Genetic Algorithm (GA) and Genetic Programming (GP), and classification task performed using Artificial Neural Network (ANN).

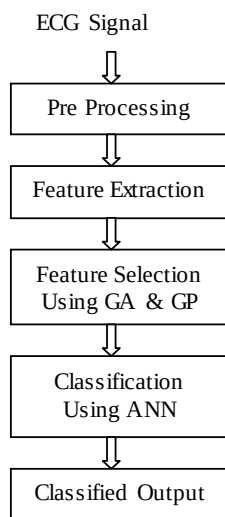


Fig 1. Work flow of Proposed Method

1. Pre processing

Wavelet transform is an efficient tool for breaking down non stationary ECG signals. The wavelet transform is used to extricate morphological data from the ECG information. Wavelet transformation maps the ECG signal into a time-scale plane. The fourth scale has the good projection between the frequency ranges 12.25 to 22.5Hz [2][15]. It is worth to use the WT which can be efficiently implemented as a filter bank. The WT function of the signal is given as

$$\int_{-\infty}^{\infty} f(x) \varphi^*_{a,b} x dx \quad (1)$$

* -denotes complex conjugate

1.1 Wavelet types

Symlet, Haar, Biorthogonal and Daubechies wavelets are used for denoising the signal. Among these wavelets chosen the best wavelet, it depends on the performance measures [11].

A. Biorthogonal Wavelet

In Biorthogonal wavelet, the wavelet transform is invertible. The degree of freedom for biorthogonal wavelets is more when compared to the orthogonal wavelets. The

symmetric wavelet functions is one of the additional degree of freedom

Where,

ϕ, θ - scaling functions in biorthogonal cases
 M and N - scaling sequences of a, a^*

$$\sum_{n \in \mathbb{Z}} a_n a_n + 2m = 2 \cdot \delta_{m,0} \quad (2)$$

Biorthogonal wavelet scaling function can be defined as

$$b_n = -1^n a_{M-1-n} \quad (n = 0, \dots, N-1) \quad (3)$$

$$b^*_{n^*} = -1^{n^*} a_{M-1-n^*} \quad (n^* = 0, \dots, N-1) \quad (4)$$

B. Haar Wavelet

Haar wavelet is also the elementary potential wavelet. The practical complexity of the haar wavelet is that it is not continuous, and therefore it is not differentiable.

The Haar mother wavelet function $\varphi(t)$ can be defined

$$\varphi(t) = \begin{cases} 1 & 0 \leq t < 1/2 \\ -1 & 1/2 \leq t < 1 \end{cases} \quad (5)$$

The scaling function of Haar can be described as

$$\phi(t) = \begin{cases} 1, & 0 \leq t < 1, \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

C. Daubechies Wavelet

The Daubechies wavelets, in view of crafted by Ingrid Daubechies, are a group of symmetrical wavelets characterizing a discrete wavelet change and portrayed by a maximal number of vanishing moments for some given help. With every wavelet sort of this class, there is a scaling capacity which creates a symmetrical multiresolution examination [9].

The scaling function of Daubechies wavelet can be defined

$$b_k = -1^k a_{N-1-k} \quad (7)$$

Compare the different wavelet transforms in terms of performance for MIT-BIH, choose the best wavelet for feature extraction to detect R-peaks, RR interval, QRS points, ST segment, Q amp and S amp in the ECG waveforms since those are going to help the classification.

2. Feature extraction

Feature extraction is extracting and converting the input data into a set of features which called feature vector, by reducing the data representation pattern [3] [12]. The features set will extract the relevant information from the input data in order to perform the classification task. The transform of a

signal is just another form of representing the signal. It does not alter the information content present in the signal.

To find QRS points Pan-Tompkins's algorithm is used. The following steps explain how to obtain the QRS points in beat [10][6]

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

Step 2: Reconstruct the signal only using approximate coefficients.

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

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

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

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

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

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

Step 9: Using that points find the QRS duration.

The gating signal set as 1 for the QRS duration, when multiplied with the denoise ECG signal, gives a signal that shows only the QRS points alone while all the others are masked. The maximum value points in this signal denote as R-peaks while the minima on the left and right of R-peaks denote as Q and S points respectively. To extract the P and T points, the gating signal set as 0 is multiplied with the filtered ECG. Doing so masks the QRS complex while only P and T waves remain there. P wave lies in a range of 120-200 ms before the QRS onset so a peak lying in this range marks the P-peak and the two minima on its either side mark the onset and offset thereof. The positive peak following the QRS offset is the T peak and the minima on left and right of the peak are the T-onset and T-offset respectively. From these points extracted a number of parameters that were applied to classification.

TABLE 1.Features extracted parameters

S.No	Features	Description
1	QRS Complex	Due to contraction of ventricles occurs duration between 80-120 ms

2	R-amp	Depolarization activity of atria's
3	S-T segment	Depolarization activity of ventricles occurs between 80-120ms (approx)
4	S amp	Downward deflection of QRS complex occurs after R-wave
5	RR interval	Interval between two consecutive R-amp
6	Q-amp	Depolarization of the inter ventricular septum

3. Classification using ANN

Artificial Neural network (ANN) used as a tool for classification problems in many fields. The feed forward network architecture such as Multi Layer Perceptron (MLP), RBF network is commonly used in pattern recognition problems [1][7]. The patterns are given as input to the input layer which propagates in the forward direction from one layer to another layer. The MLP networks are trained by back propagation algorithm which employs supervised learning technique. A well-known error correction algorithm used in MLP network is the Least Mean Square (LMS) algorithm. The MLP network consists of two passes in the network. In the forward pass, the input patterns are applied to the input neurons and the signal propagates in the forward direction resulting in a group of outputs (actual response). In the backward pass, the error correction rule will determine the synaptic weights and the error signal propagates in the direction opposite to the synaptic weight connection. The weights are modified so that the actual output approaches the desired output.

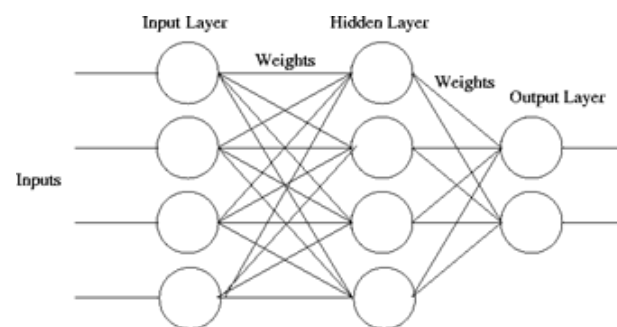


Fig 2.Basic structure of ANN model

An Artificial Neural Network (ANN) is an organized group of nodes, akin to the immeasurable network of neurons in a brain. Here, each node represents an artificial neuron and an arrow represents a connection between previous output, present input and present output.

A. Levenberg-Marquardt (Lm) Algorithm

According to Levenberg-Marquardt (LM)

optimization, It is a network training function that updates weight and bias values. It is one of the fastest back propagation algorithm and is highly suggested as a first-choice supervised algorithm, despite the fact that it needs more memory than any other algorithms.

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

$$w(k+1) = w(k) - (J^T J + \alpha I)^{-1} J^T \epsilon(k) \quad (8)$$

Where

J - Jacobian matrix,
 α - Adjustable parameter,
 ϵ - Error vector.

The parameter is modified based on the development of error function E . If the step causes a reduction of E , otherwise reset the original value and recalculate $w(k+1)$.

4. Feature selection

Reasons for doing feature selection are that the amount of features can make construction of a model harder. Removing redundant and irrelevant features can help improve the performance of classifier. This is done by reducing the potential for over-fitting and, in the case of redundant features, selecting those that work best [14].

4.1 Feature selection using Genetic Algorithm

Genetic Algorithm (GA) is derived from the natural evolution process of a man; it generally works on the population based and algorithmic search heuristic method. The various operators of GA such as reproduction crossover, mutation, selection which helps to produce a new population, GA to choose a subset of features, a fitness function must be clearly defined to evaluate the discriminative potential of each subset of features [5]. The fitness of each chromosome in the population is evaluated by means of kNN-based fitness function.

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

$$D(x_{test}, x_i) = \sqrt{\sum_{m=1}^M (x_{test} - x_i)^2} \quad (9)$$

The kNN count each category min the class information (accumulated as $(count\ x_m)$ using 3 Nearest

Neighbors and then report classification results and errors based on the

$$\arg \max (count(x_m)) \quad (10)$$

Subjected to

$$\sum_{i=1}^M count(x_m) = class \quad (11)$$

The iterations were repeated, In order to reduce the error rate and pick the individual with the least fitness value

Since error rate is reported for every chromosomes involved in process and the minimum of error rate is finally chosen by the GA.

$$fit = \frac{\alpha}{N_f} + \exp\left(\frac{-1}{N_f}\right) \quad (12)$$

α = kNN-Based classification error.

N_f = Cardinality of the selected features.

The arithmetical formation of above equation ensures the learning of the GA, error minimization and concentrated on the number of features selected which immensely reduces the error, with the application of GA for dimensionality reduction, more discriminating features were obtained.

TABLE 2.Parameters of Genetic Algorithm (GA)

Parameter	Value
Maximum number of individuals	1920
Generation	100
Crossover probability	0.8
Mutation probability	0.2
Fitness evaluation	kNN-Classification error
Selection scheme	Tournament

4.2 Feature selection using Genetic Programming

Genetic programming based feature selection was used to improve the classification results and reduce the dimensionality of the data. Genetic programming as a form of evolutionary algorithm and an extension of genetic algorithms is used as the feature selection[4][8]. The major difference between the GP and GA approaches lies in the way that each algorithm solves the problem under consideration. With a GA-based solution, the basic form of the solution is predefined; the GA is able to optimize parameters of the solution, however not the actual structure of the solution.

Fitness evaluation is calculated by same as genetic algorithm evaluation. Features can be reduced by rank selection.

Steps of Genetic Programming:

The implementation of GP involves the following steps

Initialization: Genetic Programming begins with randomly generated population of solutions of size P .

Termination: GP is terminated when termination criteria are satisfied. GP is terminated when a desired solution (may be with fitness value 1) is achieved.

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

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

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

TABLE 3.Parameters of Genetic Programming (GP)

Parameter	Value
Maximum number of individuals	1920
Generation	100
Type of cross over	One point cross over
Type of mutation	Point mutation
Fitness evaluation	kNN-Classification error
Selection scheme	Tournament
Cross over probability	0.5
Mutation Probability	0.5
Maximum tree depth	5

IV. RESULTS AND DISCUSSION

The results and analysis of various features that are extracted are discussed as follows

a. Denoising performance measures

Original ECG is denoised and obtain clean ECG signal by using different wavelet transforms. The performance measures are calculated for Symlet, Haar, Biorthogonal and Daubechies wavelets.

TABLE 4. Denoising performance measures of MIT-BIH

Performance measures	Haar	Biortho	Symlet	Db
Variance	0.55483	0.77618	0.16646	0.22365
MSE	3.4341	3.9491	3.3811	3.3424
MAE	1.8029	1.8106	1.9933	1.9964

Table 4.Shows the performance measures (MSE, MAE and Variance) of Haar, Biorthogonal, Symlet and Daubechies wavelets. Haar wavelet gives higher performance than other wavelet for MIT-BIH.

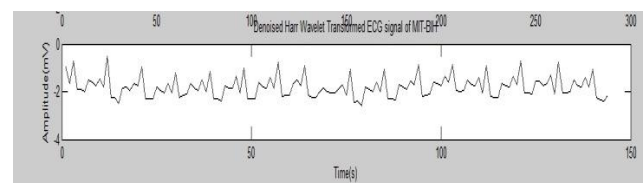


Fig 3. Denoised ECG signal using haar wavelet

b. Feature extraction

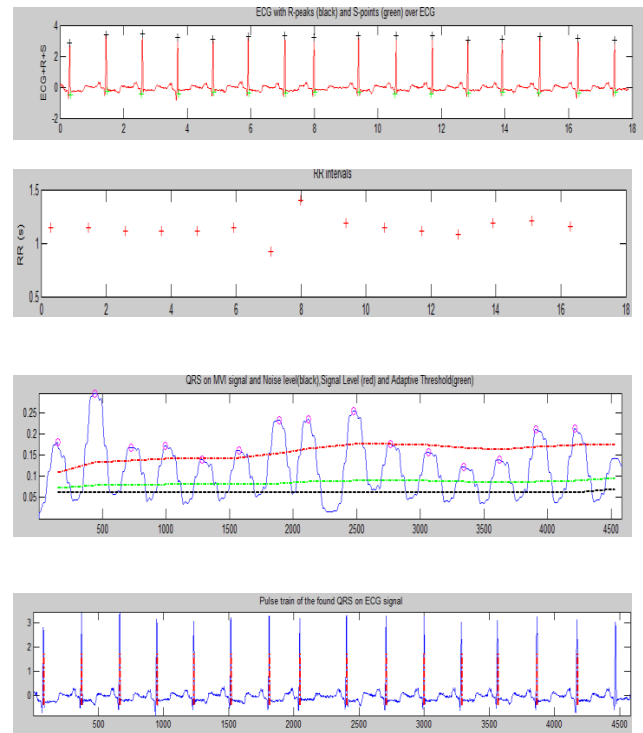


Fig 4. Extracted parameters R-amp, RR interval, QRS complex using pan tompkins's algorithm

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

c. Classification using ANN

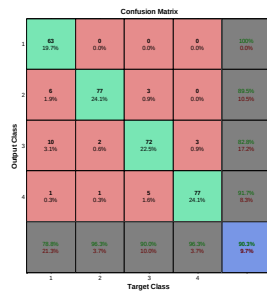


Fig 5. Performance obtained from confusion matrix for ANN

Figure 5. Shows the confusion matrix, classified and misclassified samples for four classes and its performance of ANN classifier.

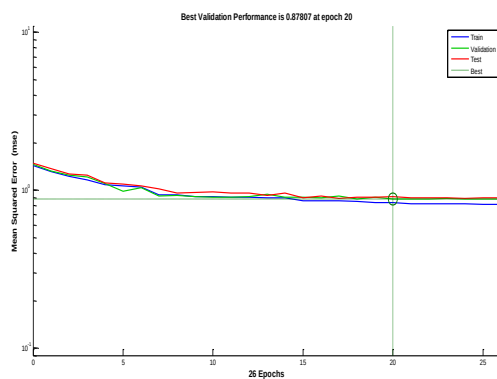


Fig 6. Best validation performance for ANN

Figure 6. Shows the best validation performance of 0.8707 at epoch 20 of ANN classifier.

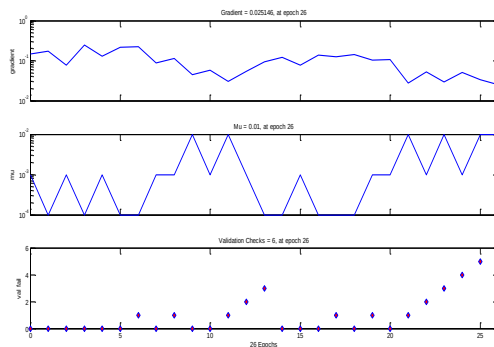


Fig 7. Best training state performance for ANN

Figure 7. Shows the best training state performance of gradient 0.025146 at epoch 26 of ANN classifier.

TABLE 5. Performance measures of ANN

	Mean Squared Error(MSE)
Train Performance	0.8340
Test Performance	0.9114
Validation Performance	0.8781
Overall performance	0.8515

Table 5. Shows the Mean Squared Error (MSE) for training, testing and validation performance of ANN classifier.

d. ANN classifier using GA based features

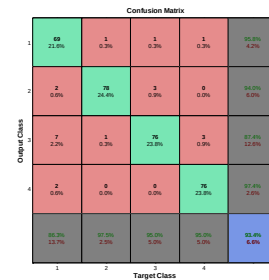


Fig 8. Performance obtained from confusion matrix for GA based ANN

Figure 8. Shows the confusion matrix, classified and misclassified samples for four classes and its performance of GA based ANN classifier.

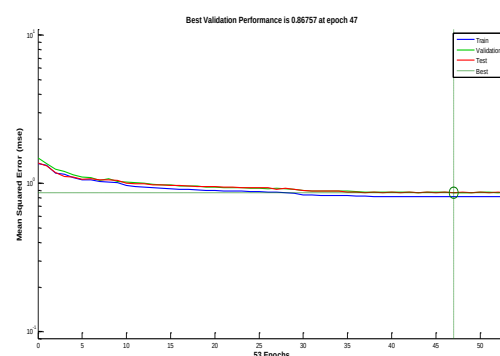


Fig 9. Best validation performance for GA based ANN

Figure 9.Shows the best validation performance of 0.88767 at epoch 47 of GA based ANN classifier.

	Mean Squared Error(MSE)
Train Performance	0.8138
Test Performance	0.8677
Validation Performance	0.8676
Overall performance	0.8293

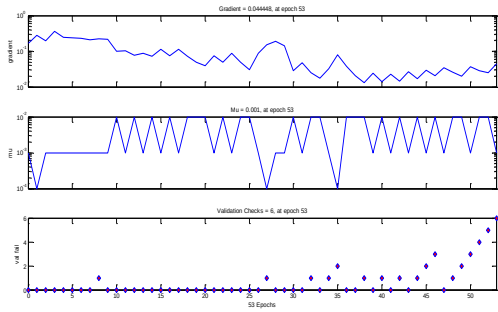


Fig 10.Best training state performance for GA based ANN

Figure 10.Shows the best training state performance of gradient 0.044448 at epoch 53 of GA based ANN classifier.

TABLE 6.Performance measures of GA based ANN

Table 6.Shows the Mean Squared Error (MSE) for training, testing and validation performance of GA based ANN classifier.

e. ANN classifier using GP based features

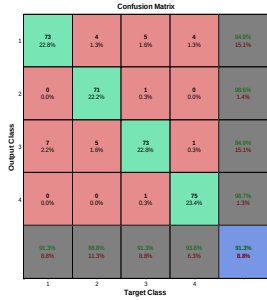


Fig 11.Performance obtained from confusion matrix for GP based ANN

Figure 11.Shows the confusion matrix, classified and misclassified samples for four classes and its performance of GP based ANN classifier.

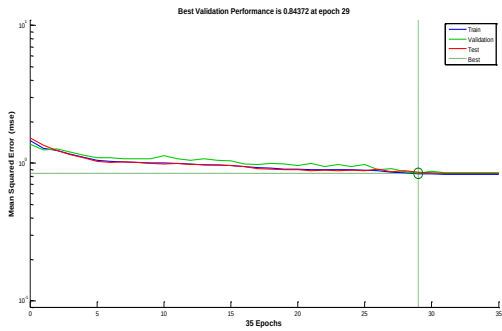


Fig 12.Best validation performance for GP based ANN

Figure 12.Shows the best validation performance of 0.84372 at epoch 36 of GP based ANN classifier,

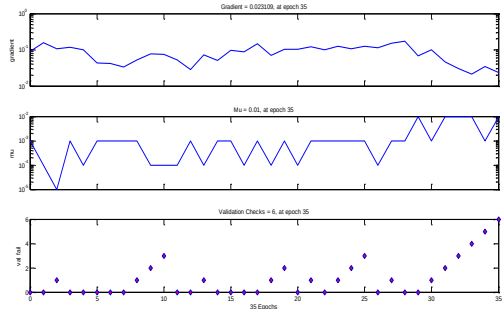


Fig 13.Best training state performance for GP based ANN

Figure 13.Shows the best training state performance of gradient 0.023109 at epoch 35 of GP based ANN classifier.

TABLE 7.Performance measures of GP based ANN

	Mean Squared Error(MSE)
Train Performance	0.8375
Test Performance	0.8556
Validation Performance	0.8437
Overall performance	0.8399

Table 7.Shows the Mean Squared Error (MSE) for training, testing and validation performance of GP based ANN classifier.

TABLE 8.Overall Performance comparison measures

Performance measures	ANN	GA with ANN	GP with ANN
----------------------	-----	-------------	-------------

Accuracy	90.3%	93.4%	91.3%
Sensitivity	91.2%	94.7%	92.4%
Specificity	78.6%	81.2%	80.6%
Mean squared error	0.8515	0.8123	0.8399

Table 8. Shows the performance measures (Accuracy, Specificity, Sensitivity and MSE) of ANN classifier, GA with ANN classifier, GP with ANN classifier. GA with ANN classifier gives higher performance than other two approaches.

V.CONCLUSION AND FUTURE WORK

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

REFERENCES

- [1] Osowski S., Markiewicz T. and Tran Hoai L., "Recognition and classification system of arrhythmia using ensemble of neural networks", *Measurement*, Vol. 41, No. 6, pp. 610-617, 2008.
- [2] Tsipouras M. G. and Fotiadis D. I., "Automatic arrhythmia detection based on time and time- frequency analysis of heart rate variability", *Computer Methods and Programs in Biomedicine*, Vol. 74, No. 2, pp. 95-108, 2004.
- [3] de Chazal P. and Reilly R. B., "A Patient- Adapting Heartbeat Classifier Using ECG Morphology and Heartbeat Interval Features"
- [4] Pooja Gupta and Sulochana Wadhwani "Feature Selection by Genetic Programming, Artificial Neural Network-based Machine Condition Monitoring" *International Journal of Engineering and Innovative Technology*, Vol 1, No 4, 2012.
- [5] Babatunde Oluleye, Armstrong Leisa and Jinsong Leng "A Genetic Algorithm-Based Feature Selection" *International Journal of Electronics Communication and Computer Engineering*, Vol 5, No 4, 2010.
- [6] S.Karpagachelvi, M.Arthanari and M.Sivakumar "ECG Feature extraction technique - A survey approach" *International Journal of Computer Science and Information Security*, Vol 4, No 1, 2010.
- [7] Cardiac Arrhythmia Classification Using Neural Networks with Selected Features - *International Conference on Computational Intelligence: Modeling Techniques and Application*, Vol 5, No 4, 2013.
- [8] Masood Ahmad Arbab, Gul Muhammad Khan and Ali Mahmud Sahibzada "Cardiac Arrhythmia Classification using Cartesian Genetic Programming Evolved Artificial Neural Network", Vol 20, No 9, 2013.
- [9] Krishna Prasad and J. S. Sahambi, "Classification of ECG Arrhythmias using Multi Resolution Analysis and Neural Networks," *IEEE Transactions on Biomedical Engineering*, vol. 1, pp. 227-231, 2003.
- [10] Y. H. Hu, S. Palreddy, and W. Tompkins, "A Patient Adaptable ECG Beat Classifier Using A Mixture Of Experts Approach", *IEEE Transactions on Biomedical Engineering* vol. 44, pp. 891-900, 1997.
- [11] S. T. Sanamdikar, S. T. Hamde and V. G. Asutkar, "Extraction of different features of ECG signal for detection of cardiac arrhythmias by using wavelet transformation Db 6," *2017 International Conference on Energy, Communication, Data Analytics and Soft Computing India*, pp. 2407-2412, 2017
- [12] A. E. Vincent and K. Sreekumar, "A survey on approaches for ECG signal analysis with focus to feature extraction and classification," *2017 International Conference on Inventive Communication and Computational Technologies*, pp. 140-144, 2017
- [13] R. R. Perez, A. Marques and F. Mohammadi, "The application of supervised learning through feed-forward neural networks for ECG signal classification," *2016 IEEE Canadian*

Conference on Electrical and Computer Engineering, Vancouver, BC, pp. 1-4, 2016.

[14] T. Mar, S. Zaunseder, J. P. Martínez, M. Llamedo and R. Poll, "Optimization of ECG Classification by Means of Feature Selection," in IEEE Transactions on Biomedical Engineering, vol. 58, no. 8, pp. 2168-2177, Aug. 2011.

[15] Y. R. Rao and N. DevadasNaik, "Stationary and non-stationary noise removal from cardiac signals using CSLMS," 2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering, pp.1208-1212, 2017

