

---

# Optimization of Multi-Layer Perceptron Neural Network Using Genetic Algorithm for Arrhythmia Classification

V. S. R. Kumari<sup>1, \*</sup>, P. Rajesh Kumar<sup>2</sup>

<sup>1</sup>Departments of Electronics and Communication, Research scholar, Andra University, Vishakhapatnam, India

<sup>2</sup>Departments of Electronics and Communication, Andra University, Vishakhapatnam, India

## Email address:

vsrk46@gmail.com (V. S. R. Kumari), rajeshauce@gmail.com (P. R. Kumar)

## To cite this article:

V. S. R. Kumari, P. Rajesh Kumar. Optimization of Multi-Layer Perceptron Neural Network Using Genetic Algorithm for Arrhythmia Classification. *Communications*. Vol. 3, No. 5, 2015, pp. 150-157. doi: 10.11648/j.com.20150305.21

---

**Abstract:** An Electrocardiogram (ECG) graphically records changes in electrical potentials between different sites on the skin due to cardiac activity. The heart's electrical activity is a depolarization and depolarization sequence. ECGs help in identifying cardiac arrhythmia because they have diagnostic information. ECG arrhythmia detection accuracy improves by using machine learning and data mining methods. This study proposes multi-layer perceptron neural network optimization using Genetic Algorithm (GA) to classify ECG arrhythmia. Symlet extracts RR intervals from ECG data as features while symmetric uncertainty assures feature reduction. GA optimizes learning rate and momentum.

**Keywords:** Arrhythmia Classification, Electrocardiogram (ECG), RR Interval, MIT-BIH ECG Dataset, Multi-layer Perceptron Neural Network, Genetic Algorithm (GA)

---

## 1. Introduction

Cardiac arrhythmias are disturbances in heart rhythm, manifested by irregular or abnormally fast rates (tachycardias) or abnormally slow rates (bradycardias). Patients with such abnormalities frequently observe palpitations, which is described as a sensation of 'my heart turning over in my chest', or being aware of their hearts beating rapidly/slowly. Other symptoms are weakness, light headedness, shortness of breath, dizziness, fainting (syncope) and occasionally, chest pain. Symptoms are severe when rate is faster, ventricular function is worse, or arrhythmia is linked to autonomic tone abnormalities. But, many arrhythmia patients report no symptoms, with the condition being discovered during routine check-ups. A rapid and long tachyarrhythmia can produce cardiomyopathy and congestive heart failure. In such cases, arrhythmia treatment ensures normal function of ventricles.

Though arrhythmia's physical signs help physician make correct diagnosis, electrocardiography is standard method used for recognizing cardiac arrhythmias. A prolonged electrocardiographic recording, often called a 'Holter monitor', or an event recorder that the patient activates when sensing an abnormality, may assist in confirming the diagnosis when the arrhythmia occurs sporadically, as is

often the case [1].

Classification of ECGs into different disease categories is a complex pattern recognition task. However, the analysis of ECG signals is the most effective available method to diagnose cardiac arrhythmias [2]. Many methods are proposed for ECG signals classification. Research presented in [3]-[8] are among the recently published works. The method in [3] is based on Fisher Linear discriminant where RR interval duration and distance between occurrence of P and T waves is perceived. Fisher's Linear Discriminant is applied using these features. In [4] a SVM-based method for PVC arrhythmia detection is seen as more efficient than Anfis. In [5] a new approach for feature selection and classification of cardiac arrhythmias based on PSO-SVM is proposed. [6] Describes a neuro-fuzzy approach to ECG-based heart rhythms classification where Hermite polynomials characterize QRS complex signals whose coefficients feed neuro-fuzzy classifier. Arrhythmia detection by Independent Component Analysis (ICA) and Wavelet transform proposes to extract important features in [7]. Finally, in [8], authors present an approach to classify beats of a large dataset by training a Neural Network (NN) classifier using wavelet/timing features. The authors found that fourth scale of a dyadic wavelet transform with a quadratic spline wavelet together with the pre/post RR-interval ratio is effective in distinguishing the normal and

PVC beats from others.

Machine learning algorithms note data changes adapting its design to accommodate new findings. It is applied in ECG classification with most works being based on NN [9, 10], Markov chain model [11] and Support Vector Machine (SVM) [12]. The Wavelets [13] and morphology [14] based methods are common feature extraction techniques.

Artificial Neural Network (ANN) is used as a base in computer based diagnostic systems to design structural designs. Improved ANN techniques like data reduction and feature extraction are presented in literature [15, 16]. MLP can recognize /classify ECG signals better than other ANN methods.

Neural networks are classification tools with research proving they are alternatives to conventional classification. Their advantages include being self-adaptive and data driven and adjusting to data without functional specification/distributional form. They accurately approximate functions. Classification requires intra group member's functional relationships, and object attributes for accurate function identification. NN are nonlinear and flexible in modelling real world complex relationships while NN estimates posterior probabilities to provide classification rule establishment and provide a base for statistical analysis.

NN are based on numeric optimization from the gradient descent method. By using different optimization methods such as PSO, GA etc., various algorithms have been established for training of NN. The inclusion of momentum leads to improvement in the presentation of gradient descent, and presents a second parameter  $\mu$ , additionally to that of the learning rate parameter  $\eta$ .

This study proposes optimization of fuzzy NN using GA to classify ECG arrhythmia. Symlet extracts RR intervals from ECG data as features while symmetric uncertainty ensures feature reduction. GA optimizes learning rate and momentum. Final evaluation is by a MIT-BIH arrhythmia database.

## 2. Literature Review

ECG arrhythmias classification using Type-2 Fuzzy Clustering Neural Network (T2FCNN) was proposed by Ceylan, et al., [17] where T2FCNN architecture for classification of electrocardiography arrhythmias was presented. The idea behind using T2FCNN clustering algorithm was reduction of neural network classification error by optimization training pattern set. Data for the study was from Physionet database belonging to MIT-BIH ECG arrhythmia database. Finally, the new T2FCNN structure classified ECG arrhythmias with a 99% detection rate.

An arrhythmia Beat Classification using Pruned Fuzzy K-Nearest Neighbor (PFKNN) Classifier was proposed by Arif, et al., [18]. This PFKNN classifier was used to classify various arrhythmia types and different beats presented in MIT-BIH Arrhythmia database. The classifier was tested on ~103100 beats for six beat types presented in database. It achieved 97% beat classification accuracy and geometric sensitivity mean was 94.5% with 19% of total training

examples. Accuracy/sensitivity was comparable to FKNN when all training data was used.

An arrhythmias Classification with MLP NN and Statistical Analysis was proposed by Raut and Dudul [19] which presented a classification system for cardiac arrhythmias using ANN with back propagation algorithm. The work concluded that the proposed MLPNN classifier estimated complex decision boundaries correctly had remarkable discriminating ability and outperformed statistical discriminant analysis and classification tree rule based prediction.

A Hybrid system for cardiac arrhythmia classification with fuzzy k-nearest neighbors and Multi-Layer Perceptron along with a fuzzy inference system was proposed by Ramirez, et al., [20]. This described hybrid architecture for cardiac arrhythmias classification taking MIT-BIH Arrhythmia database ECG records as source. The suggested method introduced a Mamdani fuzzy inference system combining each classifier's output achieving a high classification rate of 98%.

An Evolvable Block-based Neural Networks for real-time heart arrhythmia classification from ECG signals was proposed by Nambiar, et al., [21]. The author suggested a method to classify heart arrhythmia from ECG signals by using Block-based Neural Networks (BbNN). This was used in the problem's hardware implementation due to its block based structure; relatively fast computational speed, and reduced resource consumption. BbNN system on chip showed high arrhythmia classification accuracy averaging 99.64% for tested patient records.

An Arrhythmia Classification Using Hybrid Networks was proposed by Haseena, et al., [22] which focused on a Fuzzy C-Mean (FCM) clustered Probabilistic Neural Network (PNN) and Multi Layered Feed Forward Network (MLFFN) for discrimination of eight ECG beats types. Parameters like fourth order Auto Regressive (AR) coefficients with Spectral Entropy (SE) were extracted from ECG beats and featured reduction was tried out using FCM clustering

An arrhythmia disease diagnosis using NN, SVM and genetic algorithm-optimized k-means clustering was proposed by Martis and Chakraborty [23]. This method presented a ECG based arrhythmia disease detection using GA optimized k-means clustering. Open-source ECG data from MIT-BIH arrhythmia database and MIT-BIH normal sinus rhythm database were subjected to a steps sequence including segmentation using R-point detection, extraction of features using PCA and pattern classification. It concluded that GA-optimized k-means algorithm was an alternative to classification whose accuracy was near supervised (EBPNN/SVM) classifiers.

A methodology for automated creation of fuzzy expert systems for ischemic and arrhythmic beat classification based on a rules set from a decision tree was proposed by Exarchos, et al., [24]. The methodology for automated fuzzy expert systems creation was applied in ischemic and arrhythmic beat classification. The suggested system automatically created a fuzzy expert system from initial training dataset ensuring

high accuracy and ability to interpret decisions.

Wang, et al., [25] presented a feature reduction method by integrating principal component analysis (PCA) with Linear Discriminant Analysis (LDA). Classification of eight different types of arrhythmia was done using a Probabilistic Neural Network (PNN) classifier from ECG beats. The feature reduction improved average classification accuracy by 99.71%.

Zadeh, et al., [26] investigated premature ventricular contraction recognition from normal beats. The new system included three modules: de-noising, feature extraction and classifier modules. Many supervised classifiers like multi-layer perceptron neural networks, SVM with various kernel types and PNN were used.

### 3. Methodology

The Arrhythmia database of MIT-BIH contains 48 half hour recordings of two channel ambulatory ECG from 47 subjects in 1975 and 1979 by Boston’s Beth-Israel Hospital Arrhythmia Laboratory. Twenty-four hour ambulatory ECG recordings were collected from a 4000 mixed population having inpatients (60%) and outpatients (40%). Recordings were digitized at 360 samples per second per channel with 11-bit resolution over a 10mV range. Two or more cardiologists annotated each record independently; a consensus was taken to get computer-readable reference annotations for all beats [27].

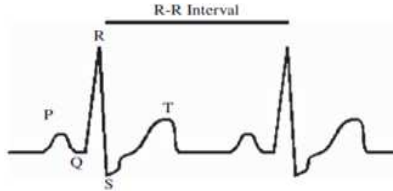


Figure 1. Structural components of an ECG signal.

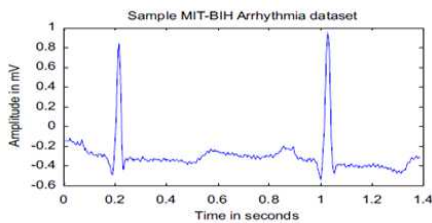


Figure 2. Sample MIT-BIH Arrhythmia dataset.

ECGs are inexpensive and non-invasive means to observe the heart’s physiology. In 1961, Holter introduced techniques for continuous ECG recording for ambulatory subjects over many hours; long-term ECG (Holter recording) with 24 hours duration is the standard technique for observing transient cardiac electrical activity aspects. From mid-1970s, cardiac rhythm abnormalities (arrhythmias) are reflected in long-term ECGs and automated methods to identify arrhythmias was used. Though recordings are plentiful, access to data is not, and recorded waveforms thorough characterization is tedious and expensive. Further, there is a wide variability in ECG

rhythms and waveform morphology details, both between subjects and within individuals over time. Hence, a representative collection of long-term ECGs for research must have many recordings [28].

R-R intervals in ECG are extracted using Symlet, and the extracted features are reduced using symmetric uncertainty in the proposed method. Features are classified as normal beats, LBBB and RBBB using proposed Fuzzy neural network. GA optimized learning rate and momentum. Various techniques used in the methodology are detailed in this section.

#### 3.1. R-R Interval

To determine an ECG signal’s RR- interval, R-waves should be detected in obtained ECG signal. As R-waves have largest amplitudes among waves it is easily detected. Specific detail components of decomposed signals are kept to detect R waves and other low frequency/ high frequency components are removed.

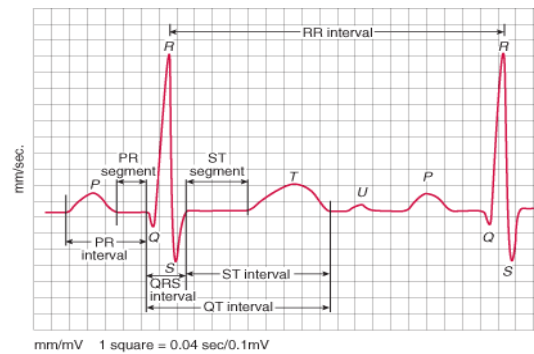


Figure 3. R-R interval pattern in ECG.

R-wave detecting includes low frequency baseline elimination and R-waves identification/detection on cleaned-up signal. Baseline removal assures correct R-wave amplitude measurements. Experiments reveal this as necessary to prevent missing R-wave in baseline’s high slope portions. Baseline removal is through low pass, zero phase shifts, moving average filter. An 800 ms window length with a 40 ms sampling interval yielded a 3 dB point of 0.3Hz.

#### 3.2. Symlet Wavelet

A Wave is an oscillating function of time or space, Wavelets are localized waves and they have their energy concentrated in time or space. The Transform of a signal is another form of representing the signal. It does not change the information content present in the signal. The Wavelet Transform provides a time- frequency representation of the signal and is well suited to the analysis of non-stationary signals [29] such as ECG. A Wavelet Transformation uses multi resolution technique by which different frequencies are analyzed with different resolutions. A Wavelet Transform, at high frequencies, gives good time resolution and poor frequency resolution, while at low frequencies the Wavelet Transform gives good frequency resolution and poor time resolutions.

Wavelets are waveforms bound in both frequency and time. Wavelet analysis splits the signal into shifted and scaled versions of the original (or mother) wavelet. The Continuous Wavelet Transform (CWT) is given by the wavelet function  $\psi$  by adding all time of the signal multiplied by scaled, shifted versions. Mathematically the continuous wavelet is defined by

$$(scale, position) = \int_{-\infty}^{+\infty} f(t)\psi(scale, position, t)dt \quad (1)$$

Many wavelet C coefficients are scale and position function due to CWT. The original signal's constituent wavelets are got through multiplying a coefficient by applicable scaled/shifted wavelet. Daubechies proposed symlets are symmetrical wavelets. Symlets are db family modifications. Both wavelet families are similar with difference of db wavelets having maximal phase whereas symlets have minimal phase. Symlets are compactly supported with slight asymmetry. Wavelet coefficient for symlet is any positive even number and highest number of vanishing moments for a specific support width [30].

### 3.3. Symmetric Uncertainty (SU)

Another method of feature selection is choosing features with highest symmetric uncertainty (SU) values between feature and target classes [31]. To locate this indicator, feature values for zero mean and unit variance are first normalized by

$$\hat{x}_{ijk} = \frac{x_{ijk} - \bar{x}_k}{\sigma_k}, \sigma_k^2 = \frac{1}{CN} \sum_{j=1}^c \sum_{i=1}^n (x_{ijk} - \bar{x}_k)^2 \quad (2)$$

Then continuous features normalized values are discretized into L finite levels to facilitate locating probabilities. Corresponding discrete values are  $x_{ijk}$ . Mutual information of kth feature is

$$I(X_k, \omega) = \sum_{l=1}^L \sum_{j=1}^c P(\tilde{x}_{ljk}, \omega_j) \log_2 \frac{P(\tilde{x}_{ljk}, \omega_j)}{P(x_{ljk})P(\omega_j)} \quad (3)$$

where  $P(\omega_j)$  is prior probability of class of  $\omega_j$ ,  $P(\tilde{x}_{ljk})$  is distribution of k<sup>th</sup> feature, and,  $P(\tilde{x}_{ljk}, \omega_j)$  the joint probability. This indicator measures how much feature values distribution and target classes differ from statistical independence. This is nonlinear correlation estimation between feature values and target classes. Symmetric uncertainty (SU) is derived from mutual information by normalizing it to feature values entropies and target classes

$$SU(x_k, w) = 2 \left[ \frac{I(x_k, w)}{h(x_k, H(w))} \right] \quad (4)$$

where entropy of variable X is found by  $H(x) = P(x_i) \log 2P(x_i)$

## 4. Feature Extraction Using Fuzzy Clustering Algorithms

Features are extracted using Fuzzy C-means algorithm from Symlet coefficients. PCA Fuzzy clustering solves problems in pattern recognition and fuzzy model identification. Various fuzzy clustering methods are proposed with most being based on distance criteria. A widely used algorithm is Fuzzy C-Means (FCM) which uses reciprocal distance to compute fuzzy weights. A better algorithm is new FCFM which computes cluster center using Gaussian weights, resorts to large initial prototypes, adding processes of eliminating, clustering and merging. The following sections discuss/compare FCM algorithm and FCFM algorithm.

### 4.1. The Fuzzy C-Means Algorithm

The Fuzzy C-Means (FCM) algorithm [32] uses weights that minimize total weighted mean-square error

$$J(w_{qk}, z(k)) = \sum_{k=1, K} \sum_{k=1, K} (w_{qk}) \|x(q) - z(k)\|^2 \quad (5)$$

$$\sum_{(k=1, K)} (w_{qk}) = 1 \text{ for each } q$$

$$W_{qk} = (1 / (D_{qk})^2)^{1 / (p-1)} / \sum_{k=1, K} (1 / (D_{qk})^2)^{1 / (p-1)}, \quad p > 1 \quad (6)$$

FCM allows feature vector to belong to all clusters with a fuzzy truth value (between 0 and 1), computed using Equation (6). The algorithm assigns feature vector to clusters based on maximum feature vector weight over other clusters.

### 4.2. PCA

PCA identifies data patterns, expressing them to highlight similarities and differences. As data patterns are hard to find for high dimension data, where graphical representation is unavailable, PCA is a powerful tool for data analysis. The other PCA advantage is that once a data pattern is located it is compressed, i.e. by reducing dimensions number without information loss [33].

Principal components are located through calculation of data covariance matrix eigenvectors and eigenvalues in computational terms. This process is equal to finding axis system where co-variance matrix is diagonal [34]. The eigenvector with largest eigenvalue is greatest variation direction; one with second largest eigenvalue is (orthogonal) direction with next highest variation etc.

Let A be an n X n matrix. Eigenvalues of A are defined as roots of:

Determinant

$$(A - \lambda I) = |(A - \lambda I)| = 0 \quad (7)$$

Where I is then n X n identity matrix. This equation is a characteristic equation (or characteristic polynomial) with n roots.

Let  $\lambda$  be Eigen value of A. There exists a vector x so that:



$$Ax = \lambda x \tag{8}$$

Vector  $x$  is called eigenvector of  $A$  associated with Eigen value  $\lambda$ .

An Artificial Neural Network (ANN) is a mathematical/computational model attempting to imitate biological neural networks structure and functionality. It is composed of a set of simple, highly interconnected computational units called nodes, each representing a biological neuron [35]. In a NN, hidden units receive a weighted inputs sum applying an activation function to it. Then, output units receive a weighted hidden unit's output sum and apply an activation function to it. Information is passed from one layer to another. The network is called a Multi-Layer Perceptron Neural Network, with specific characteristics. To easily explain MLP neural network structure, Figure 1 reveals main components. It has an input layer representing input variables to be used in NN model and can be connected with output layer directly. It has  $i$  hidden layers with each layer containing  $j$  hidden units.

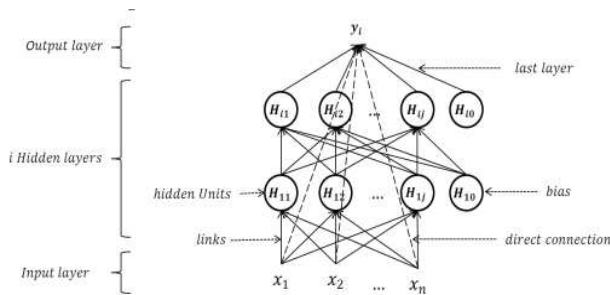


Figure 4. MLP Neural Network structure.

### 4.3. Genetic Algorithm

GA is an optimization technique trying to replicate natural evolution where individuals with best characteristics adapting to the environment are likely to reproduce and survive. Such advantageous individuals mate among themselves and produce offspring with similar characteristics thereby preserving favourable characteristics while those unfavourable are destroyed, leading to species progressive evolution.

GA iterates and evolves a population forming a new population at every step. GA iteration includes the following steps:

1. Selection: The first step selects individuals/chromosomes for reproduction. Fitness value is important in selection and completely randomly. Individuals having better fitness values are chosen more often for reproduction.
2. Reproduction: Offspring are produced by selected individuals. Recombination and mutation techniques generate new chromosomes.
3. Evaluation: Fitness of new chromosomes is evaluated.
4. Replacement: In this last step, old population individuals are removed and replaced by new ones [36].

Five genetic operators change MLP. In addition to mutation percentage and number of crossing points, each

operator's application priority has to indicate number of individuals generated by genetic operators. Tests are undertaken using higher priority for mutation and similar levels for remaining operators.

Mutation operators modify weights of specific neurons randomly, depending on application rate. The operator is used with 40% application probability, that is, 40% of weights are changed, which obtained better results empirically than lower probabilities.

Crossover operator carries out multipoint cross-over between two chromosome nets, to obtain two networks whose hidden layer neurons are a combination of both parents hidden layer neurons.

The additional operator, attempts to solve a main BP issue and its variants: difficulty in guessing hidden layer neurons number. This operator has to perform incremental learning starting with a small structure and its increments, adding new hidden units, if necessary. Simultaneously, it raises the dilemma of over fitting: small networks generalize well, but are slow learners, whereas big networks learn fast, but generalize badly.

Elimination operator eliminates a hidden neuron randomly. This operator has to perform decremental learning: it prunes certain nodes for better results in generalization and a smaller network [37].

Flowchart of the proposed method:

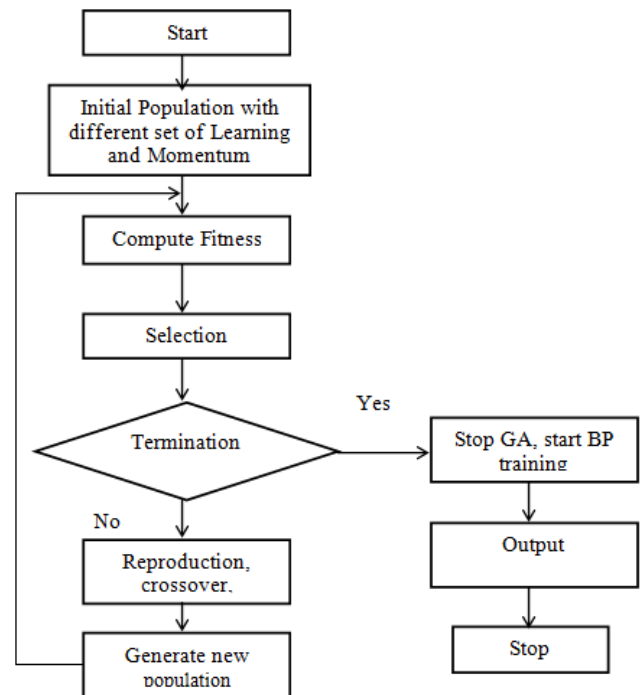


Figure 5. Flowchart of proposed method.

## 5. Results and Discussion

In this study, Optimization of Multi-layer perceptron neural network using genetic algorithm for classifying ECG arrhythmia is proposed. Symlet extracts RR intervals from ECG data as features while symmetric uncertainty ensures

feature reduction. GA is used to optimize learning rate and momentum. The results obtained are as shown in figure 6 to

10. Table 1 tabulates the result summary for various techniques.

Table 1. Summary of results.

Technique used	Classification accuracy	Precision	Recall	RMSE
MLP	0.9273	0.9281	0.9273	0.2046
Proposed MLP with GA optimization	0.9429	0.9428	0.9430	0.1924
MLP with fuzzy feature extraction	0.9543	0.9543	0.9545	0.1722
Proposed MLP with fuzzy feature extraction and GA optimization	0.9693	0.9693	0.9692	0.1542

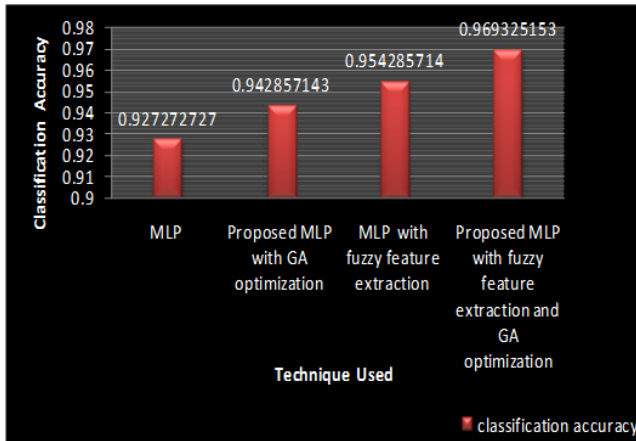


Figure 6. Classification accuracy of different techniques.

From figure 6, it is shown that the proposed MLP with fuzzy feature extraction and GA optimization has the high classification accuracy value as 0.9693.

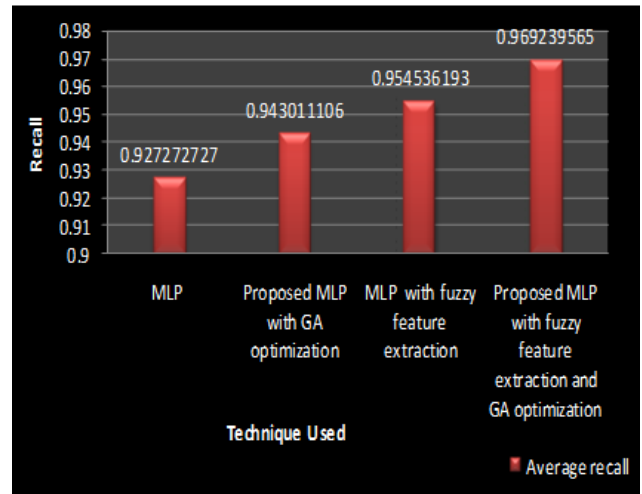


Figure 8. Average recall of different techniques.

From figure 8, it is shown that the proposed MLP with fuzzy feature extraction and GA optimization has the high average recall value as 0.96924.

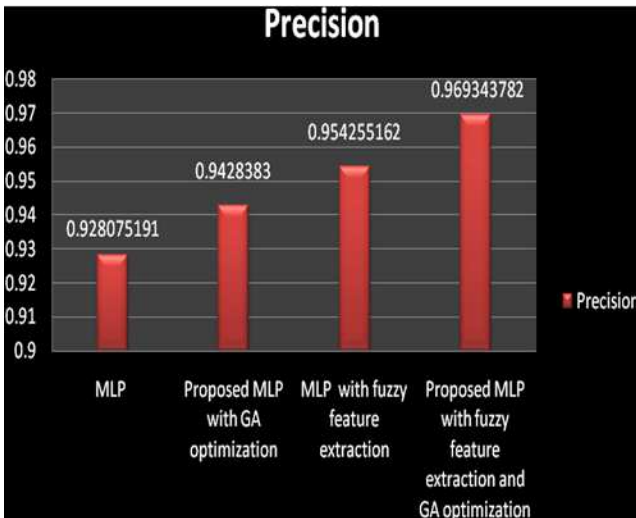


Figure 7. Average precision of different techniques.

From figure 7, it is shown that the proposed MLP with fuzzy feature extraction and GA optimization has the high average precision value as 0.969344.

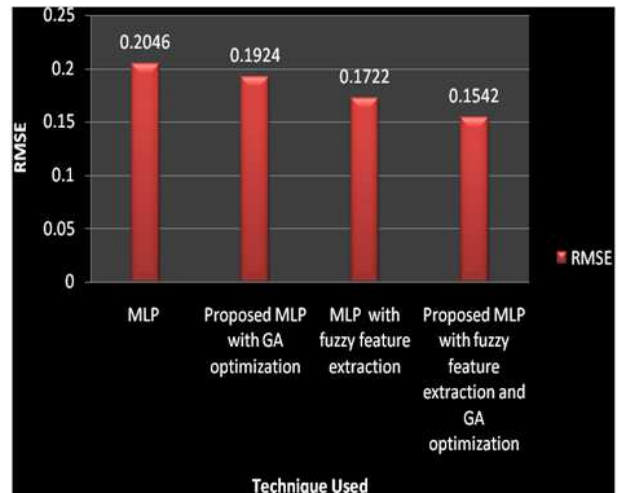


Figure 9. RMSE of different techniques.

The root mean square error of the proposed MLP with fuzzy feature extraction and GA optimization decreases to 0.1542.

Table 2. The comparative results of different approaches to best fitness.

Number of iterations	Best fitness - MLP with GA	Best fitness - MLP with fuzzy feature selection
1	0.237329364	0.178378517
5	0.229627814	0.174060391
10	0.226777397	0.172250112
15	0.224410964	0.172250112
20	0.222833936	0.172557942
25	0.221415012	0.173063692
30	0.219716221	0.164789972
35	0.21817534	0.164126341
40	0.216792272	0.164632091
45	0.215129049	0.164848841
50	0.213623449	0.164126341
55	0.212129421	0.162849929
60	0.205268552	0.163572429
65	0.203729038	0.162849929
70	0.207713508	0.165017429
75	0.206315939	0.163572429
80	0.206315939	0.162849929
85	0.206315939	0.162849929
90	0.206315939	0.162849929
95	0.206315939	0.162849929
100	0.206315939	0.162849929

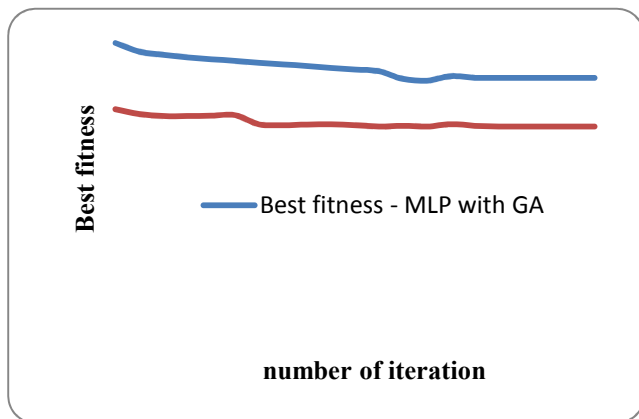


Figure 10. Best fitness value of different techniques.

From figure 10, it is shown that the proposed MLP with fuzzy feature extraction and GA optimization has best fitness function compared to other methods.

## 6. Conclusion

This study investigates ECG classification for arrhythmic beats based on RR intervals. The method includes RR interval beat extraction using Symlet conversion for ECG data. Symmetric uncertainty ensures feature reduction. Features are classified using Multi-layer perceptron Neural Network with GA optimization. GA is used to optimize learning rate and momentum. The MIT-BIH arrhythmia database is used to evaluate the classification efficiency for classifying instances like Right bunch bundle block, Left bunch bundle block and Normal RR interval. Experimental results show that the proposed optimized neural network achieved classification accuracy and Average precision of 96.93% and 96.92% Average recall. The RMSE of the proposed MLP with with fuzzy feature extraction and GA

optimization decreases to 0.1542 and it has best fitness function compared to other methods.

## References

- [1] Kastor, J. A. (2000). Arrhythmias. WB Saunders Co.
- [2] Nasiri, J. A., Naghibzadeh, M., Yazdi, H. S., & Naghibzadeh, B. (2009, November). ECG arrhythmia classification with support vector machines and genetic algorithm. In Computer Modeling and Simulation, 2009. EMS'09. Third UKSim European Symposium on (pp. 187-192). IEEE.
- [3] M. Elgendi, M. Jonkman, F. D. Boer, "Premature Atrial Complexes Detection Using the Fisher Linear Discriminant", 7th IEEE Int. Conf. on Cognitive Informatics (ICCI'08), 2008.
- [4] A. Gharaviri, F. Dehghan, M. Teshnelab, H. Abrishami, "comparison of neural network anfis, and SVM classification for PVC arrhythmia detection", Proceedings of the Seventh International Conference on Machine Learning and Cybernetics, Kunming, 12-15 July 2008.
- [5] F. Melgani, Y. Bazi, "Classification of Electrocardiogram Signals With Support Vector Machines and Particle Swarm Optimization", IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY IN BIOMEDICINE, VOL. 12, NO. 5, SEPTEMBER, 2008.
- [6] T. H. Linh, S. Osowski, and M. L. Stodoloski, "On-line heart beat recognition using Hermite polynomials and neuron-fuzzy network," IEEE Trans. Instrum. Meas., vol. 52, no. 4, pp. 1224-1231, Aug. 2003.
- [7] A. Azemi, V.R. Sabzevari, M. Khademi, H. Gholizadeh, A. Kiani, Z. Dastgheib, "Intelligent Arrhythmia Detection and Classification Using ICA", Proceedings of the 28th IEEE, EMBS Annual International Conference, New York City, USA, Aug 30-Sept 3, 2006.
- [8] T. Inan, L. Giovangrandi, and J. T. A. Kovacs, "Robust neural network based classification of premature ventricular contractions using wavelet transform and timing interval features," IEEE Trans. Biomed. Eng., vol. 53, no. 12, pp. 2507-2515, Dec. 2006.
- [9] Ceylan, R., Ozbay, Y., & Karlik, B. (2009). A novel approach for classification of ECG arrhythmias: Type-2 fuzzy clustering neural network. Expert Systems with Applications, 36, 6721-6726.
- [10] GholamHosseini, H., Luo, D., & Reynolds, K. J. (2006). The comparison of different feed forward neural network architectures for ECG signal diagnosis. Medical Engineering and Physics, 28, 372-378.
- [11] Lowitz, T., Ebert, M., Meyer, W., & Hensel, B. (2010, January). Hidden markov models for classification of heart rate variability in RR time series. In World Congress on Medical Physics and Biomedical Engineering, September 7-12, 2009, Munich, Germany (pp. 1980-1983). Springer Berlin Heidelberg.
- [12] Zhu, Y. (2012). SVM Classification Algorithm in ECG Classification. In Information Computing and Applications (pp. 797-803). Springer Berlin Heidelberg.
- [13] Khorrami, H., & Moavenian, M. (2010). A comparative study of DWT, CWT and DCT transformations in ECG arrhythmias classification. Expert systems with Applications, 37(8), 5751-5757.

- [14] Hu, J. L., & Bao, S. D. (2010, October). An approach to QRS complex detection based on multiscale mathematical morphology. In *Biomedical Engineering and Informatics (BMEI), 2010 3rd International Conference on* (Vol. 2, pp. 725-729). IEEE.
- [15] Yu, S., & Chen, Y. (2007). Electrocardiogram beat classification based on wavelet transformation and probabilistic neural network. *Pattern Recognition Letters*, 28, 1142–1150.
- [16] Yu, S., & Chou, K. (2009). Selection of significant independent components for ECG beat classification. *Expert Systems with Applications*, 36, 2088–2096
- [17] Ceylan, R., Özbay, Y., & Karlik, B. (2009). A novel approach for classification of ECG arrhythmias: Type-2 fuzzy clustering neural network. *Expert Systems with Applications*, 36(3), 6721-6726.
- [18] Arif, M., Akram, M. U., & Afsar, F. A. (2009, December). Arrhythmia Beat Classification using pruned fuzzy K-nearest neighbor classifier. In *Soft Computing and Pattern Recognition, 2009. SOCPAR'09. International Conference of* (pp. 37-42). IEEE.
- [19] Raut, R. D., & Dudul, S. V. (2008, July). Arrhythmias Classification with MLP Neural Network and Statistical Analysis. In *Emerging Trends in Engineering and Technology, 2008. ICETET'08. First International Conference on* (pp. 553-558). IEEE.
- [20] Ramirez, E., Castillo, O., & Soria, J. (2010, July). Hybrid system for cardiac arrhythmia classification with fuzzy k-nearest neighbors and Multi Layer Perceptrons combined by a fuzzy inference system. In *Neural Networks (IJCNN), The 2010 International Joint Conference on* (pp. 1-6). IEEE.
- [21] Nambiar, V. P., Khalil-Hani, M., & Marsono, M. N. (2012, December). Evolvable Block-based Neural Networks for real-time classification of heart arrhythmia From ECG signals. In *Biomedical Engineering and Sciences (IECBES), 2012 IEEE EMBS Conference on* (pp. 866-871). IEEE.
- [22] Haseena, H. H., Joseph, P. K., & Mathew, A. T. (2011). Classification of Arrhythmia Using Hybrid Networks. *Journal of medical systems*, 35(6), 1617-1630.
- [23] Martis, R. J., & Chakraborty, C. (2011). Arrhythmia Disease Diagnosis using Neural Network, SVM, and Genetic Algorithm-Optimized k-Means Clustering. *Journal of Mechanics in Medicine and Biology*, 11(04), 897-915.
- [24] Exarchos, T. P., Tsipouras, M. G., Exarchos, C. P., Papaloukas, C., Fotiadis, D. I., & Michalis, L. K. (2007). A methodology for the automated creation of fuzzy expert systems for ischaemic and arrhythmic beat classification based on a set of rules obtained by a decision tree. *Artificial Intelligence in Medicine*, 40(3), 187-200.
- [25] Wang, J. S., Chiang, W. C., Hsu, Y. L., & Yang, Y. T. C. (2012). ECG Arrhythmia Classification Using a Probabilistic Neural Network with a Feature Reduction Method. *Neurocomputing*.
- [26] Zadeh, A. E., Khazae, A., & Ranaee, V. (2010). Classification of the electrocardiogram signals using supervised classifiers and efficient features. *computer methods and programs in biomedicine*, 99(2), 179-194.
- [27] Moody, G. B., & Mark, R. G. (2001). The impact of the MIT-BIH arrhythmia database. *Engineering in Medicine and Biology Magazine, IEEE*, 20(3), 45-50.
- [28] Gustafson, D. E., Willsky, A. S., Wang, J. Y., Lancaster, M. C., & Triebwasser, J. H. (1978). ECG/VCG rhythm diagnosis using statistical signal analysis-I. Identification of persistent rhythms. *Biomedical Engineering, IEEE Transactions on*, (4), 344-353.
- [29] Vanisree, K., & Singaraj, J. (2011). Automatic Detection of ECG RR Interval using Discrete Wavelet Transformation. *International Journal on Computer Science and Engineering (IJCSE)*.
- [30] Kumari.V.S. R., & Kumar, P. R. (2012). Performance Evaluation of Boosting Techniques for Cardiac Arrhythmia Prediction. *International Journal of Computer Applications*, 57(17).
- [31] Abandah, G. A., & Malas, T. M. (2010). Feature selection for recognizing handwritten Arabic letters. *Dirasat Engineering Sciences Journal*, 37(2).
- [32] Zhang, L. (2001). Comparison of fuzzy c-means algorithm and new fuzzy clustering and fuzzy merging algorithm. University of Nevada, Reno, NV89557.
- [33] Smith, L. I. (2002). A tutorial on principal components analysis. *Cornell University, USA*, 51, 52.
- [34] Johnson, R. A., & Wichern, D. W. (2002). *Applied multivariate statistical analysis* (Vol. 5, p. 767). Upper Saddle River, NJ: Prentice hall.
- [35] Correa, A., Gonzalez, A., & Ladino, C. (2011). Genetic Algorithm Optimization for Selecting the Best Architecture of a Multi-Layer Perceptron Neural Network: A Credit Scoring Case. In *SAS Global Forum 2011 Data Mining and Text Analytics*.
- [36] Lacerda, E., Carvalho, A. C., Braga, A. P., & Ludermir, T. B. (2005). Evolutionary radial basis functions for credit assessment. *Applied Intelligence*, 22(3), 167-181.
- [37] Castillo, P. A., Merelo, J. J., Prieto, A., Rivas, V., & Romero, G. (2000). G-Prop: Global optimization of multilayer perceptrons using GAs. *Neurocomputing*, 35(1), 149-163.