



ABNORMALITY DETECTION IN ECG USING ARTIFICIAL NEURAL NETWORKS

Shahanaz Ayub¹, J. P. Saini².

¹Department of Electronics & Communication Engineering.
Bundelkhand Institute of Engineering & Technology. Jhansi.

²M.M.M. Engineering College. Gorakhpur
Uttar Pradesh, India

Email: [shahanaz_ayub @ rediffmail.com](mailto:shahanaz_ayub@rediffmail.com)

Rev Electron Biomed / Electron J Biomed 2010;3:47-52

[Comment of the reviewer Prof. Jose Luis Hernandez Caceres, PhD.](#) Center for Cybernetics Applications to Medicine, La Habana, Cuba.

[Comment of the reviewer Eduardo Tejera PhD.](#) Pharmacy Faculty, Porto University, Porto, Portugal.

ABSTRACT ECG i.e. Electrocardiogram represents electrical activity of the heart. Arrhythmias are among the most common ECG abnormalities. Millions of ECGs are taken for the diagnosis of various classes of patients, where ECG can provide a lot of information regarding the abnormality in the concerned patient, ECGs are analysed by the physicians and interpreted depending upon their experience.

The interpretation may vary by physician to physician. Hence this work is all about the automation and consistency in the analysis of the ECG signals so that they must be diagnosed and interpreted accurately irrespective of the physicians. This would help to start an early treatment for the problems and many lives could be saved. Many works have been done previously but this paper presents a new concept by application of MATLAB based tools in the same weighted neural network algorithms. This will help to reduce the hardware requirements, make network more reliable and thus a hope to make it feasible. To do so various networks were designed using the MATLAB based tools (licensed version) and parameters. Two classes of networks were designed, but with different training algorithms, namely Perceptron and Back propagation.

They were provided training inputs from the data obtained from the standard MIT-BIH Arrhythmia database. After training different forms of networks, they were tested by providing unknown inputs as patient data and the results in the whole process from training to testing were recorded in the form of tables. The results for the normal beats were best in the case of Cascade-Forward Back propagation network algorithm. The percentage of correct classification is 100%.

The results are compared with the previous work which concludes that the method proposed in this paper gives best results.

KEYWORDS: Arrhythmia. MATLAB. Artificial Neural Networks. Back propagation. Cascade Forward Network. MIT-BIH arrhythmia data base.

RESUMEN:

El electrocardiograma (ECG) representa la actividad eléctrica del corazón. Las arritmias cardiacas se encuentran entre las anomalías mas frecuentemente detectadas en el ECG. Comúnmente se registran millones de electrocardiogramas de variadas categorías de pacientes, y el ECG puede contener información valiosa acerca de las alteraciones en un paciente dado. Los ECGs son analizados por médicos especializados y son interpretados en correspondencia con su experiencia individual.

La interpretación varía de un especialista a otro, por lo que el presente trabajo se ha encaminado a la automatización y al logro de consistencia en el análisis de las señales electrocardiográficas de modo que puedan ser diagnosticadas e interpretadas independientemente del especialista. Esto ayudaría en la asunción temprana de un tratamiento y ayudaría a salvar numerosas vidas. Existen muchos intentos anteriores, pero el presente trabajo presenta una concepción novedosa al aplicar herramientas basadas en MATLAB con algoritmos de redes neurales ponderadas. Esto ayudaría a reducir los requerimientos de hardware, haría la red más confiable y al parecer más manejable. Para lograrlo, se diseñaron varias redes utilizando herramientas y parámetros basados en MATLAB (versión protegida por licencia). Se diseñaron dos clases de redes pero con diferentes algoritmos, específicamente "Perceptron" y "Backpropagation".

Las redes fueron entrenadas con datos obtenidos de la base de arritmias de MIT-BIH. Después de entrenar diferentes configuraciones de redes, estas fueron probadas con diferentes colecciones de entradas desconocidas y los resultados de todo el proceso desde el entrenamiento hasta la evaluación fueron tabulados. Los mejores resultados para las pulsaciones normales se obtuvieron con el algoritmo "Cascade-Forward Back propagation". El porcentaje de clasificación correcta alcanzó el 100%.

Los resultados fueron comparados con los de trabajos anteriores y se concluye que el método propuesto en el presente artículo resulta superior.

PALABRAS CLAVE: Arritmia. MATLAB. Redes Neuronales Artificiales. Redes de tipo Cascade Forward. Bases de arritmias de MIT-BIH.

1.- INTRODUCTION

Today many patients are suffering from cardiac problems. Heart disease is the most common cause of death in the world. In recent years considerable work has been done to assist cardiologists with their task of diagnosing the ECG recordings. Detection and treatment of arrhythmias has become one of the cardiac care unit's major functions. More than 3 million ECGs are taken worldwide each year for the patients with different cases, right from heart rhythm anomaly to the hormonal imbalances due to organ failures. All the samples taken have one thing in common and that is, they are analysed by the experienced doctors who depending upon their knowledge predict out the problem(s) associated with the patient which is disturbing the normal morphology of the signal. If this morphological disturbance becomes somewhat complex (such as the case of fusion beats) then it is analysed by them depending upon their experience. This experience based analysis gives different interpretations. Hence there is a need of a system that could analyse the ECG signals properly and with a great accuracy so that there is a less chance of mistake as well as the problem is spotted in time so that an early treatment could be started.

So to achieve this objective many works have been done in this field based on image processing, Digital Signal Processing etc and prominent among them is the use of Artificial Neural Networks¹ which has given promising results to such complex problems. Neural network based analyses made were either weight based or weightless. This work is based on weighted neurons with bias adjustments but with the application of MATLAB based algorithms and neural network structure.

The excellent features of the MATLAB²⁻³ such as wide range of tools for network structure development and adjustment according to requirements as well as tools to analyse the results, makes it a

good option to solve this complex problem in a simple way, especially the case of fusion beats. In this paper the case of normal beats is discussed so as to have an insight into the concept of identification of normal beats using cascade forward neural networks (MATLAB based) with back propagation algorithm. The data base used in this paper to train and test the neural network is the standard MIT-BIH arrhythmia database⁴.

2. OBJECTIVE OF THE WORK The objective of this work is to make the analysis of normal beats easy so that the patient could be diagnosed for the heart problems in less time as well more accurately so that the medical practitioners have primary information about the ailment and could start a treatment early. Apart from this the project has been targeted towards the rural community and so we are also considering hardware implementation of this work but in low cost and greater efficiency. The soft computing technique used is MATLAB based neural network tools to identify the normal beats from ECG beats so that other beats can be detected as abnormal beats.

3. METHODOLOGY

The database provided by the MIT-BIH arrhythmia database⁴ regarding different kinds of heart rhythm abnormalities for different class of patients, is the source of data used for training and testing of the neural networks. The data from the patient number 208 was preferred and taken out for different cases of heart beats. Other patient data were also taken so as to enhance the prediction capability of the trained neural network and make it more accurate.

The data taken was used to make training inputs which represented the whole ECG cycle as well as for making test inputs. The MATLAB based Perceptron and back propagation⁵ networks were developed and training parameters were fixed for certain quantities and varied for others. The network trained were analysed using the test inputs first for all unknowns which were not used for training and then for all inputs which included both training inputs and test inputs. Analysis plot tools^{5,6} were also used to understand the network capability and other properties such as Mean Squared Error (MSE) value and learning capability.

4. Database

The MIT-BIH Arrhythmia Database contains 48 half-hour excerpts of two-channel ambulatory ECG recordings, obtained from 47 subjects studied by the BIH Arrhythmia Laboratory between 1975 and 1979. The recordings were digitized at 360 samples per second per channel with 11-bit resolution over a 10 mV range. Two or more cardiologists independently annotated each record; disagreements were resolved to obtain the computer-readable reference annotations for each beat (approximately 110,000 annotations in all) included with the database. For our study purpose the record chosen was the MIT-BIH patient 208. This was chosen for two reasons:

- 1) The patient was without any kind of medication and had sufficient bits for normal and ventricular as well as maximum number of beats for fusion analysis among all patient data available.
- 2) Earlier weightless neural network analyses were done by other scholars using this patient⁷ so we wanted to compare our network and its result with their one. Figure 1. shows the sample record of data of patient 208.

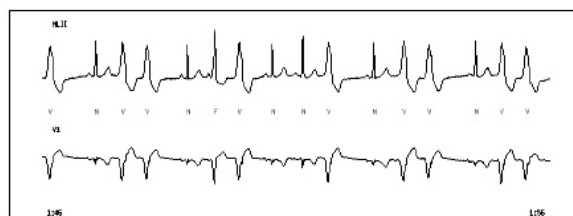


Fig. 1. A sample record of data of patient 208

5. Inputs

The inputs for the networks were selected considering two important points:

- a) The inputs must be of a standard size such that it is neither too small to cover up one ECG cycle and nor too high to increase the number of bits required to analyse the signal, thus increasing the hardware requirements.
- b) Secondly the input must be so arranged that the R peak in the QRS complex must be at the centre of the signal cycle under considerations.

The first condition was achieved by setting up an arbitrary value of 301 samples of MLII lead data⁸ obtained from the database in which the 150 samples were on the left side and 150 samples on the right side of the 151st sample value. From the database when the ventricular beats were derived we got the R peak values in the table form. This R peak was taken as centre and from the samples of the same patient 150 sample before this R peak value and 150 samples after it was taken to make a 301 sample input, where the ventricular beat was in the centre. Thus the input becomes a matrix of 301x1 and ready to be used in MATLAB. The same process was repeated to make all the inputs of all the kinds of inputs that are normal, fusion and ventricular.

The second condition was achieved by allowing the 151st sample to be the beat value of MLII lead signal obtained from the database for particular conditions. After the inputs were arranged in the matrix then the inputs for training were taken in two stages (not for all cases of network), named as input1 with total samples of 1809 and input2 with 2921 samples. This division was made with taking approximately 1/3rd of various kinds of samples in the first case and all the inputs fixed for training in the second case. This was done to check the chances of having desired network accuracy in minimum number of inputs.

E.g. If sample number 2242 of patient 208 is a R peak then from the samples of the same patient, values for MLII lead from sample number $(2242-150)=2092$ to sample number $(2242+150)=2392$ will be input data. This is illustrated in the Figure 2.

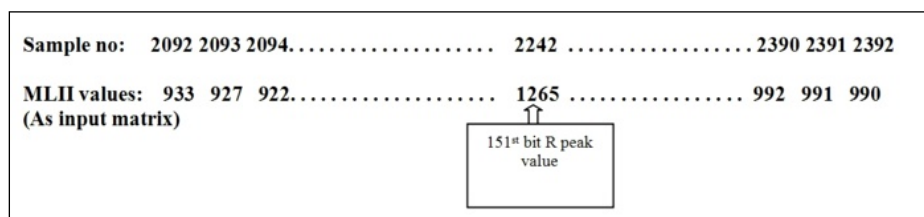


Fig. 2. Selection and Organisation of data

The Table 1 is described as; total Normal beats taken for analysis are 1586. Out of which 1286 normal beats are used for training the network. 300 normal beats are used for testing the trained network and 6 normal beats are used for validation of the network. One single beat given to the network input consists of 301 data points as shown in figure 2. Similarly fusion beats, ventricular premature beats and unclassified but abnormal beats analysis can also be explained.

Table 1. Input Matrix structure.

KIND	INPUTS	MATRIX dimension	Training(total)	Testing	Validation
Normal Beats	1586	301x1586	301x1286	301x300	301x6
Fusion Beats	734	301x734	301x584	301x150	301x6
Ventricular Premature Beats	992	301x992	301x792	301x200	301x6
Unclassified (but abnormal) Beats	259	301x259	301x179	301x80	-

6. Analysis

The following tables provide the results related to various beats analysed using the best case of cascade- forward back propagation network, among them the results for the normal beats should be observed carefully.

Table 2: Training statistics for cascade forward analysis for normal beats with 'trainbfg' training algorithm

INPUTS	HN	TIME (in SEC)	P(cc) UNKNOWN	P(cc) ALL	EPOCHS (MAX=1000)	MSE 0.0001	RESULT
1284N/2921T	5	434	100V,99.3N,94F,100U	99.7V,96.5N,94F,100U	29	0.0291	PASSED
	10	3029	100V,99N,94F,100U	99.7V,97N,94.3F,96.8U	39	0.0221	PASSED

Table 3: Training statistics for cascade forward analysis for normal beats with 'trainlm' training algorithm

INPUTS	HN	TIME (in SEC)	P(cc) UNKNOWN	P(cc) ALL	EPOCHS (MAX=1000)	MSE 0.0001	RESULT
1284N/2921T	5	24	99.5V,98N,95.3F,92.5U	99.4V,98.5F,97.8F,95.6U	12	0.0117	PASSED
	10	112	100V,97.7N,94.7F,98.8U	99.7V,98.5N,98.2F,97.7U	12	0.00552	PASSED

TRAINBFG - BGFS Quasi Newton Back propagation training algorithm

TRAINLM - Levenberg marquardt back propagation Training algorithm

P (CC) - Percentage of correct classifications

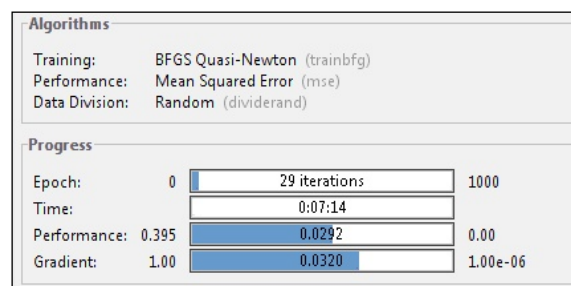
INPUTS - total number of samples is suffixed by T

HN - Hidden neurons, representing the number of neurons in the Hidden layer

MSE - Mean Squared Error; the error goal was fixed at 0.0001 and hence here the difference MSE-0.0001 is being tabulated

SUFFIXES V, N, F, AND U - in the P (cc) columns indicate respectively Ventricular, normal, fusion and unclassified beats and their percentage of correct classifications.

The following images are the analysis plots for this work where each of them interprets different properties about the network.

**Fig. 3. Training Process Results**

Fewer epochs mean network learns in small repetitions. Less time means network achieved goal easily and shortly. Performance indicates the final MSE achieved. Lower value is associated with higher network accuracy.

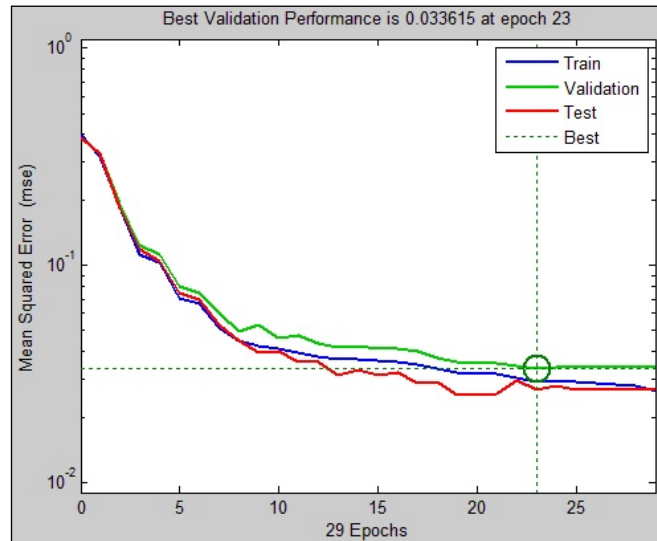


Fig. 4. Mean squared error plot for cascade forward network

Mean squared error plot shows the achieved error value. Best performance was at 23rd epoch. Lower value means the less probability of false predictions. Here network has achieved quite low error probability.

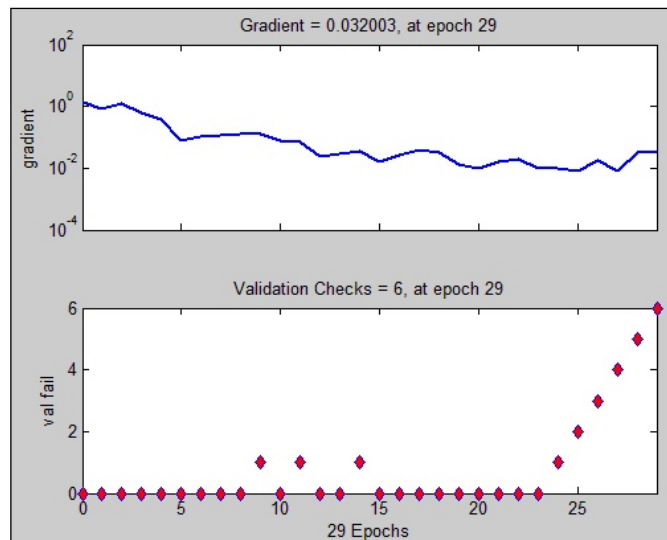


Fig. 5. Gradient and Validation Check plots

Low value of gradient plot indicates that the network is learning up to a large extent which means finer adjustments in the weights and bias. This in turn makes network more accurate and reliable, avoiding chances of false predictions. Validation plot shows the point where the network learned sufficiently and passed validation without. The point where the failures cross the defined limit is the stoppage point of training and indicates the starting of the over fitting of the data.

7. CASCADE-FORWARD NETWORK STRUCTURE

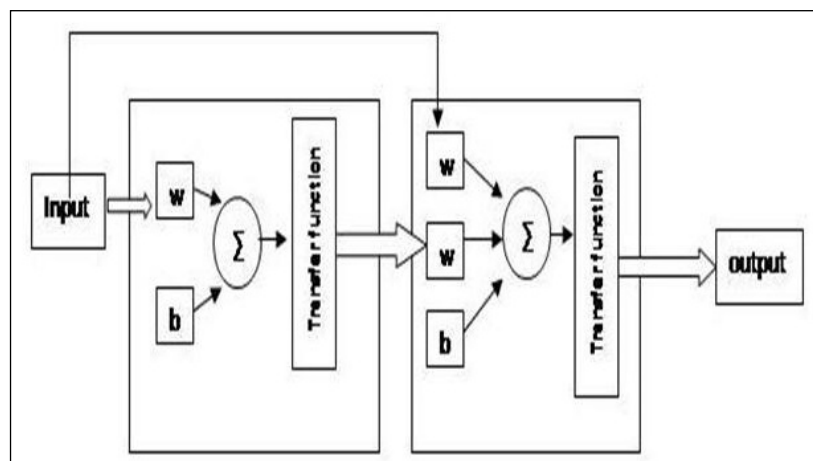


Fig. 6: Cascade-Forward Network General Structure

7.1. Inferences

- This network gives quite low value of MSE and it is near 0.033615 in just 29 epochs. Time taken for training may be large but it is giving good results especially for fusion and unclassified

bits also.

- Though one can see the trainlm case also gives good result but again we are considering this one due to its low memory usage.

Validation error plot shows the fitting confusion which is seen in the form of two different elevations or increase in validation error. But near about 25 iterations it comes to the desired adjustments.

8. RESULTS:

The results obtained with other methods like weightless neural networks, MLP^{3,7,9-12} etc. are compared with our results. Table 4 shows the comparison of the results.

Table 4: Comparative Results with other methods

Methods	MLP	HFNS	PCA	Weightless	Our method
% of correct classification	98.85%	98.78%	98.80%	99.69%	100%
Remark	Comparatively lowest accuracy	2 ECG leads are used as inputs	Same samples are used for training & testing the network	Comparatively low accuracy, little complex	Best Accuracy

9. CONCLUSION

The network based on Cascade-Forward network algorithm with trainbfg training algorithm was best for the case of normal beat analysis because it has an accuracy of about 100% as well as the memory requirements were also low. Hence we preferred this network for the normal beat analysis. The conclusion derived from this work is that, by using the MATLAB based neural network design⁶; such networks can be made which have capability to understand different class of inputs when they are fed to be analysed. Such networks can be very reliable as MATLAB provides a good set of tools so that the network parameters can be adjusted easily and precisely by just adjusting values for them and change in full length code, as was done previously, is not required. Though the objective of this research was not to use MATLAB or Neural Networks, these were used to get higher accuracy in analysis of ECG which is more useful for the mankind.

10. Limitations of present study

Though the networks are trained and tested with different data beats available in the standard data base, it could be tested with the other new current patients data. The results obtained would be the same as 100% accuracy as the training and testing of the network is done with large number of ECG beats. But still this limitation could be overcome in further studies of this work where other unclassified classes are also under study of exact classification though today they are just classified as abnormal, they need to be given the specific name of the abnormality.

REFERENCES

- 1.- Zurada J. Introduction to artificial neural systems. West Publishing House, 1999.
 - 2.- Gilat A. MATLAB- An Introduction with Applications, Wiley & Sons publications INC., 2005.
 - 3.- Demuth H, Beale M, Hagan M. Matlab Neural Network Toolbox, Version R2008b. The MathWorks, Inc. October 2008.
 - 4.- Brown G. MIT-BIH Arrhythmia database. MIT, 2006.
 - 5.- Demuth H, Beale M, Hagan M. Matlab Simulink, Version 6. The MathWorks, Inc. October 2004.
 - 6.- Schreck DM. System and method for synthesizing leads of an Electrocardiogram. United States Patent 6901285. Available online <http://www.freepatentsonline.com/6901285.html>
 - 7.- Manintveld WA. Automated ECG Analysis using Weightless Neural Networks. Master's thesis, Dept. of Applied Mathematics, Delft University of Technology, 1996; 58.
 - 8.- Chickh M. A.N. Belgacem, F. Bereksi-Reguig, Neural classifier to classify ectopic beats. Acte des IX^{emes} rencontre de la Societe Francophone de Classification, Toulouse(2002), le 16-18.
 - 9.- Thomson DC, Soraghan JJ, Durrani TS. An automatic neural-network based SVT/VT classification system. Computers in Cardiology. 1993, Proceedings: 333-336.
 - 10.- Gao D, Kinouchi Y, Ito K, Zhao X. Time series identifying and modeling with neural networks. Neural Networks. [Proceedings Int Joint Conf Portland, USA]. 2003;4:2454-2459.
 - 11.- Gao D, Kinouchi Y, Ito K, Zhao X. Neural networks for event extraction from time series: A back propagation algorithm approach. Future Generation Computer Systems 2005; 21: 1096-1105.
 - 12.- Nadal J, de C. Bossan M. Classification of cardiac arrhythmias based on principal component analysis and feed forward neural networks. Computers in Cardiology 1993; 341- 344.
- Chow HS, Moody GB, Mark RG. Detection of ventricular ectopic beats using neural networks. Computers in Cardiology 1992; 659-662.

CORRESPONDENCE:

Shahanaz Ayub
Department of Electronics & Communication Engineering,
Bundelkhand Institute of Engineering & Technology,
Jhansi, Uttar Pradesh
284128 - India
Email:shahanaz_ayub@rediffmail.com

Comment of the reviewer Prof. Jose Luis Hernandez Caceres, PhD. Center for Cybernetics Applications to Medicine, La Habana, Cuba.

The extraction of useful information from medical datasets is one of the most promising lines in health informatics. In particular, automatic diagnosis of ECG recordings can potentiate health services allowing reliable early detection of abnormality in primary attention setting where the opinion of highly qualified experts is not always at hand.

Authors incorporate MATLAB-based tools for automatic detection of arrhythmias in ECGs. Thanks to the widely use of MATLAB worldwide the implementation of the proposed networks is feasible, and the level of accuracy reached with them makes the work done by colleagues from Bundelkhand Institute of Engineering & Technology worth of recognition.

Comment of the reviewer Eduardo Tejera PhD. Pharmacy Faculty, Porto University. Porto. Portugal.

The classification of abnormal beats during ECG measurements is not a new area, however, the author approach is novel and interesting mainly in the kind of neural network (NN) applied. The methodology is appropriated and well executed.

I recommend to the author (in further works) to increase the size for validation group. Is well known that NN could be extremely useful in classification problems but at the same time with a risk of overfitting.

The use of Matlab is a very intelligent choice however, Matlab is not a free platform. The author could test and use Octave, a free alternative and very similar to Matlab.

In general the author provide a convincing approach and clear results to justified the methodology application in abnormal beats recognition.

Received November 22, 2010. Received reviewed December 12, 2010
Published: December 20, 2010