Indian Journal of Science and Technology, *Vol 10(12), DOI: 10.17485/ijst/2017/v10i12/112970, March 2017*

ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

Optimization of ECG Peaks (Amplitude and Duration) in Predicting ECG Abnormality using Artificial Neural Network

Fakroul R. Hashim*, Nik G. N. Daud, Anis S. N. Mokhtar, Amir F. Rashidi, Ja'afar Adnan and Khairol A. Ahmad

Department of Electrical and Electronic Engineering, Faculty of Engineering, Universiti Pertahanan Nasional Malaysia, Kuala Lumpur, Malaysia; fakroul@upnm.edu.my

Abstract

Artificial Neural Networks (ANN) adapted from neuron concept's, generally applied in various applications especially the fields of biomedical engineering. ANN techniques have been applied in order to provide educated solutions to assist in decision making for the medical purpose. The study was conducted for the purpose of determining the suitability and implementation of ANN to detect ECG abnormalities by using six features from ECG signal, both amplitude and duration of P, QRS and T peaks and used as input vector for ANN. In this study, Multilayer Perceptron (MLP) network is trained by using three different training/learning algorithms. The network is trained by using Bayesian Regularization (BR) algorithm has provided the highest accuracy performance (93.19%), followed by Levenberg Marquardt (LM) (92.88%) and Backpropagation (BP) (88.63%).

Keywords: Amplitude, Duration, ECG Abnormality, Multilayer Perceptron Network

1. Introduction

Cardiovascular disease is still the leading cause of death in the whole of mankind. Most cases can be reduced if early detection, such as pre-diagnostic or pre-monitoring can be provided. Early detection of abnormal conditions on the function of the heart, called arrhythmias, can be valuable to doctors[1–3]. Prevention and monitoring of humans from a heart attack is an important issue and played by an Electrocardiogram (ECG). P, QRS and T peaks contain the typical ECG signals. The a trial depolarization, P, ventricular depolarization, QRS complexes, and ventricular repolarization, T, is based on the wave of the far field caused by a phenomenon specific to the surface of the heart[1,2].

In clinical tests, ECG analysis is a technique commonly used. Industrial plants in this technique began in the late 1950s, and each year's how the total increase in the recording was analyzed and found. Thus, research on automated ECG analysis grew and became one of the exciting fields of biomedical engineering at the branch. Various types of ECG analysis algorithms have been developed, over the last two decades. Most of the studies focused on the detection of QRS complexes. In ECG recordings, there are the numbers of parameter with useful information to indicate the status of patients[4–6].

An ANN can be viewed as an application that is designed based on the model of brain function[7]. Based on the characteristics of ANN are capable of solving the problems of linearity or non-linearity, map of input-output data and analogy of neurobiology. ANN has been applied in many fields, including engineering, finance and mathematic. In these areas, ANNs are used for system identification, image processing, and pattern recognition and data classification with a significant level of success.

This study has dedicated to the development of smart system prediction system of ECG abnormalities based data for MLP network. In order to do so, the study employed 3 different training/learning algorithms, namely the Bayesian Regularization (BR)[8], the Lavenberg Marquardt

* *Author for correspondence*

(LM)[8] and Back Propagation (BP)[9]. A system is developed and trained based on those 3 algorithms in classify the condition of patient either in normal condition or vice versa based on amplitude and duration of P, QRS and T peaks.

2. Multilayer Perceptron Network

Artificial Neural Networks (ANN) is a computing system and inspired from a biological nerve cells known as neurons. Neurons are small cell which consists of our brain. By looking at the concept of neuron, the proposed ANN, which is able to model the biological structure of neurons in based on their operations and architectures. They serve as a mathematical model for clustering/non-parametric regression, data classification and computing non-linear function approximation. They also operated as a simulation for the behavior of the neuron model of human biology. ANN is able to provide reliable performance, especially in making decision in place of the human brain. Therefore, the developed system could be an alternative choices system to the intelligent developed system. One of the ANNs used for this system is the MLP network[10–12].

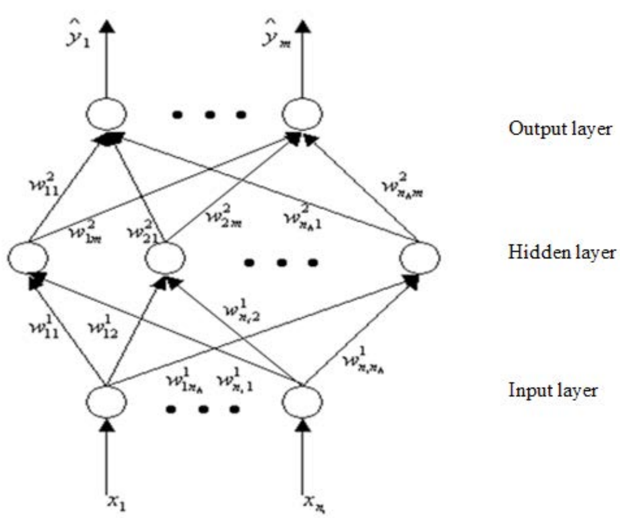

Figure 1. Structure of MLP network[13].

Rosenblatt in 1958 introduced an artificial neuron model, called Perceptron model[7]. He cascaded several number of Perceptron model in an array/layer network has been shown in Figure 1. ANN in the rest of paper is referred as MLP network. By referring to Figure 1, the MLP network with the hidden layer, n_h and output layer, m respectively, is expressed as:

$$\hat{y}_k(t) = \sum_{j=1}^{n_h} w_{jk}^2 F\left(\sum_{i=1}^{n_i} w_{ij}^1 x_i^0(t) + b_j^1\right) \text{for } 1 \le k \le m \quad (1)$$

Based on the Equation (1), w_{ij}^1 is denote as the weight connection between input layer to hidden layer while w_{jk}^2 denote as the weights connection between hidden layer to the output layer, respectively. b_j^1 in the Equation 1 represent as hidden node's threshold and x_i denote as inputs parameter which are fed to the input layer. In this study, sigmoid activation function has been chosen and denotes as $F(\bullet)$ in Equation (1). By referring to Equation (1), the values of w_{ij}^1, w_{jk}^2 and b_j^1 are determined based on an appropriate algorithm. In this study, three different training algorithms are used which are Back Propagation (BP), Lavenberg Marquardt (LM) and Bayesian Rule (BR) algorithms.

3. Back Propagation Algorithm

Back Propagation (BP) training algorithm, usually to train network since the algorithm is gradient descent procedure. The BP capable calculates the values of derivatives inefficient doing adjustment of the weight according to the parameter/protocol known as learning phase[9]. Back Propagation is a steep decent type which the connection weight between *j*-th neuron of hidden layer and the neuron to enter *i*-layer is updated according to:

$$\begin{aligned} w_{ji}(t) &= w_{ij}(t-1) + \Delta w_{ij}(t) \\ b_j(t) &= b_j(t-1) + \Delta b_j(t) \end{aligned} \quad (2)$$

The updated weight, $\Delta w_{ij}(t)$ and $\Delta b_j(t)$ given by:

$$\begin{aligned} \Delta w_{ij}(t) &= \eta_w \rho_j(t) x_i(t) + \alpha_w \Delta w_{ij}(t-1) \\ \Delta b_j(t) &= \eta_b \rho_j(t) + \alpha_b \Delta b_j(t-1) \end{aligned} \quad (3)$$

where the subscripts *w* represent the weight and *b* represent the threshold, respectively. The α_w and α_b are the constants momentum, in order to determine the changes which influences of past parameter to the current parameter. The η_w and η_b noted of training/learning rates while $\rho_j(t)$ is the signal error of the *j*-th neuron (of the hidden layer), propagated back to the network.

The signal error at the output node is shows in Equation (4), with the consideration of the activation function is functioned as linear parameter:

$$\rho\ (t) = y_k(t) - \hat{y}_k(t) \quad (4)$$

where $y_k(t)$ is the predicted output and $\hat{y}_k(t)$is the desired output. The output at the hidden layer is

$$\rho_j(t) = F'(x_i(t)) \sum_j \rho_j^k(t) w_{jk}^2(t-1) \quad (5)$$

where $F'(x_i(t))$ is derived from $F(x_i(t))$ by respecting to $x_i(t)$. The BP algorithm is a type of steepest decent method so; it's suffers slow convergence speed. Search for global minimum could be trapped in local minimum. It is also sensitive to users elected parameters[8].

4. Levenberg Marquardt Algorithm

The Levenberg Marquardt (LM) training algorithm, do the optimization based on deterministic gradient-based local. Once used to train the MLP network, the profit of the LM algorithm to be compared with the BP in providing faster convergence rate and maintain a relative stability[8]. As a quasi-Newton method, LM training algorithm is designed to reach the second order training speed and bypassed the Hessian matrix. During the time LM function had formed sum of squares, the Hessian matrix would be estimated with:

$$H = J^T J \quad (6)$$

and the gradient of the matrix is calculated as:

$$g = J^T \rho \quad (7)$$

where J is the Jacobian matrix which derived network errors by respecting network's weights and biases. The matrix can be calculated by using a standard BP technique, less complexity than computed by Hessian matrix[8]. The LM algorithm used the estimator to update Newton-like in Hessian matrix as:

$$\Delta w = -\left[J^T J + \mu I\right]^{-1} J^T \rho \quad (8)$$

Where updated weight, Δw is controlled by μ.. The Hessian in order to update Newton-like equation at the moment scalar μ is small (approach to zero). Newton method capable to converge faster than BP with more accurate prediction and results minimum error. The μ is decreased at each successful step(minimize error), however, the μ is increased when a tentative step function (increment error), in order to increase in performance. Therefore, the performance function need to be decrease at each iteration[8].

5. Bayesian Regularization Algorithm

Training/learning algorithm introduced by Thomas Bayes called D. Bayes' rule is applied to estimate the posterior probability of θ given the data D[9]. Normally, the posterior probability is provided to entire tabulation over possible values of θ. The Bayes rule given as

$$p(\theta \mid D) = \frac{p(D \mid \theta)}{p(D)} \quad (9)$$

with $p(\theta)$ is the pre-probability of a parameter θ before the data is seen while and $p(\theta \mid D)$ is called the likelihood, the probability of data. This process is applied to ANN and results the probability distribution over the network weights. The w given in the training data $p(w \mid D)$ is the network weights. The posterior distribution is observed by

$$p(w \mid D) = \frac{p(D \mid w) p(w)}{p(D)} = \frac{p(D \mid w)}{\int p(D \mid w) p(w) dw} \quad (10)$$

By refer to Bayesian formulation, the training to the weights that may change our belief of the weights from the pre *p(w)*, to the posterior, $p(w \mid D)$ as shown in Figure 2.

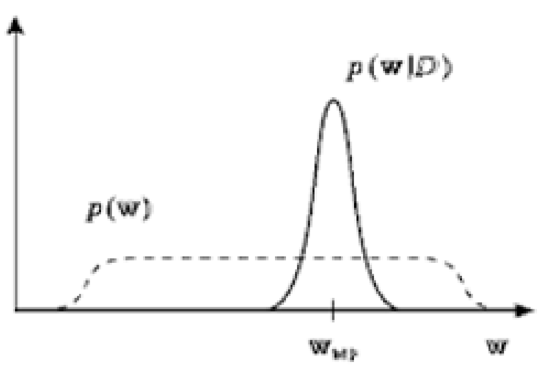

Figure 2. Changes pre weights to posterior/post weights[12].

6. Data Sample

Recorded ECG samples were taken from MIT-BIH repository is used in this research. Each ECG signal is extracted and 6 features such as amplitude of P peak, duration of P peak, amplitude of QRS peak, duration of QRS peak, amplitude of T peak and duration of T peak is measured[14]. The combination of these features will be applied as the input vector for the MLP networks. In the study, the MLP network used six input nodes represents the features or input parameter of MLP network. During the research, 1800 dataset are used to train the MLP network, where 1000 of them are used to determine the optimum structure of MLP network while the rest is applied to test the capability of MLP network.

7. Results and Discussions

By using the same analysis procedure used by[15–16] on data prediction and classification, the research is applying two types of analysis. On the first analysis, the optimum structure of the MLP network is determined. The optimum hidden nodes numbers of the MLP network is determined, the complexity of the MLP network may result high computational network and time consuming. During the analysis, MLP network is trained up to 100 epochs. Table 1 shows the analysis performance of three different types of training algorithm.

Table 1. Optimum structure performance MLP network

MLP Training Algorithm	Optimum number of hidden nodes
BP	10
LM	5
BR	4

Further comparison analysis comes as second analysis then is applied by using MLP optimum structure's obtained during the first analysis. The second analysis is determined based on the occurrence of correct classifications of ECG abnormality. Table 2 below shows the result of the second analysis both during the training and testing phase.

The results tabulated in Table 1 shows BR performance in trained the MLP network to form the simplest network architecture, by requiring only 4 hidden nodes as comparing with BP and LM training algorithms. For others training algorithms they need 5 and 10 hidden nodes layer to reach the optimum level for both LM and BP, respectively. For the analysis performance, as tabulated in Table 2 shows the MLP network trained using BR training algorithm produces the highest accuracy performance among LM and BP with 93.50% and 92.88% for training and testing phase respectively. Results for LM training algorithm are 93.25% and 92.50% for training and testing phase while for BP training algorithm, the performance are at 89.50% and 87.75% accuracy for both training and testing, respectively.

Table 2. Accuracy performance analysis of MLP network

Training	Accuracy Performance		
Algorithm	Training (%)	Testing (%)	Overall (%)
BP	89.50	87.75	88.63
LM	93.25	92.50	92.88
BR	93.50	92.88	93.19

8. Conclusion

This paper analyses ANN approach in terms to monitor the ECG This study proves that ANN is capable to result high accuracy prediction on ECG abnormality by using amplitude and duration of P, QRS and T peak of ECG signal as the input parameters. Further improvements can be developed in improving the performance of the designed system. Moreover, numbers of studies will be conducted order to improve the capability of prediction. A good advice is to use various types of ANNs and training algorithms can be used.

9. Acknowledgement

This research work is supported by the research grant (RAGS/1/2014/TK03/UPNM/2).

10. References

1. Barker RL, Burton BJ, Zieve PD. Principles of Ambulatory Medicine. 6th ed. Philadelphia: Lippincott, Wilkins and Williams; 2003.
2. Custer JW, Rue RE. The Harriet Lane Handbook. 18th ed. Philadelphia: Mosby Elsevier Inc; 2008.
3. Mark JB. Atlas cardiovascular monitoring. New York: Churchill Livingstone; 1998.

4. Houghton AR, Gray D. Making Sense of the ECG. 3rd ed. Hodder Education; 2012. p. 214.
5. Plesnik E,Malgina O, Tasic JF, Zajc M. ECG signal acquisition and analysis for telemonitoring.15th IEEE Mediterranean Electrotechnical Conference (MELECON); 2010. Crossref
6. Sonali SO, Sunkaria RK. ECG signal denoising based on Empirical Mode Decomposition and moving average filter. IEEE International Conference on Signal Processing, Computing and Control (ISPCC); 2013. Crossref
7. Haykin S. Neural network: A comprehensive foundation. New Jersey: Prentice Hall; 1994.
8. Payal A, Rai CS, Reddy BVR. Comparative analysis of Bayesian regularization and Levenberg-Marquardt training algorithm for localization in wireless sensor network. 15th International Conference on Advanced Communication Technology (ICACT); 2013.
9. Kaensar C. Analysis on the parameter of back propagation algorithm with three weight adjustment structure for hand written digit recognition.10th International Conference on Service Systems and Service Management (ICSSSM); 2013. Crossref
10. Gas B. Self-organizing multilayer perceptron. IEEE Transactions on Neural Networks. 2013; 21(11):1766–79. PMid:20858579. Crossref
11. Sivaram GSVS, Hermansky H. Multilayer perceptron with sparse hidden outputs for phoneme recognition. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP); 2011. Crossref
12. Isa NAM, Hashim FR, Mei FW, Ramli DA, Omar WMW, Zamli KZ. Predicting quality of river's water based on algae composition using artificial neural network. 2006 IEEE International Conference on Industrial Informatics; 2006. p. 1340–5.
13. Ramli DA, Saleh JM, Hashim FR, Isa NAM. Multilayered Perceptron (MLP) network trained by recursive least squares algorithm. International Conference on Computers, Communications, and Signal Processing with Special Track on Biomedical Engineering; 1st CCSP 2005; 2005. p. 288–91. Crossref
14. MIT-BIH Arrhythmia Database – PhysioNet. www.physionet.org/physiobank/database/mitdb
15. Mat Isa NA, Mashor MY, Othman NH. Classification of cervical cancer using HMLP network with confident level analysis. International Journal of Computer, the Internet and Management. 2003; 11(1):7–29.
16. Mitra P, Mitra S, Pal SK. Staging of cervical cancer with soft computing. IEEE Transaction on Biomedical Engineering. 2000; 47(7):934–40. PMid:10916265. Crossref