2022 IEEE 12th International Conference on Control System, Computing and Engineering (ICCSCE), 21–22 October 2022, Penang, Malaysia

Multilayer Perceptron Optimization of ECG Peaks for Cardiac Abnormality Detection

Addzrul Hi-Fi Syam Ahmad Jamil
Department of Electrical Engineering
Polytechnic of Seberang Perai
Penang, Malaysia
addzrulhifi@psp.edu.my

Jailani Abdul Kadir
Department of Electrical Engineering
Polytechnic of Tuanku Syed Sirajuddin
Perlis, Malaysia
jailani@ptss.edu.my

Johanis Mohd Jamil
Department of Mathematics, Science & Computer
Polytechnic of Tuanku Syed Sirajuddin
Perlis, Malaysia
johanis@ptss.edu.my

Fakroul Ridzuan Hashim
Faculty of Engineering
National Defence University of Malaysia
Kuala Lumpur, Malaysia
fakroul@upnm.edu.my

Shazreen Shaharuddin
Faculty of Medicine & Defense Health
National Defence University of Malaysia
Kuala Lumpur, Malaysia
shazreen@upnm.edu.my

Nazrul Fariq Makmor
Faculty of Engineering
National Defence University of Malaysia
Kuala Lumpur, Malaysia
nazrulfariq@upnm.edu.my

Abstract— **The development of artificial neural networks (ANNs) was founded on computer alterations of human biology (the concept of neurons). The practicality of applying ANNs to various problems has been the subject of numerous studies, particularly in the field of biomedical engineering. Medical and educational decision-making regularly use applications to ANNs. Using a range of reference data, the ANNs used in the current study were trained to recognise cardiac abnormalities. Typically referred to as reference parameters, electrocardiogram (ECG) signal amplitude and duration are employed as input parameters for cardiac issues. An ECG complex consists of a P peak, QRS wave, and T peak. The amplitude and length of each P peak, QRS wave, and T peak are measured, resulting in a total of six input parameters for the artificial neural network. The artificial neural network (ANN) structure in this study is a multilayer perceptron (MLP), and the training techniques are Bayesian Regularization (BayR), Lavenberg Marquardt (LevM), and Backpropagation (BackP). The influence of the Tansig activation function on the MLP structure. The MLP network that achieved the highest accuracy (94.44%) utilising the BayR training method and Logsig activation function surpassed all others.**

Keywords- ECG abnormality; amplitude; MLP network; Logsig activation function.

I. INTRODUCTION

An arrhythmia is a particular kind of irregular heartbeat or cardiac activity that affects people and frequently causes unanticipated heartbeats. Activity-triggered, re-entry, and induced fibrillation are a few of these aberrations. Cardiac rhythm irregularities relate to a broad group of diseases where the electrical activity of the heart is abnormal. Abnormal rhythms might be rapid, slow, regular, irregular, or a combination of these [1-2]. Depression is currently the second largest cause of death in Malaysia, after cardiovascular illness, which has exhibited a growing tendency. If monitoring and/or an early diagnosis are carried out as soon as feasible, this condition may be reduced. Arrhythmia, a cardiac activity that can be used to detect early abnormalities, can give clinicians helpful information for figuring out the appropriate level of treatment [1-3]. Based on electrocardiograms, monitoring heart attack episodes is crucial, followed by their prevention (ECG). The ECG signal can be interpreted using the P peak, QRS wave, and T peak. A P peak will result from atrial depolarization, a QRS complex from ventricular depolarization, and a T peak from ventricular depolarization. All of these peaks could be the result of heart-related electrical activity [1-2].

The examination of ECG data is typically employed during clinical trials as a strategy for spotting any heart activity. In the late 1950s, the cardiac activity detecting industry got its start. Each year, more recordings are evaluated, according to statistics. This has led to an increase in study into automatic ECG signal processing and the emergence of the engineering discipline known as biomedical engineering. Over the past two to three decades, numerous types of ECG signal processing algorithms have been created and put to use. The QRS complex detection is the main topic of most studies. There are a number of factors in the ECG signal recording that can help to show the patient's state [4-6]. Applications like ANNs are even adjusted based on actual models of how the brain functions [7]. ANNs are equipped with features such as mapping input data to output data, solving linear and non-linear problems, and neurobiological parallels. In terms of larger-scale applications, ANN have been widely used in the fields of research, technology, finance, and education. Engineering uses artificial neural networks (ANN) extensively for tasks like pattern recognition, data categorization, system identification, image processing, and accuracy.

The focus of this study is on how well forecasts can recognise cardiac failure. The prediction process is carried out using an MLP network after important information has been extracted from ECG signals. Backpropagation (BackP), Lavenberg-Marquardt (LevM), and Bayesian Regularization (BayR) training techniques are used in the study [8–10]. The architecture of the system is founded on a selected algorithm that classifies the patient's condition (either under normal conditions or vice versa). The P peak, QRS wave, and T peak are used to extract the information regarding amplitude and duration as the input parameters for the MLP network. In this instance, the MLP network was activated via the Logsig activation function.

The remaining content in the paper is formatted as follows. In part II, the research on ANNs and MLP networks is reviewed, and an explanation of ECG signals is also provided. The data sample is explained in Section III. Section V offers the conclusion after Section IV shows the results and addresses certain issues of debate.

II. Multilayer Perceptron Network

A computer network called ANN is constructed from neural stem cells taken from the human brain (neuron). The tiny cells known as neurons are what make up the human brain. ANNs offered to represent the structure of neurons by emulating the functions and creativity of the human brain by considering the idea of neurons. For non-parametric grading/regression, data/database classification, and nonlinear function calculation, these ANNs may be represented by mathematical modelling. The behaviour of biological models of human neurons is also simulated using ANN. ANNs can be trained to produce accurate prediction outcomes. Although ANNs might be thought of as an alternative to brain function, they must be able to execute at a level that matches that of the human brain

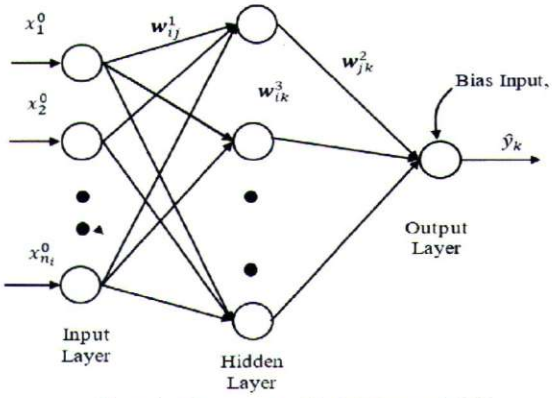


Figure 1. The structure of the MLP network [13].

. Consequently, the created system may serve as a substitute for the system. The MLP network is one of the most widely utilised applications for numerous purposes [11–12].

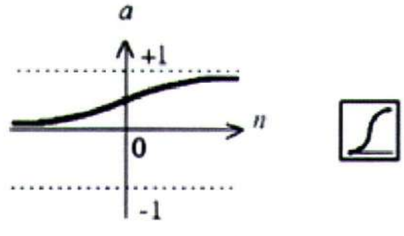
In his research, Perceptron model, Rosenblatt created an artificial neuron model in 1958 [7]. He installs a number of the network layers or sequences seen in Figure 1. In other sections of the study, ANNs are known as MLP networks. According to Figure 1, the input layer, hidden layer, and output layer of an MLP network are represented as follows:

$$\sum_{j=1}^{n_h} w_{jk}^2\, \partial\left(\sum_{i=1}^{n_i} w_{ij}^1 x_i^0(t) + w_{k0}^1 b_j^1\right) \quad (1)$$

$for\ 1 \le j \le n_h$ and $1 \le k \le m$

The weight that connects the MLP network's input layer to the hidden layer itself is shown in Equation 1 as w_{ij}^1. The weights w_{jk}^2 and w_{jk}^2 reflect the connections between the hidden and output layers of the MLP network, respectively. In Equation 1, the input parameter, indicated as x_i^0, is feeding the MLP network through the input layer, whereas b_j^1 is the threshold of the hidden node. The MLP network, indicated as $\partial(\cdot)$ in Equation 1, has been activated in this work using the Tansig activation function. Equation (1) can be used to find the values of w_{ij}^1, w_{jk}^2 and b_j^1 based on a suitable technique and put at the minimal vector but allow for recurrence at each iteration.

Three different training algorithms are utilised to train the MLP network. The BackP, LevM, and BayR algorithms. While BayR uses stochastic tabulation, the training algorithms for BackP and LevM both rely on deterministic tabulation. The Logsig transfer function is related to a bipolar sigmoid. Recently, study was done on the MLP network that Purelin activated. The study compares several activation functions and training techniques. The range of the Logsig activation function returns is 0 to +1. The choice of activation functions for the input data/signal is shown in Figure 2. The threshold of the incoming data or signal must be represented by any location between 0 and +1.



a = logsig(n)

Figure 2. Log-sigmoid activation function [16].

A. Bayesian Regularization (BayR) Algorithm

D. Bayes rule is a training/learning method that was developed by Thomas Bayes. Given the D data, the Bayes rule can be used to determine θ, also known as posterior probability [8]. Typically, the posterior probabilities produced cover the whole range of potential values for. Bayes rule could have come from

$$p(\theta|D) = \frac{p(D|\theta)}{p(D)} \quad (2)$$

Following $p(\theta|D)$, which is referred to as the likelihood of the data before the probability of data, comes $p(\theta)$, which is the initial probability of a parameter. The probability distribution of the network weights was then modified by this general principle as it was applied to the MLP network. The weights of the MLP network are the w that it receives from the training data, $p(w|D)$. One can visualise the posterior distribution by

$$p(w|D) = \frac{p(D|w)p(D)}{p(D)} = \frac{p(D|w)}{\int p(D|w)\,p(w)dw} \quad (3)$$

Using the Bayesian equation as a guide, it can be shown that, as illustrated in Figure 3, the BayR training algorithm taught and optimised the weights that may modify our belief regarding the influences, from the initial $p(w)$ to the posterior $p(D|w)$.

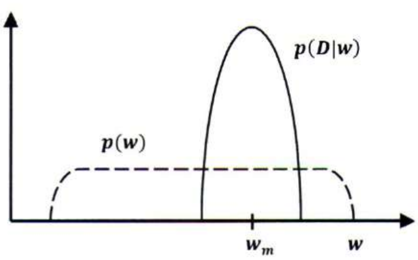


Figure 3. Change pre weights to posterior/post weights [12].

B. Levenberg Marquardt (LevM) Algorithm

A deterministic gradient tabulation underlies the operation of the Levenberg Marquardt (LevM) training algorithm. The BackP algorithm has been improved by the LevM algorithm. The LevM algorithm's profitability after training the MLP network is assessed by contrasting it with the traditional BackP training technique. In addition to being able to converge slightly more quickly, the LevM training algorithm also outperforms the traditional BackP training algorithm in terms of prediction performance. The relative stability is still present [9]. The LevM training technique, which is based on the quasi-Newton approach, was created to accelerate second-order training and get around the Hessian matrix. The sum of the squares is formed by the LevM function, and the estimated Hessian matrix is as follows:

$$H = J^T J \quad (4)$$

and the gradient of the matrix is calculated as:

$$g = J^T \rho \quad (5)$$

where J is a Jacobian matrix that calculates network error using bias and weight of the network. The array can be calculated using conventional BackP methods, which are less difficult than the calculations involving the Hessian matrix [9]. The Newton-like equation estimator is used by the LevM algorithm to update the equations in the Hessian matrix as follows:

$$\Delta w = -[J^T J + \mu I]^{-1} J^T \rho \quad (6)$$

where Δw is the estimator's updated weight, which is determined by the μ parameter. To prevent the parameter reaching 0, the Hessian matrix updated to the Newton-like equation. Since it increases BackP's momentum, the Newton estimator-based BackP training algorithm can deliver better results. With more momentum, the BackP training algorithm can find the lowest local minima and produce the least amount of error. At each step that is effective, the decreases (minimize error). However, when the step function develops, the continues to rise. At each cycle, the minimum error must, however, decrease [9].

C. Back Propagation (BackP) Algorithm

The back propagation training approach has been used in a wide variety of control system and ANN applications. Electrical signals generated by brain activity and coupled to electrical activity or electrical signals in the heart are regularly read ECGs. On the hands, feet, and chest are some electrodes. The electrical activity and electrical impulses generated will be recorded, and this information will be used to train the BackP algorithm. Typically, the algorithm instructs the network using the steepest descent approach. The BackP method can use the learning phase as a parameter or protocol to determine the value of derivatives for practical weight adjustments [8]. According to, BackP has updated the weight between the j^{th} neuron of the hidden layer and the i-layer directly.

$$w_{ji}(t) = w_{ji}(t-1) + \Delta w_{ji}(t) \quad (7)$$

and

$$b_j(t) = b_j(t-1) + \Delta b_j(t) \quad (8)$$

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

$$\Delta w_{ji}(t) = \eta_w \gamma_j(t) x_i(t) + \alpha_w \Delta w_{ji}(t-1) \quad (9)$$

and

$$b_j(t) = \eta_b \gamma_j(t) + \alpha_b \Delta b_j(t-1) \quad (10)$$

The subscripts w and b in Equations 7 and 8 respectively denote the weight and threshold of the steepest decent estimator. The constants α_w and α_b regulate the likelihood of the previous parameter to the present benchmark. The likelihood momentum must be carefully chosen. BackP won't be able to draw lessons from the past if the momentum is set too high. If it is adjusted too low, however, the BackP will continue to behave as it did previously. The BackP training algorithm's η_w and η_b learning rates are noted, whereas $\gamma_j(t)$ represents an error that occurs at the j^{th} neuron. The error afterwards changed into an input and returned to the weight. Equation 11 indicates error at the output node on the assumption that the activation function is functioned in a linear parameter:

$$\gamma(t) = y_k(t) - \hat{y}_k(t) \quad (11)$$

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

$$y_j(t) = F'(x_i(t)) \sum_j \rho_j^k(t) w_{jk}^2(t-1) \quad (12)$$

where $F'(x_i(t))$ is derived from $F(x_i(t))$ by respecting to $x_i(t)$. As a steepest descent approach, the BackP algorithm suffers from a poor convergence speed or rate. The pursuit of a global minimum may lead to local minima. It responds similarly to user-selected parameters [8].

III. Data Sample

The recordings preserved in the MIT-BIH Repository served as the source of all ECG signal samples used in this study. Three of the six different parameters extracted from each signal sample are the amplitude P, QRS and T peaks and the duration of P, QRS and T waves [14-15]. These traits will all serve as input parameters for the newly formed MLP network. The MLP network, which was built using six input parameters, was used to map the system's input-output correlation. MLP will be modelled the relationship between input and the output (target) of the dataset. The MLP network was trained on a total of 1000 datasets, of which 800 were used to ensure that the networks were generalised and optimised and the remaining 200 were used to assess the MLP network's performance.

IV. Results and Discussions

This study used Matlab software and the same method analysis as the study [16-18]. In two alternative interpretations, this work provides data prediction and categorization. The initial study carried out to determine the ideal MLP network structure The maximum MLP network arrangement during this analysis phase was set at 10 hidden nodes. The MLP network's ideal number of iterations will be determined by this analysis. The computation will take longer the more intricate the design is based on structure, training algorithm and activation function has been selected.

TABLE I. MLP OPTIMUM STRUCTURE

Training Algorithm and Algorithm	Num. of iteration
BayR using Logsig	412
BayR using Purelin	716
LevM using Logsig	27
LevM using Purelin	35
BackP using Logsig	21
BackP using Purelin	23

TABLE II. ACCURACY PERFORMANCE ANALYSIS

Training Algorithm	Accuracy Performance (%)		
	Training	Testing	Overall
BayR using Logsig	94.12	94.76	94.44
BayR using Purelin	93.14	93.24	93.19
LevM using Logsig	92.48	92.25	92.37
LevM using Purelin	89.89	89.56	89.73
BackP using Logsig	87.52	86.27	86.90
BackP using Purelin	86.23	85.67	85.80

The accuracy performance of the MLP network, which was trained using the BackP, LevM, and BayR training methods, is displayed in Table I. On the other side, the Logsig, and Purelin activations functions are used to activate each training algorithm. The optimal structure, which was developed from the findings of the first study, is used again in the second analysis. The second analysis is carried out to categorise ECG signals according to whether the patient is in a normal state or not. The ideal MLP network structure is displayed in Table I. To assess the precision of the provided predictions, each BackP, LevM, and BayR training algorithm was put to the test using two different activation functions. As a result, Table II below displays the performance of the prediction capability as the second analysis for both the training and testing phases.

The BayR approach with Logsig activation function may train the MLP network by constructing the best accuracy networks feasible, but it requires 412 hidden nodes, as shown in Tables I and II. The results show that BayR can train MLP networks to build the most fundamental network design, as shown in Table I. In contrast to the BayR training algorithm and the LevM training algorithm activated by Logsig, they require up to 716 nodes to build the simplest network. Table II, on the other hand, demonstrates that, when compared to the performance offered by the others, the MLP network with the BayR training algorithm activated by Logsig is capable of offering the highest accuracy.

V. CONCLUSION

Analyze the ANN approach in use to monitor and classify cardiac abnormalities, determining whether they are normal or not. The results of the study show that ANN is capable of producing excellent prediction accuracy when using the gathered amplitude and duration characteristics from ECG signals as input parameters to MLP networks. The results also show that the BayR algorithm can build the network with the largest topology while producing more precise prediction results when compared to the LevM and BackP training techniques. The activation function for this result was decided to be Logsig, which can activate the input parameter, then Purelin.

ACKNOWLEDGMENT

The authors fully acknowledged the Department of Polytechnic Education & Community College, Ministry of Higher Education of Malaysia (MOHE) and National Defence University of Malaysia (NDUM) for the approved fund, making this important research viable and effective.

REFERENCES

[1] M. A. Quiroz-Juárez, O. Jiménez-Ramírez, R. Vázquez-Medina, E. Ryzhii, M. Ryzhii and J. L. Aragón, "Cardiac Conduction Model for Generating 12 Lead ECG Signals With Realistic Heart Rate Dynamics," in IEEE Transactions on NanoBioscience, vol. 17, no. 4, 2018, pp. 525-532.

[2] U. Satija, B. Ramkumar and M. S. Manikandan, "A New Automated Signal Quality-Aware ECG Beat Classification Method for Unsupervised ECG Diagnosis Environments," in IEEE Sensors Journal, vol. 19, no. 1, 2018, pp. 277-286.

[3] A. Walinjkar and J. Woods, "Personalized wearable systems for real-time ECG classification and healthcare interoperability: Real-time ECG classification and FHIR interoperability," 2017 Internet Technologies and Applications (ITA), Wrexham, 2017, pp. 9-14.

[4] U. Satija, B. Ramkumar and M. S. Manikandan, "Automated ECG Noise Detection and Classification System for Unsupervised Healthcare Monitoring," in IEEE Journal of Biomedical and Health Informatics, vol. 22, no. 3, 2017, pp. 722-732.

[5] F. R. Hashim, J. J. Soraghan, L. Petropoulakis and N. G. N. Daud, "EMG cancellation from ECG signals using modified NLAVMS adaptive filters," 2014 IEEE Conference on Biomedical Engineering and Sciences (IECBES), Kuala Lumpur, 2014, pp. 735-739.

[6] F. R. Hashim, J. Adnan, K. A. Ahmad, S. H. S. Ahmad Jamil and Y. Januar, "MLP Based Tan-Sigmoid Activation Function for Cardiac Activity Monitoring," MATEC Web Conf., 255, 2019.

[7] S. Haykin. Neural Network and Learning Machine Learning, rentice Hall, New Jersey, 2008.

[8] A. Talwar, G. Lokhande, R. Jain and S. Singh, "Estimation of aerodynamic parameters using Cascade Forward Back Propagation," 2017 2nd International Conference on Telecommunication and Networks (TEL-NET), Noida, 2017, pp. 1-6.

[9] Ü. Şentürk, İ. Yücedağ and K. Polat, "Cuff-less continuous blood pressure estimation from Electrocardiogram(ECG) and Photoplethysmography (PPG) signals with artificial neural network," 2018 26th Signal Processing and Communications Applications Conference (SIU), Izmir, 2018, pp. 1-4.

[10] D. Yue, F. Xu, Z. Zhang and Y. Jin, "Quad-pol reconstruction with wishart-Bayesian regularization," 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, 2017, pp. 3933-3936.

[11] N. A. Mat Isa, F. R. Hashim, F. W. Mei, D. A. Ramli, W. M. Wan Omar and K. Z. Zamli, "Predicting quality of river's water based on algae composition using artificial neural network," Industrial Informatics, 2006 IEEE International Conference on, 2006, pp. 1340-1345.

[12] D. A. Ramli, J. M. Saleh, F. R. Hashim and N. A, Mat Isa, "Multilayered perceptron (MLP) network trained by recursive least squares algorithm," Computers, Communications, & Signal Processing with Special Track on Biomedical Engineering, 2005. CCSP 2005. 1st International Conference on, 2005, pp. 288-291.

[13] MIT-BIH Arrhythmia Database – PhysioNet cited from www.physionet.org/physiobank/database/mitdb

[14] J. Adnan, N. G. Nik Daud, S. Ahmad, M. H. Mat, M. T. Ishak, F. R. Hashim and M. M. Ibrahim, "Heart abnormality activity detection using multilayer perceptron (MLP) network," AIP Conference Proceedings, 1930(1), 2018, pp. 020013.

[15] P. Mitra, S. Mitra and S. K. Pal, "Staging of Cervical Cancer with Soft Computing," IEEE Transaction on Biomedical Engineering. vol. 47, no. 7, 2016, pp. 934-940.

[16] S. H. Syam Ahmad Jamil, J. A. Kadir, F. R. Hashim, B. Mustapha, N. S. Hasan and Y. Januar, "Optimization of ECG Peaks for Cardiac Abnormality Detection using Multilayer Perceptron," 2020 10th IEEE International Conference on Control System, Computing and Engineering (ICCSCE), Penang, Malaysia, 2020, pp. 169-173.

[17] F. R. Hashim, S. H. Syam Ahmad Jamil, J. A. Kadir, N. S. Hasan, B. Mustapha and Y. Januar, "Tansig Based MLP Network Cardiac Abnormality," 2019 9th IEEE International Conference on Control System, Computing and Engineering (ICCSCE), Penang, Malaysia, 2019, pp. 199-203.

[17] A. C. Idris, M. R. Saad, Rahman, M. R. A. Rahman, F. R. Hashim and K. Kontis, " Experimental Validation of Artificial Neural Network (ANN) Model for Scramjet Inlet Monitoring and Control," International Journal of Recent technology and Engineering, 7(5), pp. 558-563.

[18] A. C. Idris, M. R. Saad, Rahman, M. R. A. Rahman, F. R. Hashim and K. Kontis, " Experimental Validation of Artificial Neural Network (ANN) Model for Scramjet Inlet Monitoring and Control," International Journal of Recent technology and Engineering, 7(5), pp. 558-563.