BACKPROPAGATION LEARNING ALGORITHM BASED ON LEVENBERG MARQUARDT ALGORITHM

S.Sapna[1], Dr.A.Tamilarasi[2] and M.Pravin Kumar[3]

[1] Assistant Professor, Department of Master of Computer Applications, K.S.R College of Engineering, Tiruchengode-637215, Tamilnadu, India.
sapnaanjusrijumani@rediffmail.com, sapnaanjusri@gmail.com
[2] Prof. & Head, Department of Master of Computer Applications, Kongu Engineering College, Perundurai, Tamilnadu, India.
[3]Assistant Professor, Department of Electronics and Communication Engineering, K.S.R College of Engineering, Tiruchengode-637215, Tamilnadu, India
erodepravin@gmail.com

ABSTRACT

Data Mining aims at discovering knowledge out of data and presenting it in a form that is easily compressible to humans. Data Mining represents a process developed to examine large amounts of data routinely collected. The term also refers to a collection of tools used to perform the process. One of the useful applications in the field of medicine is the incurable chronic disease diabetes. Data Mining algorithm is used for testing the accuracy in predicting diabetic status. Fuzzy Systems are been used for solving a wide range of problems in different application domain and Genetic Algorithm for designing. Fuzzy systems allows in introducing the learning and adaptation capabilities. Neural Networks are efficiently used for learning membership functions. Diabetes occurs throughout the world, but Type 2 is more common in the most developed countries. The greater increase in prevalence is however expected in Asia and Africa where most patients will likely be found by 2030. This paper is proposed on the Levenberg – Marquardt algorithm which is specifically designed to minimize sum-of-square error functions. Levernberg-Marquardt algorithm gives the best performance in the prediction of diabetes compared to any other backpropogation algorithm.

KEYWORDS

Data Mining, Diabetes, Fuzzy Systems, Genetic Algorithm (GA), Levenberg – Marquardt, Neural Networks.

1. INTRODUCTION

In this paper an intellectual and efficient diabetic prediction method with the support of fuzzy neural network is proposed. The network is trained using multi layer feed forward back propagation algorithm to test its performance. The proposed method will not only assist medical practitioners but also supports special educators, occupational therapists and psychologist in better assessment of diabetes disease. Usually the decisions made by the medical experts are seldom based on the single symptom because of the complication of the human body, as one symptom could indicate any number of problems. A skilled medical expert is far more likely to make a sound assessment than a beginner, because from his past knowledge he knows what to

Sundarapandian et al. (Eds): CoNeCo,WiMo, NLP, CRYPSIS, ICAIT, ICDIP, ITCSE, CS & IT 07,
pp. 393–398, 2012. © CS & IT-CSCP 2012 DOI : 10.5121/csit.2012.2438

Digital Global Systems
Exhibit 1011

look out for and what to inquire, and may have etched on his mind a past inaccuracy, which he neither will nor repeat. Thus the senior medical expert is in a higher position than the trainee. Similarly it would be helpful if machines, too, could use past events as part of the factors on which their decisions are based, and this is the role that neural network seeks to fill.

The back propagation algorithm was developed by Paul Werbos in 1974 and rediscovered independently by Rumelhart and Parker [4]. Since its rediscovery, the back propagation algorithm has been widely used as a learning algorithm in feed forward multilayer neural networks. ANN using the Back Propagation (BP) algorithm performs parallel training for improving the efficiency of Multilayer Perceptron (MLP) network. It is the most popular, effective, and easy to learn model for complex, multilayered networks. A Backpropagation is a supervised learning technique which is based on the Gradient Descent (GD) method that attempts to minimize the error of the network by moving down the gradient of the error curve as stated [2,6]. The most popular in the supervised learning architecture because of the weight error correct rules [3,5]. It is considered a generalization of the delta rule for nonlinear activation functions and multilayer networks.

The structure of layered feedforward neural networks is considered and each of these networks consists of a set of inputs and one or more layers of parallel neurons. Inputs are connected only to neurons in the first layer with the exception of the special input X_0, representing the bias of each neuron, which is connected to all neurons in the network. Neurons in one layer are connected to all neurons in the next layer. No other connections exist in neural networks of this type. The last layer, which produces the output of the network, is called an output layer. Any layers that precede the output layer are called hidden layers. The set of inputs is sometimes referred to as an input layer. Inputs don't do any computation; their only role is to feed input patterns into the rest of the network.

2. BACK PROPAGATION LEARNING ALGORITHM BASED ON LEVENBERG MARQUARDT ALGORITHM (LM)

Levenberg – Marquardt algorithm is specifically designed to minimize sum-of-square error functions [1], of the form.

$$E = \tfrac{1}{2} \sum k (e_k)^2 = \tfrac{1}{2} \|e\|^2 \qquad \ldots (2.1)$$

Where e_k is the error in the kth exemplar or pattern and e is a vector with element e_k. If the difference between the pervious weight vector and the new weight vector is small, the error vector can be expanded to first order by means of a Taylor series.

$$e(j+i) = e(j) + \partial e_k / \partial w_i \left(w(j+1) - w(j)\right) \qquad \ldots (2.2)$$

As a consequence, the error function can be expressed as

$$E = \tfrac{1}{2} \left\| e(j) + \partial e_k / \partial w_i \left(w(j+1) - w(j)\right) \right\|^2 \qquad \ldots (2.3)$$

Minimizing the error function with respect to the new weight vector, gives

$$w(j+1) = w(j) - \left(Z^T Z\right)^{-1} Z^T e(j) \quad \ldots (2.4)$$

where $(Z)_{ki} \equiv \partial e_k / \partial w_i$

Since the Hessian for the sum-of-square error function is

$$(H)_{ij} = \partial^2 E / \partial w_i \partial wj = \Sigma \left\{ (\partial e_k / \partial w_i)(\partial e_k / \partial w_i) + e_k \partial^2 e_k / \partial w_i \partial wj \right\} \quad \ldots (2.5)$$

Neglecting the second term, the Hessian can be written as $H = Z^T Z$ … (2.6)

Updating of the weights therefore involves the inverse Hessian or an approximation thereof for nonlinear networks. The Hessian is relatively easy to compute, since it is based on first order derivatives with respect to the network weights that are easily accommodated by backpropagation. Although the updating formula could be applied iteratively to minimize the error function, this may result in a large step size, which would invalidated the linear approximation on which the formula is based.

In the Levenberg-Marquardt algorithm, the error function is minimized, while the step size is kept small in order to ensure the validity of the linear approximation. This is accomplished by use of a modified error function of the form.

$$E = \tfrac{1}{2} \left\| e(j) + \partial e_k / \partial w_i \left(w(j+1) - w(j) \right) \right\|^2 + \lambda \left\| w(j+1) - w(j) \right\|^2 \quad \ldots (2.7)$$

where λ is a parameter governing the step size. Minimizing the modified error with respect to $w(j+1)$ gives

$$w(j+1) = w(j) - \left(Z^T Z + \lambda I \right)^{-1} Z^T e(j) \quad \ldots (2.8)$$

very large values of λ amount to standard gradient descent, while very small values λ of amount to the Newton method.

3. PERFORMANCE EVALUATION USING LEVENBERG-MARQUARDT ALGORITHM

For this implementation, the data set collected from SRC Diabetic Care Centre, Erode District, Tamilnadu in South India is used. Levenberg-Marquardt back propagation algorithm is used for training the network [7,8]. Training automatically stops when generalization stops improving, as indicated by an increase in the Mean Square Error (MSE) of the validation samples. The Mean Squared Error (MSE) is the average squared difference between outputs and targets. Lower values are better while zero means no error. Regression R analysis is performed to measure the correlation between outputs and targets.

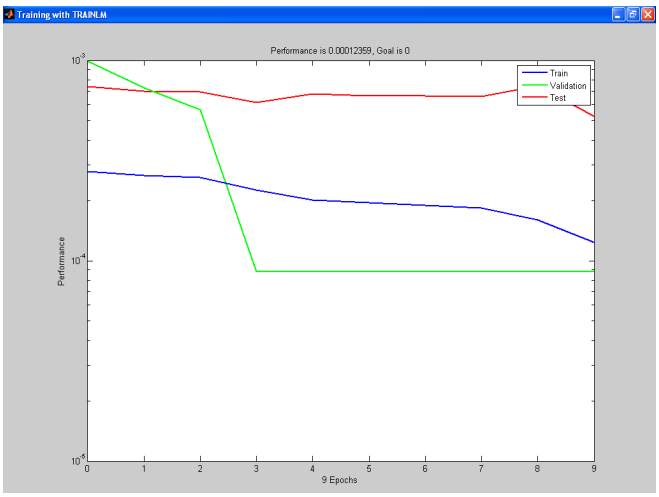

Figure 3.1 Performance of Levenberg-Marquardt Backpropagation Algorithm

An R value of 1 means a close relationship, 0 a random relationship. The performance of the proposed network when trained with Levenberg-Marquardt backpropagation algorithm using Matlab R2007b, is shown in Figure 3.1.

From Figure 3.1 it is observed that the best validation performance 0.00012359 at epoch 9 is obtained. The Regression plot shown in Figure 3.2 shows the perfect correlation between the outputs and the targets.

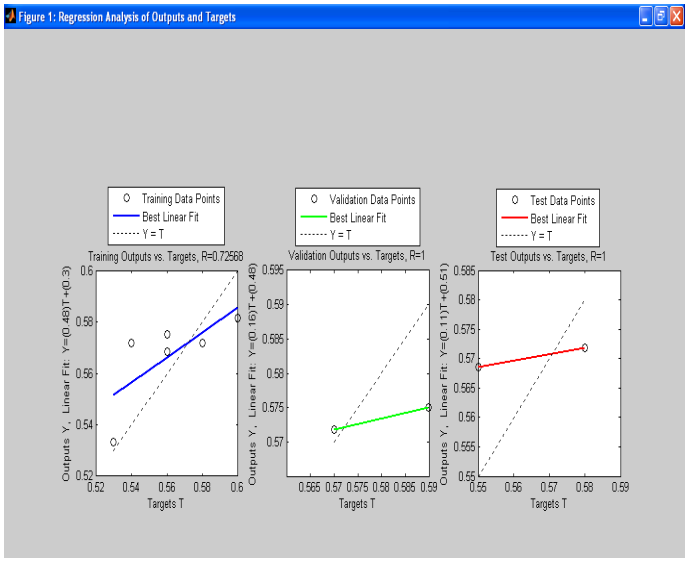

Figure 3.2 Regression Analysis Plot - Levenberg-Marquardt Backpropagation Algorithm

Table 3.1 The Mean Square Error (MSE) and Regression (R) values for the Training, Validation and Testing

	MSE	R
Training	2.25880e-4	0.902515
Validation	8.83702e-5	1.000000
Testing	6.15467e-4	1.000000

The result for training, validation and testing samples is illustrated in Table 3.1. It is observed that the value of R is closest to 1 indicating the accurate prediction. When the data set was trained in Levenberg – Marquardt algorithms the performance obtained was in 9 epochs. Levenberg – Marquardt algorithm (LM) is the most widely used optimization algorithm. LM algorithm is an iterative technique that locates a local minimum of a multivariate function that is expressed as the sum of squares of several non-linear, real-valued functions. It has become a standard technique for non linear least-square problems, widely adopted in various disciplines for dealing data-fitting applications. Levenberg-Marquardt curve-fitting method is actually a combination of two minimization methods the gradient descent method and the Gauss-Newton method.

This paper presented a standard nonlinear least squares optimization algorithm, and showed how to include it into the backpropagation algorithm. The Marquardt algorithm was experienced on several function approximation problems, and it was compared with the conjugate gradient algorithm and with variable learning rate backpropagation. The results indicated that the Marquardt algorithm is very efficient when training networks which have up to a few hundred weights. Although the computational requirements are much higher for each iteration of the Marquardt algorithm, this is more than made up for by the increased efficiency. This is specially true when high accuracy is essential. It is also found that in many cases the Marquardt algorithm converged when the conjugate gradient and variable learning rate algorithms failed to converge.

4. CONCLUSION

This paper aimed to evaluate the artificial neural network in predicting diabetes disease. The feed-forward backpropagation neural network with supervised learning is proposed to diagnose the disease. The reliability of the proposed neural network method depends on data collected from the patients and experts opinion. Backpropagation learning algorithm is used to train the feedforward neural network to perform a given task based on Levenberg-Marquardt algorithm and also the performance is analyzed. It is analyzed that Levernberg-Marquardt algorithm gives the best performance in the prediction of diabetes compared to any other backpropogation algorithm. The proposed diagnosis based on neural network showed significant results in identifying the diabetes.

REFERENCE

[1] Aldrich .C (2002), “Exploratory Analysis of Metallurgical Process Data with Neural Networks and Related Methods”, Elsevier Science, British Library Cataloguing in Publication Data, Netherlands, Vol.1, pp.56-57.

[2] Alsmadi M. K. S., K. Omar, and S. A. Noah (2009), “Back Propagation Algorithm: The Best Algorithm among the Multi-layer Perceptron Algorithm”, International Journal of Computer Science and Network Security, pp. 378-383.

[3] Insung Jung, Lockjo Koo and Gi-Nam Wang (2007), "Two States Mapping Based Neural Network Model for Decreasing of Prediction Residual Error", International Journal of Electrical and Computer Engineering, Vol.2, No.8, pp.531-537.

[4] Kevin L. Priddy and Paul E.Keller (2005), "Artificial Neural Networks An Introduction", SPIE-The International Society of Optical Engineering, Washington.

[5] Mutasem khalil Sari Alsmadi, Khairuddin Bin Omar and Shahrul Azman Noah (2009), "Back Propagation Algorithm: The Best Algorithm among the Multi-layer Perceptron Algorithm", International Journal of Computer Science and Network Security, Vol.9, No.4, pp.378-383.

[6] Norhamreeza Abdul Hamid and Nazir Mohd Nawi (2011), "Accelerating Learning Performance of Back Propagation Algorithm by Using Adaptive Gain Together with Adaptive Momentum and Adaptive Learning Rate on Classification Problems", International Journal of Software Engineering and its Applications, Vol.5, No.4, pp.31-44.

[7] Martin T. Hagan and Mohammad B. Menhaj (1994), "Training Feedforward Networks with the Marquardt Algorithm", IEEE Transactions on Neural Networks, Vol. 5, No. 6, November 1994.

[8] Sivanandam .S.N, Sumathi.S, Deepa .S.N (2008), "Introduction to Neural Network Using MATLAB 6.0", Tata McGraw-Hill Publishing Company Limited, New Delhi.

BIOGRAPHY

S.Sapna received her B.Sc Degree, M.C.A and M.Phil Degree, from Bharathiar University. She is currently working as Assistant Professor in the Department of MCA, K.S.R. College of Engineering. She has presented more than 20 papers on various topics including national, international conference and journals. She is a research scholar of Mother Teresa Women's University, Kodaikanal. Her research interest includes Soft Computing, Data Mining, Mathematical Computations and Networks. She is a life member of ISTE and CSI.

Dr.A.Tamilarasi, currently serving as Professor and Head, Department of M.C.A, Kongu Engineering College. She has published various papers in the filed of Fuzzy Logics. She has published various books in mathematical field. She is guiding several research scholars in her area of interest like Data Mining, Soft Computing and Networks.

M.Pravin Kumar received his B.E & M.E Degree, from Anna University. He is currently working as Assistant Professor in the Department of ECE, K.S.R. College of Engineering. He has presented more than 8 papers on various topics including national, international conference and journals. He is a research scholar of Anna University, Coimbatore. His research interest includes Soft Computing and Networks. He is a life member of ISTE.