International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169
Volume: 3 Issue: 2 378– 382

Comparative Study of various training algorithms of Artificial Neural Networks on Diabetes dataset

Sumi Alice Saji
Department of Computer Science and Engineering
Christ University
Bangalore, India
sumialicesaji@gmail.com

Balachandran K
Department of Computer Science and Engineering
Christ University
Bangalore, India
balachandran.k@christuniversity.in

Abstract-Artificial Neural Networks play an important role in diverse applications. They are a system of interconnected neurons that perform computation similar to a biological neural network. This study focuses on various training algorithms used in artificial neural networks. We have used the Pima Indian diabetes data set from UCI Machine Learning Repository for constructing and training the neural network. The system is implemented in MatlabR2013. About 768 instances were used for training the neural network. The input vector has patient history and the target output will be class label as tested positive or tested negative. Results show that out of the training algorithms, Levenberg-Marquardt Algorithm performed the best results. Even though, Resilient Backpropagation and Conjugate gradient with Powell/Beale Restarts Algorithms perform invariably, LM performs faster than other two algorithms.

Keywords—*Diabetes Mellitus, Levenberg-Marquardt, Artificial Neural Network, Backpropagation, Powell-Beale*

*****

I. INTRODUCTION

An artificial neuron can be defined as a function that processes and transmits the information as a biological neuron does. Artificial neurons can mimic biological neurons in various aspects. They receive more than one input and perform the summation and outputs the result as the biological neuron does with dendrites. There are different types of neural networks from simple to complex. There are mainly three types of layers in a neural network namely input layer, hidden layer and the output layer. One of the main features of a neural network is its iterative learning process. The data values are presented to the input layer one at a time and weights are adjusted each at a time. Thus after all data values are presented, learning starts over again. The structure of a biological neuron is given in Fig. 1.

A typical biological neuron consists of a cell body, dendrites, axon, axon hillock, synaptic terminals and synapse. In terms of an artificial neuron, neuron can be termed as a cell capable of transmitting nerve impulses; a nerve cell [lm]. Axon can be defined as a long process of a nerve fiber that conducts impulses away from the body of the nerve [5]. Synapse is the link between two nerve cells, consisting of a minute gap across which impulses pass by diffusion of a neurotransmitter [5]. Dendrites are the short branched extension of a nerve cell. Along which impulses received from other cells at synapses are transmitted to the cell body.

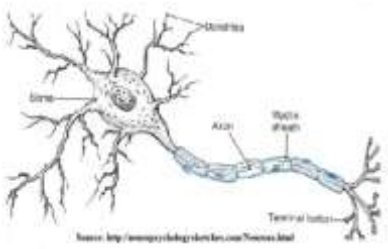


Fig.1. Structure of a neuron

Diabetes Mellitus is a group of metabolic diseases characterized by high blood sugar (glucose) levels that result from defects in insulin secretion, or its action or both. Diabetes Mellitus, commonly referred to as Diabetes was first identified as a disease associated with “sweet urine” and excessive muscle loss in the ancient world.

There are mainly three types of Diabetes Mellitus. Type1 Diabetes Mellitus results due to the body’s failure in the production of enough insulin. This condition is also called as Juvenile Diabetes. Type2 is a condition in which the cells fail to produce insulin. This condition is also called as Non-Insulin Dependent Diabetes Mellitus or adult-onset Diabetes. The primary reason is excessive body weight and not enough exercise. Another type is Gestational Diabetes which occurs when pregnant women without a previous history of Diabetes develop a high blood glucose level.

According to American Diabetes Association study, it is said that in 2012, 29.1 million Americans, or 9.3% of the population had Diabetes. Also Diabetes was the seventh leading cause of

death in US in 2010. In addition World Health Organization estimates that about 200 million people all over the world suffer from Diabetes and there is a probability to be doubled by 2030. A study done by Public Health Foundation of India shows that about 44lakh Indians aged from 20 to 79 years is not aware that they are diabetic. In India, there are nearly 50 million Diabetic patients as per the statistics given by International Diabetes Foundation.

Data mining and Artificial Intelligence finds its tremendous applications in medical field. Data mining mainly uses classification and prediction schemes in medicine. In this paper, we focus on various training algorithms used in artificial neural networks.

II. RELATED WORK

Numerous research works is going on in the field of Data Mining as well as Artificial Intelligence. Different system models are being proposed for artificial neural networks in various domains. Artificial Neural Networks was applied in medical field for various tasks [7], [8]. Numerous techniques are being used for Diabetes Prediction. An Adaptive Neuro Fuzzy Inference System (ANFIS) is being used for the diagnosis of Diabetes and prediction of Cancer [1]. In paper [1], the experimental results show that the classification accuracy is better than existing approaches. In [2], author proposes an Alzheimer's disease prediction model that can assist medical professionals in predicting the status of the disease based on medical data about patients.

A genetic neural network based data mining technique is used for the heart disease prediction [3]. In this paper, a set of risk factors are analyzed. In PVARM algorithm [3], large number of irrelevant rules is reduced and a new set of rules with high confidence is produced. It is very much used in mining medical data sets.

In, [4], Zhao et. al. does the comparative study of BP algorithms in MATLAB neural network toolbox and proves that trainlm is having the best convergence.

A multi-layer perceptron is created by LMA for classifying Million Song Dataset from UCI Machine Learning Repository using error back propagation in [5]. In this paper, an optimized Levenberg-Marquardt algorithm is implemented by harnessing thread level parallelism in Java, to enhance performance of training a neural network [5].

Artificial Neural Networks is also applied using back propagation neural network along with binary classification to predict Diabetes [6].

Rajendra Acharya et al. used artificial neural network and Fuzzy equivalence relation to classify several diseases. They did a comparison between ANN and Fuzzy classifier on the same dataset [9]. In [10], a naïve creedal classifier was used to perform mining of hidden patterns on the lung cancer dataset

Another interesting approach was done by Agarwal et al in [11]. They proposed a predictive model for Diabetes by using Hybrid Kernel approach. They have utilized the benefits of both Gaussian and Polynomial kernels function and proposed a hybrid kernel function which was based on Twin Support Vector Machine [11].

III. TRAINING OF ARTIFICIAL NEURAL NETWORK

The training of artificial neural network mainly consists of two types of training, supervised and unsupervised training. Both of the training differ only in the causal structure of the model.

A. Supervised Training

Supervised Training involves the process in which we have the inputs as well as desired outputs.In this type of training, a predefined idea of desired output is considered.

B. Unsupervised Training

Unsupervised Training is the process in which all observations are assumed to be caused by a set of latent variables.

C. Backpropagation Algorithms

Backpropagation is considered as a supervised learning method in which a known desired output is needed to calculate the loss function gradient. The different backpropagation algorithms used in this paper are Levenberg-Marquardt, Resilient Backpropagation, Conjugate gradient with Powell/Beale Restarts, Fletcher Powell Conjugate Gradient and Polak- Ribiere Conjugate Gradient Algorithms.

Levenberg–Marquardt Algorithm is a Gauss Newton method for solving least-squares problem. The main objective of this algorithm is to find a minimum of a function F(x) that is a sum of squares of non-linear functions. Levenberg's damping factor is adjusted at each iteration and that guides the optimization problem. If a smaller value of damping factor is chosen, the algorithm behaves closer like a Gauss-Newton whereas if larger damping factor is used, then the algorithm behaves like a gradient descent algorithm.

Resilient Backpropagation is a first order optimization algorithm for supervised learning in feedforward artificial neural networks. It is one of the fastest weight update algorithm.

Conjugate gradient with Powell/Beale Restarts is a backpropagation algorithm which updates the weight in the steepest descent direction (negative of the gradient). In this direction, the performance function decreases almost rapidly.

In MATLAB, different functions represent different training algorithms. Table 1. shows the functions used for different backpropagation algorithms.

TABLE 1. BACKPROPAGATION ALGORITHMS

Acronym	Function	Algorithm
LM	trainlm	Levenberg-Marquardt Algorithm
RP	trainrp	Resilient Backpropagation
CGB	traincgb	Conjugate gradient with Powell/Beale Restarts

D. Training Data set Classification

Table2. shows the attributes chosen from the Pima Indian Diabetes data set. In the dataset, each of the attribute is a numeric value and a class label is assigned to identify whether the patient is tested positive or tested negative to Diabetes. About 768 instances are taken for training, testing and validation of data set. The data set is divided in the ratio of 70:15:15 for training the neural network.

All the input parameters are having numeric values. The first attribute is the total number of times the patient is pregnant. The second attribute denotes the plasma glucose level 2 hours in an oral. The third attribute is the diastolic blood pressure rate measured in mm by Hg. The next attribute shows the triceps skin fold thickness measured in mm. The fifth attribute denotes 2-hours serum insulin test. This test finds the amount of insulin produced in the patient's body. The sixth attribute denotes the patient's body mass index. Body Mass Index (BMI) is calculated as

BMI = Patient's weight in kg/ (Patient's height in meter) ^2

The seventh attribute is the Diabetes Pedigree. It denotes the function value based on diabetes family hierarchy. The last attribute is age.

TABLE 2. TRAINING DATASET

Number of times pregnant
Plasma Glucose Level
Diastolic Blood Pressure (mm Hg)
Skin rashes and thickness (mm)
2-Hrs Serum Insulin
Body Mass Index
Diabetes Pedigree
Age

E. Neural Network Model

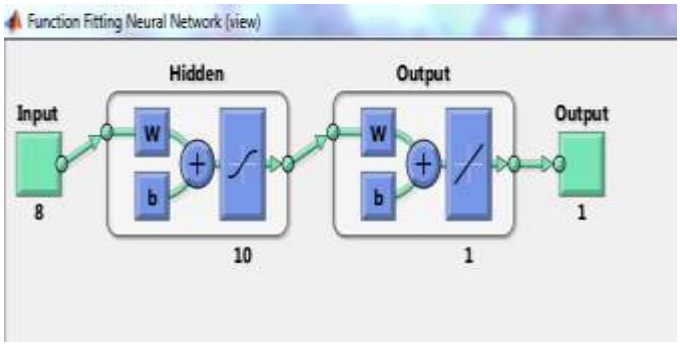


Fig. 2. Model of a neural network in Mat lab

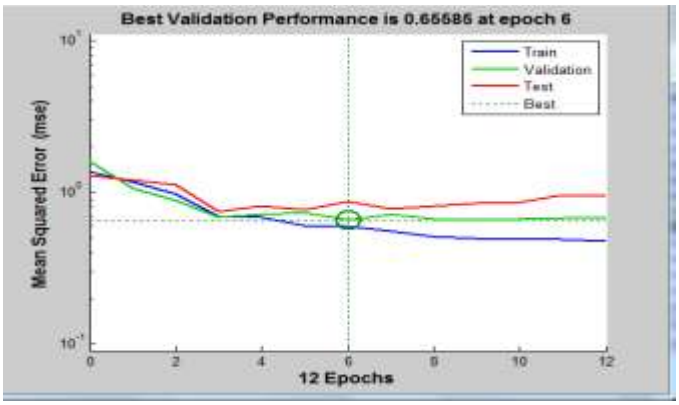
In Fig.2 the model consists of three layers namely, input, hidden and output layers. The input layer has 8 neurons which represents the 8 attributes present in the Pima Indian Diabetes Dataset. There are about 10 hidden layers and the output layer as 1. The number of neurons taken in output layer is 1.

F. Experimental Setup

- Levenberg-Marquardt Algorithm based Trainin

In Levenberg-Marquardt Algorithm based training, a trainlm function which is configured in Neural Network Toolbox was used. The parameters of trainlm function used are …..

Fig. 3. Performance Plot of Levenberg-Marquardt Training



- Resilient Backpropagation Algorithm based training

In Resilient Backpropagation Algorithm based training, a trainrp function which is configured in Neural Network Toolbox was used. The parameters of trainrp function used are

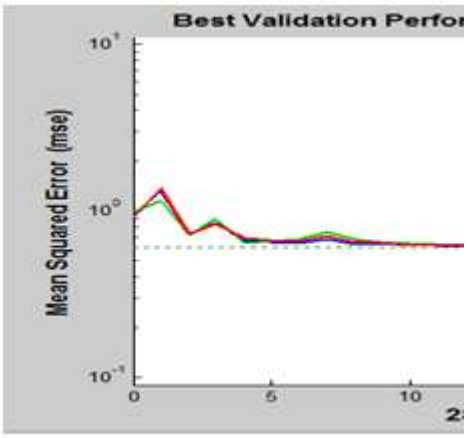


Fig. 4. Performance Plot of Resilient Backpropagation Algorithm based Training

- Conjugate gradient with Powell/Beale Restarts Algorithm based training

In Conjugate gradient with Powell/Beale Restarts Algorithm based training, a traincgb function which is configured in Neural Network Toolbox was used. The parameters of traincgb function used are

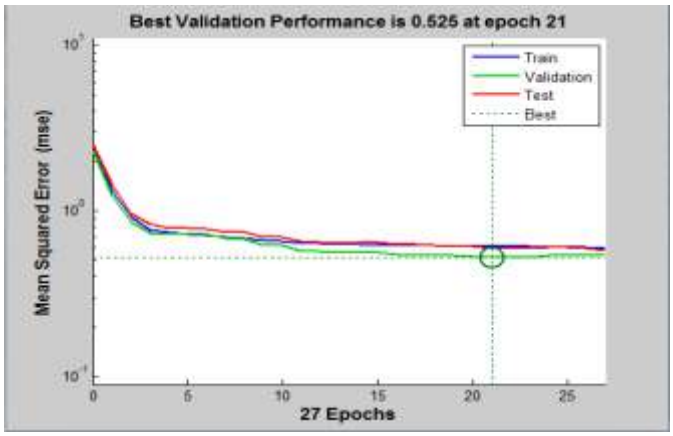


Fig. 5. Performance Plot of Conjugate gradient with Powell/Beale Restarts Training

IV. RESULTS

The comparative study of training algorithms is implemented in the working platform of MATLAB. A graph is plotted in Fig. 6 which indicates the precedence of Levenberg-Marquardt Algorithm over other two training algorithms.

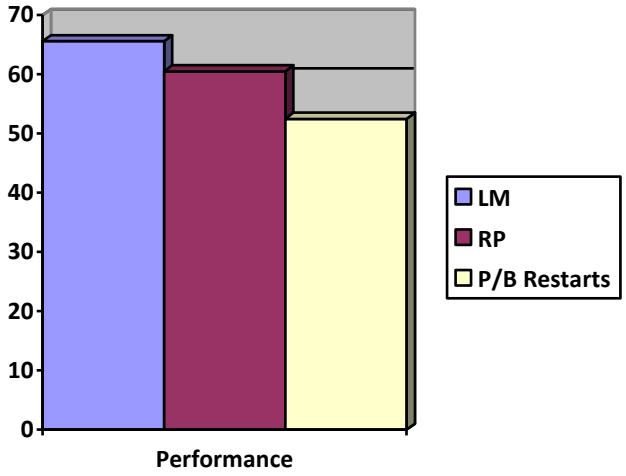


Fig. 6. Levenberg-Marquardt, Resilient and Powell/Beale Restarts in terms of Performance

There are different algorithm properties that can be deduced from the experiments. In general, on function approximation problems, for networks that contain up to a few hundred weights, the Levenberg-Marquardt algorithm will have the fastest convergence. This advantage is especially noticeable if very accurate training is required. In many cases, trainlm is able to obtain lower mean square errors than any of the other algorithms tested. However, as the number of weights in the network increases, the advantage of trainlm decreases. The storage requirements of trainlm are larger than the other algorithms tested.

The trainrp function is the fastest algorithm on pattern recognition problems. However, it does not perform well on function approximation problems. Its performance also degrades as the error goal is reduced. The memory requirements for this algorithm are relatively small in comparison to the other algorithms considered. The Powell/Beale Restarts algorithm also performs well for function approximation problems but it does not provide convergence easily

V. CONCLUSION

The work we have done in this research is a comparative study of three different training algorithms in artificial neural networks. The system was implemented on a huge set of diabetes data and it was observed that LM algorithm outperforms and achieves faster convergence compared to Resilient Backpropagation and Powell/Beale Restarts Algorithm. In future, the performance analysis of LM Algorithm can be compared with fuzzy systems and other optimization based algorithms.

REFERENCES

[1] C. Kalaiselvi and D. G. .. Nasira, "A New Approach for Diagnosis of Diabetes and Prediction of Cancer using ANFIS," in 2014 World Congress on Computing and Communication Technologies, 2014.

[2] D. AL-Dlaeen and A. Alashqur, "Using Decision Tree Classification to Assist in the Prediction of Alzheimer's Disease," in 6th International Conference on CSIT, 2014.

[3] S. U. Amin, K. Agrawal and D. R. Beg, "Genetic Neural Network Based Data Mining in Prediction of Heart Disease Using Risk Factors," in Proceedings of 2013 IEEE Conference on Information and Communication Technologies, 2013.

[4] Z. Zhao, H. Xin, Y. Ren and X. Guo, "Application and Comparison of BP Neural Network Algorithm in MATLAB," 2010.

[5] N. L. R. S and K. Balachandran, "Performance Optimization of Levenberg-Marquardt Algorithm with Parallelization," 2014.

[6] R. Dey, V. Bajpai, G. Gandhi and B. Dey, "Application of artificial neural network technique for diagnosing diabetes mellitus," in IEEE, Kharagpur, India, 2008.

[7] I. Kononenko, "Machine learning for medical diagnosis: history,state of the art and perspective," Artificial Intelligence in Medicine, vol. 23, no. 1, pp. 89-109, 2001.

[8] R. Andrews, J. Diederich and A. B. Tickle, "Survey and critique of techniques for extracting rules from trained artificial neural networks," Knowledge-Based Systems, vol. 8, no. 6, pp. 373-389, 1995.

[9] U. R. Acharya and e. al., "Classification of heart rate data using artificial neural network and fuzzy equivalence relation," Pattern Recognition, vol. 36, pp. 61-68, 2003.

[10] V. Krishnaiah, G. Narsimha and N. S. Chandra, "Diagnosis of Lung Cancer Prediction System Using Data Mining Classification Techniques," International Journal of Computer Science and Information Technologies, vol. 4, pp. 39-45, 2013.

[11] T. Divya and A. Sonali, "Predictive Model for diabetic patients using Hybrid Twin Support Vector Machine," in 5th International Conferences on advances in Communication Network and Computing, In Press 2014.

[12] H. Jiawei, M. Kamber and J. Pei, Data mining: concepts and techniques, Morgan kaufmann, 2006.

[13] [Online]. Available: http://archive.ics.uci.edu/ml/datasets. [Accessed December 2014].

[14] J. G. Yang and S. Y. Weng, Applied Textbook of Artificial Neural Network, Hangzhou: Zhejiang University Press, 2001.

[15] K. Levenberg, "A method for the solution of certain problems in least squares," Quart. Appl. Math, vol. 2, pp. 164-168, 1944.

[16] D. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," SIAM J.Appl.Math, vol. 11, pp. 431-441, 1963.

[17] "Wikipedia,"[Online].Available: https://en.wikipedia.org/wiki/Sensitivity_and_specificity.

[18] S. F. B. Jaafar and D. M. Ali, "Diabetes mellitus forecast using artificial neural networks," in Asian Conference of paramedical research proceedings, Kuala Lumpur, Malaysia, 5-7 September,2005.

[19] T. Jayalakshmi and D. A. Santhakumaran, "A novel classification method for classification of diabetes mellitus using artificial neural networks," in International Conference on Data Storage and Data Engineering, 2010.

[20] A. G. Karegowda, A. Manjunath and M. Jayaram, "Application of Genetic Algorithm Optimized Neural Network Connection Weights For Medical Diagnosis Of Pima Indian Diabetes," International Journal on Soft Computing, vol. 2, no. 2, May,2011.