

Diabetes Prediction with Supervised Learning Algorithms of Artificial Neural Network

Muhammad Akmal Sapon¹, Khadijah Ismail¹, Suehazlyn Zainudin² and Chew Sue Ping¹

¹ Universiti Pertahanan Nasional Malaysia, Kuala Lumpur

² Department of Endocrine, Pusat Perubatan Universiti Kebangsaan Malaysia, Kuala Lumpur

Abstract. Diabetes is one of metabolic diseases where a patient has high blood sugar either caused by the body failure to produce enough insulin or the cells failure to respond to the produced insulin. Diabetes can be identified by studying on several readings taken from a patient such as fasting glucose, sodium, potassium, urea, creatinine, albumin and many others. This paper presents a study on the prediction of diabetes with different supervised learning algorithms of Artificial Neural Network. Fletcher-Powell Conjugate Gradient, Polak-Ribienko Conjugate Gradient and Scaled Conjugate Gradient algorithms are used for diabetes prediction with the collected data. The network is trained using the data of 250 diabetes patients between 25 to 78 years old. Discussion on the performance of each algorithm is provided by performing regression analysis to investigate the correlation between the predicted and target results. The values of correlation coefficient, R of the algorithms are compared to determine the best algorithm for the diabetes prediction task.

Keywords: diabetes, supervised, regression.

1. Introduction

Many studies have been carried out on supervised learning neural networks. A new supervised training algorithm known as Multi-Spike Prop is developed for multiple spiking neural networks for the epilepsy and seizure detection [1]. On the other hand, supervised neural network is used to control industrial processes [2]. The performance of supervised training methods such as Backpropagation with Variable Stepsize (BPVS), Fletcher Reeves (FR), Polak-Ribienko (PR), Davidon-Fletcher-Powell (DFP) and Broyden-Fletcher-Goldfarb-Shanno (BFGS) are analyzed [3]. The algorithms are used for pattern classification and continuous function approximation problems.

Supervised neural networks also find applications in medical areas. Multilayer neural network is trained using Levenberg-Marquardt (LM) algorithm to diagnose diabetes [4]. The results obtained showed that neural network could produce successful diabetes predictions compared to some conventional validation methods. Neural network algorithms are proven useful for early diagnosis of acute myocardial infarction and prediction of infarct size in chest pain patients [5]. Neural network with different network structures are implemented to diagnose chest diseases such as tuberculosis, COPD, pneumonia, asthma and lung cancer [6].

2. Methodology

2.1. Dataset

The data of 250 diabetes patients is obtained from Pusat Perubatan Universiti Kebangsaan Malaysia, Kuala Lumpur. The data consists of 27 variables such as blood pressure, creatinine, urine PH, fasting glucose and many others are coded as numeric values. The patients are among both male and female between 25 to 78 years old. The data is trained to identify the diabetes pattern using MATLAB neural network toolbox. 125 data is used for the training purpose, whereas 125 data is used for testing stage.

2.2. Supervised Learning

Supervised learning is an approach to find the input-output relationship based from the training using a set of data. Fig. 1 represents the block diagram of supervised learning. Learning system is fed with the input data and generates output, which is then compared with the target to compute the error signal by arbitrator. The error is sent to the learning system for further training until the minimum value of error is generated.

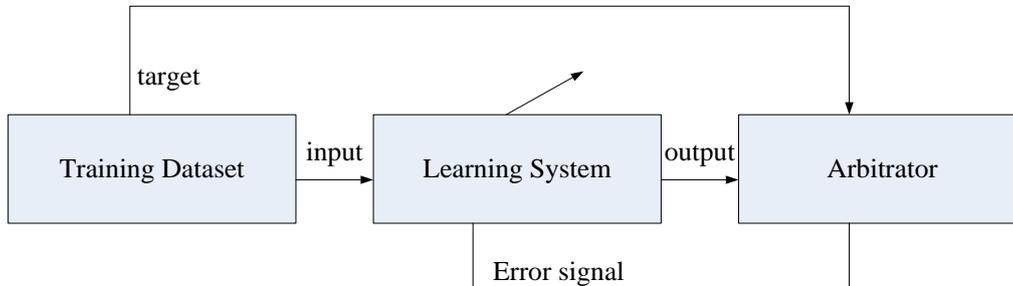


Fig. 1: Block diagram of supervised learning.

Backpropagation algorithm is an example of this type of learning and is widely implemented for various purposes. The learning process takes place by inputting the data to be trained by the network. The information from the input layer is distributed to the hidden layer for information process. Then the output layer will further process the information to obtain the results. The predicted outputs are then compared with the desired values for error computations.

2.3. Training Algorithms

There are many training algorithms available to be used in MATLAB neural network toolbox. In this study Fletcher-Powell Conjugate Gradient, Polak-Ribière Conjugate Gradient and Scaled Conjugate Gradient are used to train the network to predict diabetes.

Fletcher-Powell Conjugate Gradient algorithm updates weight and bias with Fletcher-Reeves updates. Two new dependent methods is proposed and their global convergence properties are discussed with different line searches [7]. On the other hand, discussions on the global convergence of the Polak-Ribière-Polyak (abbreviated PRP) conjugate gradient method is presented for unconstrained optimization problems [8].

Scaled Conjugate Gradient algorithm is introduced for fast supervised learning and the performance is compared with the standard backpropagation algorithm (BP), conjugate gradient algorithm with line search (CGL) and one-step Broyden-Fletcher-Goldfarb-Shanno memoriless quasi-Newton algorithm (BFGS) [9].

2.4. Regression Analysis

Regression analysis is performed to investigate correlation between the desired and predicted results based on the value of correlation coefficient, R. The perfect fit between the training data and the produced results is indicated by the value of R which is equal to 1.

3. Results

The performances of all algorithms are presented in Fig. 3.

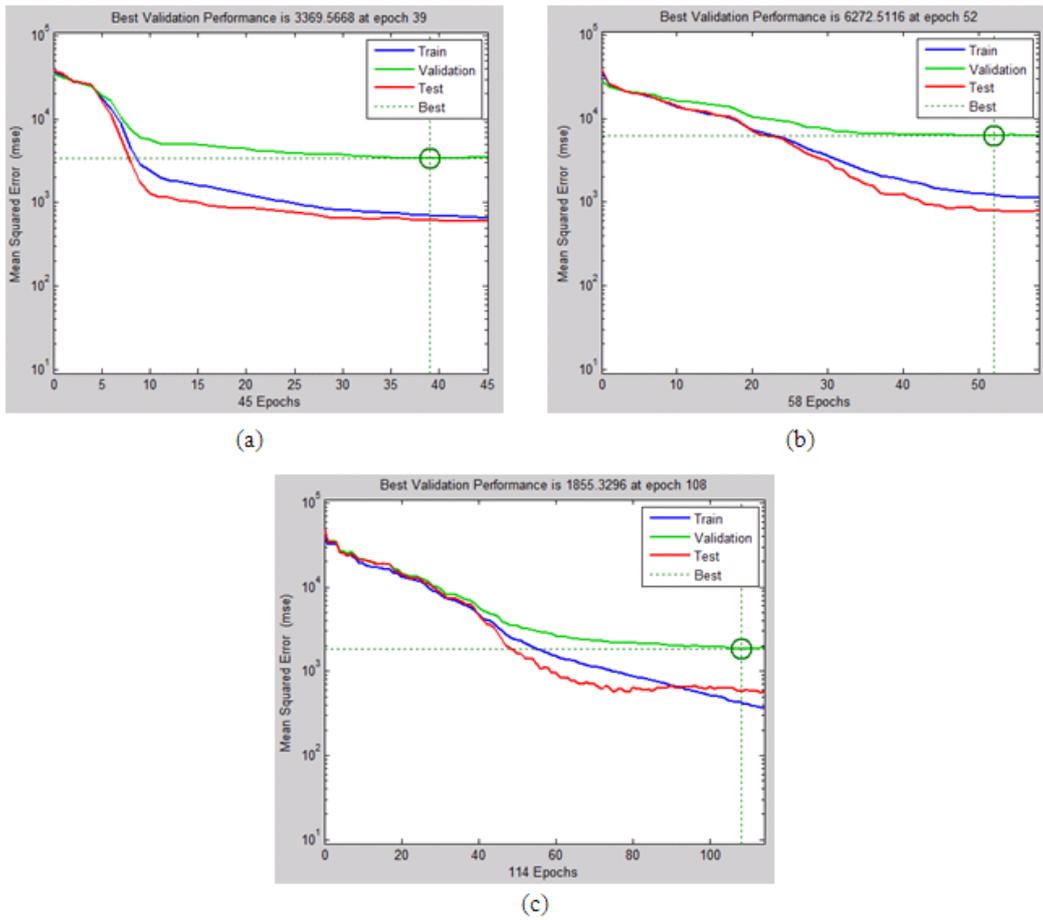


Fig. 1: Performance of different training algorithms. (a) Fletcher-Powell Conjugate Gradient (b) Polak-Ribière Conjugate Gradient (c) Scaled Conjugate Gradient

In a regression plot the perfect fit which shows the perfect correlation between the predicted and targets is indicated by the solid line. The measure on how well the produced results correlate with the target is indicated by correlation coefficient, R. The dashed line indicates the best fit produced by the algorithm. Regression plots of all algorithms are presented in Fig. 4.

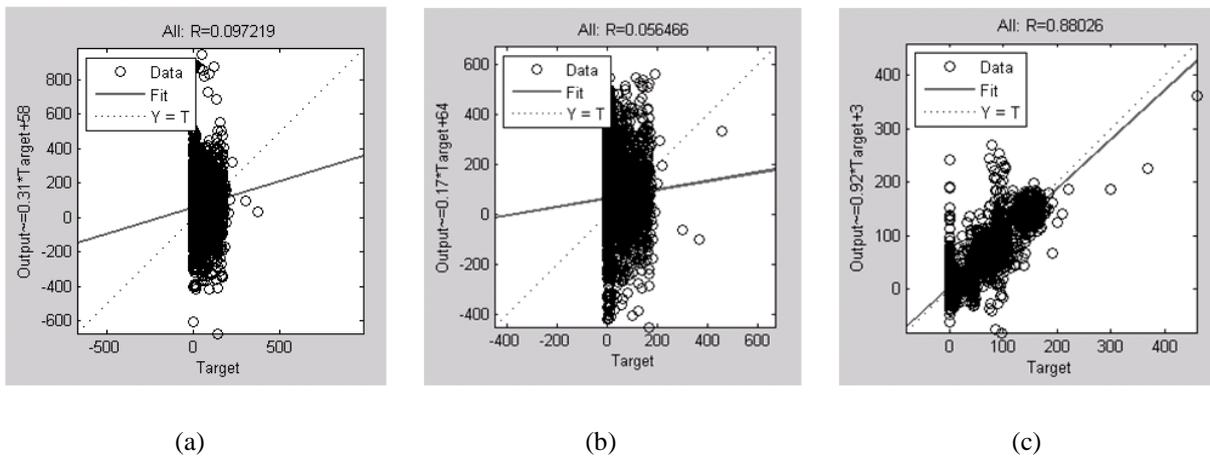


Fig. 4: Regression analysis plots. (a) Fletcher-Powell Conjugate Gradient (b) Polak-Ribière Conjugate Gradient (c) Scaled Conjugate Gradient

Based on Fig. 4 the solid line in the regression plot of Scaled Conjugate Gradient is nearly overlap with the dash line indicate a close match between the desired and predicted outputs. However the solid lines of both Fletcher-Powell Conjugate Gradient and Polak-Ribière Conjugate Gradient are far from overlapping with the dashed lines indicating poor correlation between the predicted and target outputs.

Table 1 summarizes the values of parameter based on the performance and regression plots of all algorithms.

Table 1: Network simulation parameters.

Algorithm	Epochs	Correlation Coefficient, R
Fletcher-Powell Conjugate Gradient	39	0.097219
Polak-Ribière Conjugate Gradient	52	0.056466
Scaled Conjugate Gradient	108	0.88026

The objective of this study is to determine the algorithm which provides the best performance using the collected data. The performances of the different algorithms are studied based on the regression analysis which investigates the correlation between the produced results with the target outputs. The value of R of Scaled Conjugate Gradient algorithm is the closest to 1 compared to the other algorithms indicating accurate prediction. The training and testing of the network using this algorithm is completed after producing 108 epochs.

Polak-Ribière Conjugate Gradient algorithm produces poor prediction performance by providing the smallest value of R which is equal to 0.05646. On the other hand, Fletcher-Powell Conjugate Gradient provides larger value of R of 0.097219. Training and testing stages using Polak-Ribière Conjugate Gradient and Fletcher-Powell Conjugate Gradient algorithm are completed after providing 52 and 39 epochs respectively.

4. Conclusion

Based on the analysis conducted Scaled Conjugate Gradient algorithm produces the best performance in the prediction of diabetes compared to Fletcher-Powell Conjugate Gradient and Polak-Ribière Conjugate Gradient algorithms based on the values of R from the regression analysis. This algorithm produces R of 0.88026 which indicate a good correlation between the targets and predicted outputs. Polak-Ribière Conjugate Gradient algorithm produces poor performance with the smallest value of R of 0.056466.

5. Acknowledgements

This work is completed with the cooperation provided from the Department of Endocrine, Pusat Perubatan Universiti Kebangsaan Malaysia, Kuala Lumpur by allowing the first author to collect some data to be studied. Special thanks are dedicated to all staffs involved through the completion of this work.

6. References

- [1] S. Ghosh-Dastidar, H. Adeli. A New Supervised Learning Algorithm for Multiple Spiking Neural Networks with Application in Epilepsy and Seizure Detection. *Neural Networks*. 2009, pp. 1419-1431.
- [2] A. E. Smith, C. H. Dagli. Controlling Industrial Processes through Supervised, Feedforward Neural Networks. *Computers in Engineering*. 1991, pp. 247-251.

- [3] G. S. Androulakis, G. D. Magoulas, M. N. Vrahatis. Geometry of Learning: Visualizing the Performance of Neural Network Supervised Training Methods. *Nonlinear Analysis, Theory, Methods & Applications*. 1997, pp. 4539-4544.
- [4] H. Temurtas, N. Yumusak, F. Temurtas. A Comparative Study on Diabetes Disease Diagnosis using Neural Networks. *Expert systems with Applications*. 2009, pp. 8610-8615.
- [5] K. M. Eggers, J. Ellenius, M. Dellborg, T. Groth, J. Oldgren, E. Swahn, B. Lindahl. Artificial Neural Network Algorithms for Early Diagnosis of Acute Myocardial Infarction and Prediction of Infarct Size in Chest Pain Patients. *International Journal of Cardiology*. 2007, pp. 366-374.
- [6] O. Er, N. Yumusak, F. Temurtas. Chest Diseases Diagnosis using Artificial Neural Networks. *Expert Systems with Applications*. 2010, pp. 7648-7655.
- [7] C. Wang, S. Lian. Global Convergence Properties of the Two New Dependent Fletcher-Reeves Conjugate Gradient Methods. *Applied Mathematics and Computation*. 2006, pp. 920-931.
- [8] Z. Shi, J. Shen. Convergence of the Polak-Ribière-Polyak Conjugate Gradient Method. *Nonlinear Analysis*. 2007, pp. 1428-1441.
- [9] Moller, M. F. A Scaled Conjugate Gradient Algorithm for Fast Supervised Learning. *Neural Network* .1993, pp.525-533.