ISSN 2229-5518

Comparison of Backpropagation and Resilient Backpropagation Algorithms in Non-Invasive Blood Glucose Measuring Device

Avan[1], Erfiani [2], Bagus Sartono [3]

Abstract— Artificial intelligence technology, especially Artificial Neural Network (ANN) is a method of learning that is widely used in estimating a value. Backpropagation method is an algorithm that is widely used to recognizing complex patterns. Application of Backpropagation (BP) is often too slow for practical purposes so that some modifications are made. One of the developed modifications is by changing the learning rate called resilient backpropagation algorithm (Rprop). BP and Rprop algorithms will be applied to estimating blood glucose concentrations in the non invasive device using several models. The results of the analysis show that the Rprop algorithm gives better estimation value on all tested models. Rprop generates a smaller RMSE and greater R2. In addition, Rprop is much more efficient in the learning process. Rprop requires fewer iterations and faster learning process to assume the estimation value. In the back propagation algorithm, there is a heteroscedasticity violation on the variance errors so that the variance error is not homogeneous.

Index Terms— Artificial Neurol Network, Backpropagation, Feed-Forward Nural Network, Glukosa, Learning Rate, Non-Invasive, Resilient Backpropagation,

1 Introduction

Artificial neural network (ANN) is a computational system which architecture and operations inspired from the knowledge of biological neurons in the human brain. ANN can be trained to learn and analyze patterns of input data. ANN tries to find a formula or function that will connect it with the desired output. The accuracy of ANN prediction is measured by R2 and RMSE [1]. One of the most widely used ANN algorithms is back propagation (BP). BP method is a supervised learning technique that is widely used in many layers to recognize complex patterns. Besides the advantages, BP has a weakness of taking a long time in the learning process. A lot of researches conducted to fix the weakness, one of them is resilient backpropagation algorithm (Rprop). At Rprop, learning rate parameter will always change according to the condition of the error of each iteration [4].

This research tried to model the equations of spectral calibration of voltage value excreted by Non-Invasive glucose measuring device to predict blood glucose using a resilient backpropagation algorithm and compare it with a standard backpropagation algorithm.

2 Background

2.1 Diabetes

Diabetes Mellitus is a disease that causes high levels of glucose in the body because the body can not produce or lack of the insulin hormone [10]. Diabetics need routine monitoring for glucose control. However, routine measurements are rarely done because the measurement is still by an invasive method which causing pain and highly cost of measuring. It is needed a device that can quickly detect the glucose levels without injuring the body at a more affordable cost. Therefore, a non-invasive glucose measuring device is being developed. An alternative method that can express the relationship between the spectrum output of non-invasive measuring device and the blood glucose level is the Artificial Neural Network.

Ozgur and Erdal (2005) compared 3 backpropagation development algorithms namely Levenberg-Marquardt, conjugate gradient and resilient back-propagation [8]. Iftikhar et al (2008) implemented a backpropagation algorithm with Resilient Backpropagation to detect interference on a computer [3]. And Navneel et al (2013) also compared resilient backpropagation and backpropagation algorithms to classify spam emails [6].

2.3 Backpropagation (BP)

Backpropagation is one of the weighting algorithms. BP is commonly used by perceptron with many layers to change the weights associated with existing neurons in the hidden layer. The BP algorithm uses output errors to change the value of the weights in the backward propagation. The error is obtained from the difference between the target value and the value of the learning output. Learning is done by using gradient descent. The BP error of the output is pushed back through the network to estimate the error in the hidden layer [7]. In forward propagation the neurons will be activated by using a differentiated activation function..

2.2 Artificial Neural Network (ANN)

Artificial Neural Network (ANN) is an information processing system that adopts the workings of biological neural networks in the brain [2]. ANN is not programmed to produce a particular output. All outputs generated by the network are based on experience during the learning process. In the ANN learning

- *Avan is currently pursuing masters degree program in statistics in Bogor Agriculture University, Indonesia, PH +6281231117500. E-mail: avandjauhari@gmail.com*
- *Erfiani is Lecturer, Departement of Statistics, Bogor Agriculture University, Bogor, Indonesia. E-mail: erfiani_ipb@yahoo.com*
- *Bagus Sartono is Lecturer, Departement of Statistics, Bogor Agriculture University, Bogor, Indonesia. E-mail: bagusco@gmail.com*

ISSN 2229-5518

process, the input and output patterns are inputted and the network will be taught to provide an acceptable answer. In the ANN, the input will be processed by neurons with certain weights. In general, the working is to process the received signal then distributed across the network and stored as a weight in each neuron. During the training process, the process of adjusting the weights and the limit values obtained the desired output.

2.4 Resilient Backpropagation (Rprop)

Backpropagation is an excellent method and is widely used for recognizing the complex patterns. Besides the advantages, BP has a weakness of taking a long time in the learning process. Resilient Backpropagation or called Rprop is one of the modifications in backpropagation to accelerate learning rate. Rprop performs a direct adaptation of the weighted value based on information from its local gradient [9]

The weight changes in BP are influenced by the learning rate. The smaller the learning rate, the longer the learning process. While the greater the learning rate, the weights will be far from the minimum weight. The Rprop algorithm is the result of developed backpropagation algorithms to overcome these weaknesses. BP has a single learning rate determining size with adjustment value (Δ_0)[9]. On the other hand, Rprop has 2 learning rate namely the increase factor (η^+) and decrease factor (η^-). The Rprop algorithm uses positive or negative marks from gradient to indicate the direction of weight adjustment. The value of the learning rate that is often used is 1.2 for the value of η + and 0.5 for the value of η- the adjustment value rule [9].

$$\Delta_{ij}{}^{(t)} \begin{cases} \eta^{+}\Delta_{ij}{}^{(t)}\,, if\ \dfrac{\partial E^{t}}{\partial W_{ij}}\dfrac{\partial E^{t-1}}{\partial W_{ij}} > 0 \\ \eta^{-}\Delta_{ij}{}^{(t)}\,, if\ \dfrac{\partial E^{t}}{\partial W_{ij}}\dfrac{\partial E^{t-1}}{\partial W_{ij}} < 0 \\ \Delta_{ij}, \qquad els \end{cases}$$

3 METHOD

3.1 Data

This study used primary data which are part of research of development and clinical trial prototype of non-invasive glucose monitoring tool. The data used were the output voltage values of non-invasive glucose measuring device and invasive blood glucose data. This study used 120 respondents who are students from several departments in IPB. Measurements and blood sampling were conducted at the Biochemical Laboratory of the Community Nutrition Department of the Bogor Agriculture University.

In the first stage of the study, respondents' weight and height were measured. In the second stage, blood glucose measurements were performed on the respondents using non-invasive measuring instruments. In the final stages of research, respondents' blood glucose was measured by invasive device. Invasive measurements were made by taking respondents' blood samples. The blood is taken to Prodia's laboratory and was measured the glucose levels through a chemical process.

3.2 TOOLS AND FRAME WORK

Non-invasive blood glucose measuring devices are coupled with sensors and infrared lamps with a wavelength of 1600 nm. The devices are combined to infrared lights with controlled time periods and light intensities. The lamp has been set to have 10 periods. Each time period is 500 ms with continuously added intensity of light. Each period produces many different transmittance outputs. The process of blood glucose measuring is conducted by firing infrared light to the ring finger for 5 seconds. The light fired will partially be retained by the compounds in the body, one of them is glucose and some will be missed. The intensity of light that is passed will be captured by the sensor in the form of a continuous voltage of the spectrum. The continuous voltage is transformed by the ADC into a discrete output value.

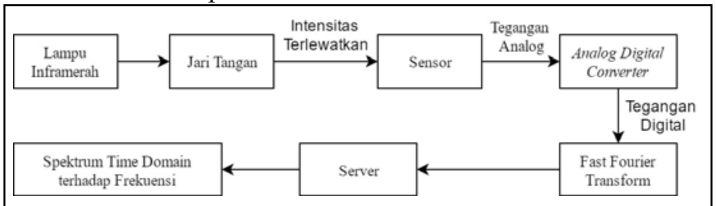

Figure 1 Flow chart of frame work of non-invasive blood glucose measuring device

3.3 DATA ANALYSIS

Stages of analysis in this study are:

1. Preparation of the data
   a. Transformation of transmittance value in each period into a summary of values that describes the average of passed light intensity. The value will be the input value (X)

$$\bar{x}_i = \sum_{j=1}^{n} x_{ij} \qquad i=1,..10;\ j=1\ldots n$$

   b. Enter the value of the invasive measurement result which then becomes the target value (Y)
2. Plot spectrum output of non-invasive blood glucose measuring device for low and high blood glucose levels
3. Arificial Neural Network (ANN)
   a. Transform the data with scale [0,1]
   b. Divide the data into two parts i.e 80% of training data and 20% testing data.
   c. Determine the activation function to be used
   d. Determine the number of hidden layers
   e. Determine the number of nodes in each hidden layer
   f. Analyze ANN with BP and Rprop algorithm by testing several models in training data
   g. Observe the plot of standardized conjectural values ($\hat{Y}$) with standardized errors (ε) in training data
   h. Observe the plot of the estimation value ($\hat{Y}$) of the ANN with the target value (Y) of each model in the training data and the testing data
   i. Select 4 models with criterias 2 of the initial model and 2 modes of the best development results by looking at the plot point g and h

ISSN 2229-5518

4. Analyze ANN with BP and Rprop algorithm on 4 selected models with 50 replications.
5. Evaluation of the model by comparing the mean values of RMSE and R^2 from the 4 models in the both best algorithm evaluation by comparing the level of accuracy and effectiveness of the three algorithms of the RMSE, R^2 and number of iterations needed in the learning process

6 Discussion

6.1 Exploration

Figure 2 is an example of a non-invasive blood glucose measurement result with 10 periods for blood glucose values of Y = 67 mg/dl and Y = 123 mg/dl. Each period has many different transmittance values. The value is summarized to the average transmitter value per period (X) that is used as input value on ANN.

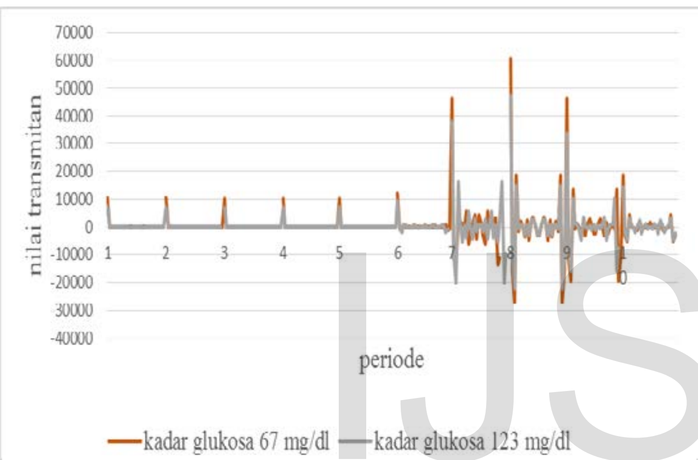

Figure 2 Example of spectrum image of non-invasive blood glucose measuring device at 67 mg/dl and 123 mg/dl of blood glucose

Figure 2 shows the difference in the spectrum image between low blood glucose level Y = 67 mg/dl with high blood glucose level Y = 123 mg/dl. The spectrum with blood glucose levels Y = 67 mg/dl image, the graph has higher peaks than Y = 123 mg/dl. The conclusion of figure 2 is when the infrared ray of non-invasive blood glucose is fired on the respondent's fingers with low blood glucose levels, the intensity of the light passed is greater than the respondent with high blood glucose levels. The graph shows that there is a spectral difference between high blood glucose levels and low blood glucose levels. So that the spectrum of non-invasive blood glucose measurements can be modeled.

Figure 4 shows the presence of outlier of Y = 276 mg/dl. This outlier value will change the entire pattern so that the observed value Y = 276 mg/dl will be eliminated in the further analysis. After the outlier value is removed, the plot of variable X and Y looks to have a quadratic relationship so that in this study will be tried several models to look for the best model.

6.2 Arificial Neural Network (ANN)

The preparation of ANN starts by dividing data into 80% of training data and 20% of data testing. The activation function used is linear function y = x and treshold = 0.01. Learning rate = 0.01 for BP. At Rprop of positive learning rate = 1.2, negative learning rate = 0.5. This research will try some models. The model is written in the following order: (input variable, target value, number of hidden layers, number of nodes in each hidden layer).

Table 1 Type of algorithm and model used in ANN

Algorithm	Model
[1] .*Backpropagation* (BP)	a. (X, Y, 1, 10)
[2] *Resilient Backpropagation* (*Rprop*)	b. (X, Y, 2, (10 10))
	c. (X^2, ln Y, 1, 10)
	d. (X^2, $\sqrt{Y}$, 1, 10)

The selection of the best model is determined by looking at homoscedasticity of the variance error. The Variance of error from one observation to another is expected to be homogeneous. Homoscedasticity of variance error can be seen through plot between the estimation value ($\hat{Y}$) which is the output of ANN learning outcomes with the target value (Y) result of invasive measurement. A good model is when the plot between Y and the dotage follows the linear line with the equation y = x so that the variance error between one observation and another tends to be homogeneous.

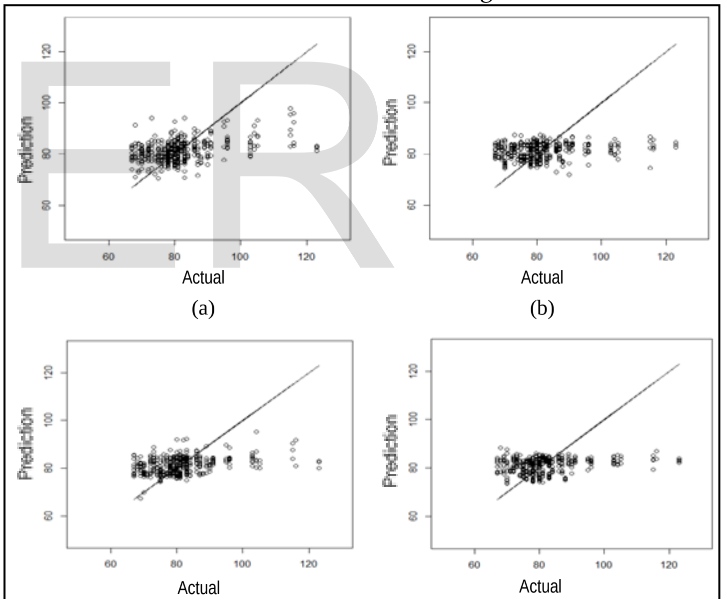

Figure 3 Plot of *invasive* blood glucose measurement result (Y) with estimation value of ANN ($\hat{Y}$) on *training* data. (a) model a of BP algorithm, (b) model a of Rprop algorithm, (c) model b of BP algorithm, and (d) model b of Rprop algorithm.

Figure 3 plot Y and of training data for models a and b show the variance error between one observation and other observations is not homogeneous for BP and Rprop algorithms. Figure 3 shows the greater the initial value (Y) to be estimated the far the estimated value ($\hat{Y}$) from the initial value. The greater the value of Y the greater the error value. This suggests that the variance error between one observation and another is not homogeneous. The variance error tends to increase in the greater Y value. In model a and model b for BP and Rprop algorithm, there has been heteroscedasticity violation.

The further analysis developed other models for improving heteroscedasticity in variance errors. After trying many models then chosen 2 best models. The models are c (X2, Y, 1, 10) and model d X2, 1, 10). Figure 4 shows the model c and d for the algorithm Rprop of dots following the linear line y = x so that the error range between one observation with the other is tended to be constant. Heteroscedasticity in the variance errors does not occur. While the algorithm BP dot plot do not follow linear lines so that the models still occur heteroskedasticity in variance error.

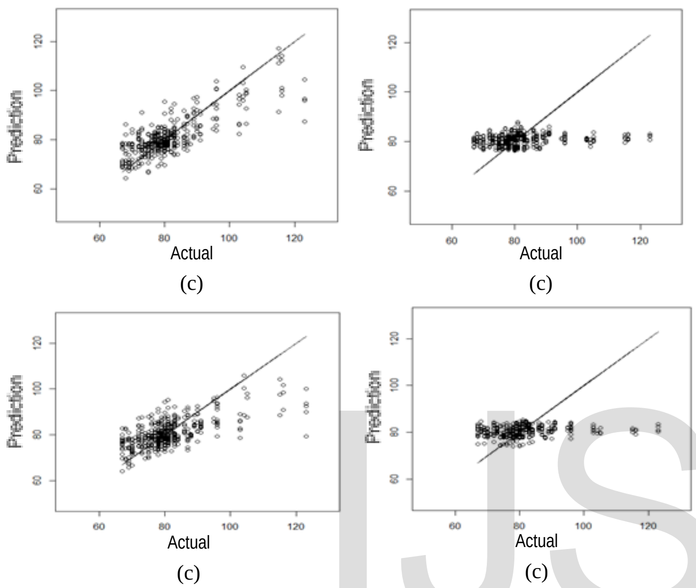

Figure 4 Plot of *invasive* blood glucose measurement result (Y) with estimation value of ANN ($\hat{Y}$) on *training* data. (a) model c of BP algorithm, (b) model c of Rprop algorithm, (d) model d of BP algorithm, and (d) model d of Rprop algorithm.

6.3 Evaluation of Results

ANN performed on 50 types of training data and data testing on 4 models with BP and Rprop algorithms. Each model was tested with 50 replicates with different training and testing data. In ANN result the value of RMSE, R^2, number of iteration and duration of learning process was observed.

Comparison of Rmse Value between Models and Algorithms

The RMSE expresses the magnitude of the error generated by an estimation model of its original value. Figure 5 The average plot of RMSE data testing shows that the Rprop model c and d algorithms provided better results with smaller RMSE values of 9.58, 9.54, compared to a and b models of 9.55, 9.71. As for the 4th BP, algorithm model produces relatively equal value RMSE 9.60, 9.66, 9.63, and 9.72 for BP. The best algorithm determination in generating the minimum RMSE can be seen in Figure 5. The Rprop algorithm is better than BP because it produces a smaller RMSE for model c and d.

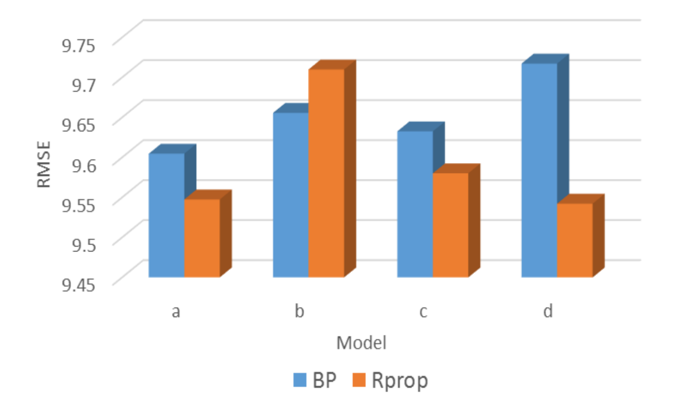

Figure 5 Plot of RMSE average on the *testing* data

R2 shows the ability of independent variables in explaining the variance of the response variables. Figure 6 The plot of average R2 data testing shows the algorithm Rprop model c and d has R2 value greater than model a and b that are 37.13%, 24.69% in Rprop. As for model a and b the value of R2 is only 9.64%, 7.56%. In the BP model development algorithm has not succeeded in increasing the R2 value because the 4 models have a relatively small R2 value of 10.76%, 8.07%, 8.16%, and 6.21%. The Rprop algorithm is better than BP according to the R2 results. R2 for the Rprop algorithm is bigger than BP.

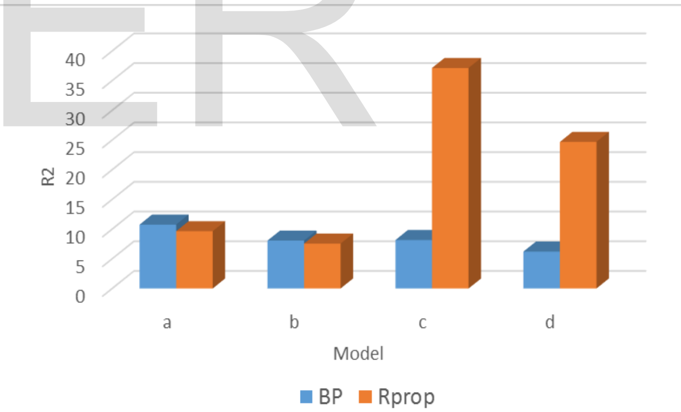

Figure 6 Plot of R^2 average on the *testing* data

Comparison of Number of Iterations between Algorithms

Figure 7 shows the average number of iterations required at 50 replications in the learning process of the two algorithms. Rprop is much more efficient than BP. The Rprop algorithm requires fewer iterations in the learning process i.e 58, 30, 1158, 439.96 for models a, b, c, and d. BP requires iterations of 2058, 1181, 2103, 2046.

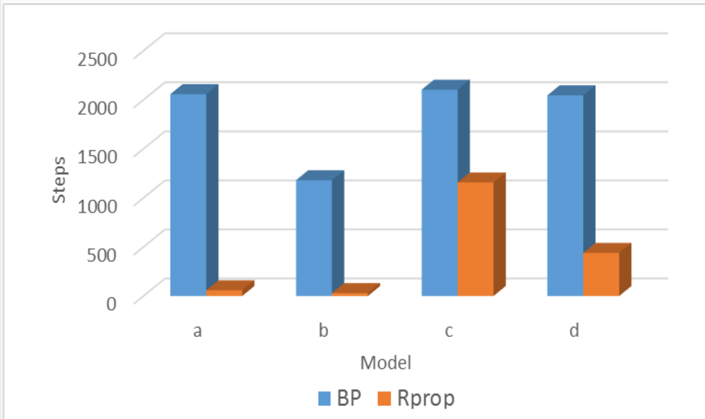

Figure 7 Plot of average number of iterations required

Comparison of Duration of Learning Process between Algorithms

Figure 8 shows the average duration of the learning process in ANN for both algorithms from 50 replications. According to the time of learning process Rprop much more efficient than BP. The Rprop algorithm requires a faster learning time of 0.12, 0.12, 1.75, 0.68 seconds for models a, b, c, d. While backpropagation takes longer learning process time i.e 3.31, 3.43, 2.9, 2.84 for model a, b, c, d.

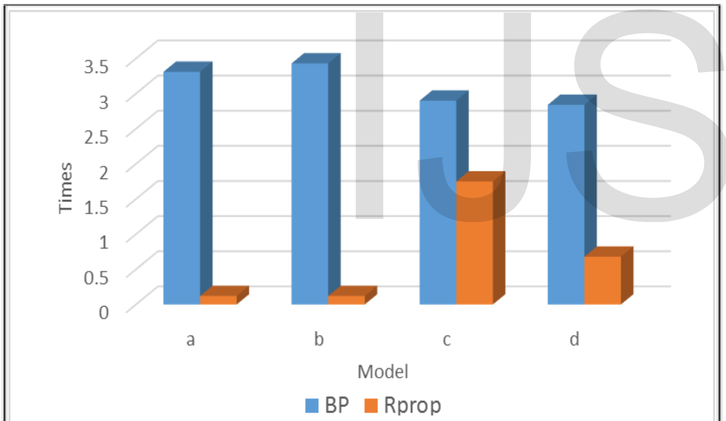

Figure 8 Plot of average duration of the learning process

7 Conclusion

7.1 Appendices

Estimation of blood glucose level with ANN using developed model in model c (X2, Y1, 1, 10), and model d e (X^2, $\sqrt{Y}$, 1, 10) yield better estimation value. Improvement models can avoid the heteroscedasticity of error violations in model a and b. These models are also able to increase the value of R^2 and minimize RMSE. However, this only works on the backpropagation resilient algorithm. While on the backpropagation, the development model does not provide significant change. Heteroscedasticity of variance errors still occur in the BP algorithm models.

The resilient backpropagation algorithm provides better estimation values than the backpropagation algorithm. The advantages of Rprop not only can be seen from the RMSE and R^2 generated, but also from the efficiency of a learning process. Rprop requires less iteration and faster learning duration in the process of estimation

7.3 Suggestions

Suggestions for further research are to use simulation data in comparing the accuracy and efficiency of multilayer perception, backpropagation, and resilient backpropagation algorithms. Try to use methods other than ANN such as bayesian models in modeling blood glucose levels using a non-invasive measuring device because it is assumed that ANN is not the best model in the preparation of calibration model.

References

[1] Cowan MK, *Machine Learning,* California (US): Stanford University, 2013.

[2] Fausett, L. 1994. *Fundamentals of Neural Networks*. New Jersey: Prentice Hall Inc

[3] Iftikhar A, M.A. Ansari and Sajjad M. "Performance Comparison between Backpropagation. Algorithms Applied to Intrusion Detection in Computer Network Systems", *Proc. ninth WSEAS International Conference on NEURAL NETWORKS (NN'08)*, pp. 231-236, may 2008.

[4] Kriesel D. 2005. *A Brief Introduction to Neural Network*, available at http://www.dkriesel.com

[5] Moreira M dan Fiesler E. *Neural Network with Adaptive Learning Rate AND Momentum Terms*, SUISSE :Institut Dalle Molle, 1995.

[6] Navneel P, Rajeshni S. and Sunil P, "Comparison of Backpropagation and Resilient Backpropagation Algorithm for Spam Classification", *Proc. fifth international conference on computational Intellegence, Modelling and Simulation* pp. 29-34, Sept 2013.

[7] Normakristagaluh P, "Penerapan jaringan syaraf tiruan untuk peramalan curah hujan dalam statistical downscaling" , thesis, Bogor Agricultural university, 2004.

[8] Ozgur K and Erdal U, "Comparison of three backpropagation training algorithms for two case studies". *J. of Engineering and material sciences*. vol 12, pp. 434-4442, Oct 2005.

[9] Riedmiller M dan Braun H, "RPROP A fast adaptive learning algorithm". *International Symposium on Computer and Information Sciences.* 7 pp 279-285, 1992.

[10] Setiabudi. 2008. *Referensi kesehatan Diabetes Melitus* . available *at* http://creasoft.wordpress.com/2008/04/15/diabetes-melitus