Analysis Resilient Algorithm on Artificial Neural Network Backpropagation

Widodo Saputra1, 3, Tulus1, Muhammad Zarlis1, Rahmat Widia Sembiring2, Dedy Hartama1, 4

1Univesitas Sumatera Utara
2Politeknik Negeri Medan
3AMIK Tunas Bangsa
4STIKOM Tunas Bangsa

Abstract. Prediction required by decision makers to anticipate future planning. Artificial Neural Network (ANN) Backpropagation is one of method. This method however still has weakness, for long training time. This is a reason to improve a method to accelerate the training. One of Artificial Neural Network (ANN) Backpropagation method is a resilient method. Resilient method of changing weights and bias network with direct adaptation process of weighting based on local gradient information from every learning iteration. Predicting data result of Istanbul Stock Exchange training getting better. Mean Square Error (MSE) value is getting smaller and increasing accuracy.

1. Introduction

Previous Artificial Neural Network (ANN) backpropagation researches have some weaknesses, like as local minimum problem. Speed of convergence Backpropagation algorithm too slow, dependent on initial parameters, such as the number of hidden nodes, input, output, the learning rate and the weights of the connections in the network (Park, et.al, 2009). Another issue is the training time too long, especially if used parameters are less precise, for example that the selection of learning rate effect on changes in weights. Thus need algorithms to speed up the training. Many techniques applied to increase the speed of convergence in the process of training, such as optimization of gradient descent techniques (Kusumadewi & Sri, 2004).

Resilient method can be used to optimize accuracy. Resilient method is local adaptive learning algorithm, when compared with the backpropagation, its faster convergence and needs of training tends to be less. Compared with the backpropagation, resilient can provide faster of training and the rate of convergence and has the ability to stay away from the local minimum. Resilient is very strong with respect in internal parameters and its considered as one of the best learning method in ANN (Sheng, 2011).

2. Literature Review

2.1. Artificial Neural Network

ANN is a system of information processing that has certain performance characteristics as the neural network of the human brain, by doing the learning process in changes the weights. ANN is able to recognize activities with past data. The data of the past will be studied by ANN so it has the ability to give a decision of data that hasn't been studied (Hermawan, 2006).

2.2. Backpropagation

Backpropagation is a type of ANN used in solving problems of forecasting. This is possible because it is one type of training methods ANN with supervision. On the network was given a pair of patterns composed of input patterns and the desired patterns. When a pattern is given to networks, weights are modified to minimize the difference in patterns of output and the desired patterns. This exercise is done
over and over again so that the entire patterns issued network can satisfy the desired pattern. This network architecture consists of the input layer, hidden layer, and output layer. Its architecture can be seen in Figure 2.1 (Fausett, 1994).

![Figure 2.1. Backpropagation Architecture Network with a single Hidden Layer](image)

Any some algorithms that can be used to make backpropagation, but was chosen because it may be the easiest to implement, while keeping the network efficiency. It’s using more than one input layer (usually 3). Each of these layers:

- **Input Layer** - This layer store the input to the network.
- **Output Layer** - This layer store the output data, usually an input Identifier.
- **Hidden Layer** - This layer is between the input and the output layer. They as a point of backpropagation to send data from the previous layer to the next layer

### 2.3. Resilient Backpropagation (Rprop) Algorithm

Rprop algorithm is the result of backpropagation. Change weights on backpropagation are affected of the learning rate and depending on the slope of the error curve. The smaller the learning rate, learning the longer. While the greater the learning rate, the weighting values will be far from the minimum weights. To overcome this, developed a new algorithm is Rprop. This algorithm uses the sign (positive or negative) of the gradient to show the direction of the adjustment weight. While size of the change the weights is determined by the value adjustment (0Δ) (Riedmiller, et.al, 1992).

The algorithm is Resilient to change weights and bias network with direct adaptation process of weighting based on local gradient information from iteration of learning, so that the number of iterations needed to reach the target (Sulistijanti, 2013).

For it, given the size of the change of weight is improvement value \( \Delta jk \). The value of the repair adjusting the development during the learning process based on local vision towards the function of error with learning the rules as follows:

\[
\Delta jk(m) = \Delta jk (m-1) \times \eta^+, \text{ if } \frac{\partial E}{\partial w_{jk}} (m) \times \frac{\partial E}{\partial w_{jk}} (m-1) > 0 \\
\Delta jk(m) = \Delta jk (m-1) \times \eta^- , \text{ if } \frac{\partial E}{\partial w_{jk}} (m) \times \frac{\partial E}{\partial w_{jk}} (m-1) > 0 \\
\Delta jk(m) = \Delta jk (m-1) , \text{ for more } \eta^-
\]

Where \( 0 < \eta^- < 1 < \eta^+ \)

The rules of adaptation works is every partial derivative of weights and bias on two different successive iterations in order to show that the improvements of the last value that is too large and the algorithm has to
leap over a local minimum, the value of the repair $\Delta_jk$ lowered by a factor of $\eta^-$. If in two iterations sequential derivative marks constant, the value of the repairs is raised with the factor $\eta^+$ to speed up convergence in the area of surface errors. And if its derivative is zero, then the value of the update stays the same. Each time the repair, oscillating weights reduced. If changing weights in the same direction for a few iterations, then the resulting increases in the magnitude of the change in weight (Sulistijanti, 2013).

After repair value for each of the weights are adjusted, improved weights follow the rules that simple if the derivative is positive (increasing errors) then the weighted down with the value of the value of the repair, if the derivative is negative then the value of the repairs is added as follows:

$$\Delta w_{jk}(m) = -\Delta_jk (m) \text{ jika } \frac{\partial E}{\partial w_{jk}}(m) > 0$$

$$\Delta w_{jk}(m) = +\Delta_jk (m) \text{ jika } \frac{\partial E}{\partial w_{jk}}(m) < 0$$

$$\Delta w_{jk}(m) = 0 \text{ For more}$$

$$W_{jk}(m+1) = w_{jk} + \Delta w_{jk}(m)$$

$W_{jk}(m)$ which is the weighting between $j$ and $k$ neurons in two successive layers on the iteration $m$, $W_{jk} (m + 1)$ is the new weight.

3. Methods
3.1. Data

Required data from imkb.gov.tr and finance.yahoo.com. The data is compiled based on Labor Day in Istanbul's stock exchange. (http://archive.ics.uci.edu/ml/datasets/ISTANBUL+STOCK+EXCHANGE).

3.2. Data Retrieval

This study, created a pattern recognition and prediction of Istanbul Stock Exchange. This process has two stages where the first is to do pattern recognition by way of finding the best architecture of the ANN models. The process of training and testing data to get the best models from the Istanbul Stock Exchange data in working is 60 Patterns from January 5, 2009 – April 4, 2009. Criteria of pattern 1 to Pattern 30 is training data, while 31 Patterns up to 60 Patterns is test data. The second is to do predictions with the best architectural pattern obtained in the first. The process of Testing had done entering in Istanbul Stock Exchange data and to comparing the minimum error value from the best architectural patterns that are done on the first.

3.3. Training Network

Training in ANN needed input data and target data. Training covers the process of iterative input data are entered into the network so that the network can learn and adjust data that trained with the desired target data. Training is done to look for the values of the weights that connect all the neurons so that minimize the error output that is produced by the network.

Training using much as 50% are the amount of data that is comprised of data input and output targets that have been normalized. In this process will be initiated with the architecture of the ANN from a number of hidden neurons of different. Each of the tested architecture that will result in weight training that will be used as the initial weight in the process of testing. Then initialize the weights and biases to calculate the value of the output of each neuron will be multiplied by the function activation and learning rate.
After the value of the output of the network on the output layer is obtained, calculate the value of a network error, an error value is then compared to the value of the error the target that has been set. If a network error that is generated is not less than or equal to the value of the error that has been established, then the process will be done by modifying the backpropagation weighting network and a bias on a particular iteration until the minimum error value was obtained by approaching a predetermined target error before. When the error condition is less than the weighting then the target error will be stored as weights in training process.

3.4. Testing Network

Testing used as much as 50% ANN from amount of data that has been selected for each category. At this stage of the network will be in testing with the new data that hasn't been trained into a network to find out the capabilities of the network doing the generalizing the case at hand and then draw the inclination towards a specific output.

The results of the testing process in the form of a matrix which forms in compliance with output targets. The output of the network will then be determined on a particular pattern. If the output of the network is greater than or equal to 0.5 the network considered have of the output 1 and if the network has the output of less than 0.5 then it shall be deemed to have the output 0. The data is said to be known is when data normalization results are used as input values can produce the same output value chain with the desired target value.

4. Results and Discussion

At this point would do the testing system. The system created using MATLAB R2011b. System of Testing for the final review of the analysis and implementation.

| Karakteristik         | Spesifikasi         |
|-----------------------|---------------------|
| Arsitektur            | 1 hidden layer      |
| Neuron input          | 12                  |
| Neuron Hidden         | 10                  |
| Fungsi Aktivasi       | Sigmoid Biner       |
| Inisialisasi bobot    | Random              |
| Target Error          | 0.01                |
| Maksimum Epoch        | 10000               |
| Learning Rate         | 0.01                |

4.1. Results

Backpropagation algorithm is an algorithm that will produce the best value through potentially experiencing unplanned test trial and error. So this research, conducted a few experiments to increase the speed of learning (accelerated learning).

4.1.1. Training of Standard Backpropagation algorithm with Standard Backpropagation and Resilient Algorithm
In the process of training will be seen from Epoch and limits of error. In recent times an experiment with varying architecture that can be seen from table 4.2:

Table. 4.2. Training of Standard Backpropagation algorithm with Standard Backpropagation and Resilient Algorithm

| Experiment   | Architecture | Epoch   | MSE       | Accuracy (%) |
|--------------|--------------|---------|-----------|--------------|
| Backpropagation Standard | 1 12-5-1   | 37 Iteration | 0.009926  | 6.7          |
|               | 2 12-8-1   | 33 Iteration | 0.009566  | 43.3         |
|               | 3 12-10-1  | 32 Iteration | 0.009923  | 30.0         |
|               | 4 12-14-1  | 24 Iteration | 0.009894  | 33.3         |
|               | 5 12-20-2  | 36 Iteration | 0.001356  | 56.7         |

Backpropagation Standard and Resilient | 1 12-5-1 | 6 Iteration | 0.009260 | 20.0 |
|                                        | 2 12-8-1 | 3 Iteration | 0.002731 | 56.7 |
|                                        | 3 12-10-1 | 3 Iteration | 0.002731 | 56.7 |
|                                        | 4 12-14-1 | 6 Iteration | 0.002223 | 63.3 |
|                                        | 5 12-20-2 | 5 Iteration | 0.004229 | 66.7 |

4.1.2. Testing of Standard Backpropagation algorithm with Standard Backpropagation and Resilient Algorithm

In the process of testing will be seen from Epoch and limits of error. In recent times an experiment with varying architecture that can be seen from table 4.3:

Table 4.3. Testing of Standard Backpropagation algorithm with Standard Backpropagation and Resilient Algorithm

| Experiment   | Architecture | Epoch   | MSE       | Accuracy (%) |
|--------------|--------------|---------|-----------|--------------|
| Backpropagation Standard | 1 12-5-1   | 37 Iteration | 0.006238  | 10.0         |
|               | 2 12-8-1   | 33 Iteration | 0.000538  | 86.7         |
|               | 3 12-10-1  | 32 Iteration | 0.002255  | 60.0         |
|               | 4 12-14-1  | 24 Iteration | 0.031164  | 33.3         |
|               | 5 12-20-2  | 36 Iteration | 0.010406  | 56.7         |

Backpropagation Standard and Resilient | 1 12-5-1 | 6 Iteration | 0.008060 | 13.3 |
|                                        | 2 12-8-1 | 3 Iteration | 0.000771 | 90.0 |
|                                        | 3 12-10-1 | 3 Iteration | 0.000771 | 90.0 |
|                                        | 4 12-14-1 | 6 Iteration | 0.014335 | 63.3 |
|                                        | 5 12-20-2 | 5 Iteration | 0.000553 | 86.7 |

From the results of testing conducted as many as 5 experiments on backpropagation with varying architecture is 33 iteration. Whereas with same architecture of Backpropagation with Resilient Backpropagation (RPROP) is 5 iterations. Accuracy of training conducted the same experiment, an increase in the accuracy of training in experiment 1 with 12-5-1 network architecture. The same thing happens also on a second experiment with network architecture 12-10-1 to trial to 5 results accuracy Standard Backpropagation reaches 93% and results accuracy Standard Backpropagation Optimizations Resilient Backpropagation (RPROP) get the same results.
5. Conclusion

Neural Network algorithm gives satisfactory results in terms of predictions. That is as follows:

1. Changes to the parameters of the learning rate and the epoch affecting the changes up to the level of accuracy, so the selection of the parameters greatly influence on the results.
2. Artificial Neural networks are able to do predictions against data time series which is quite accurate. Looks at the cases taken, namely the Istanbul Stock Exchange.
3. By looking at the test results can be drawn a conclusion that is testing the resulting Backpropagation experience increased after optimized Resilient Backpropagation (RPROP).

Bibliography

Apriliyah, Mahmudy, WF and Widodo, AW 2008. Perkiraan penjualan beban listrik menggunakan jaringan syaraf tiruan resilient backpropagation (RProp), Kursor, vol. 4, no. 2, pp. 41-47.

Chen, C, and Lin, J, Applying Rprop Neural Network for the Prediction of the Mobile Station Location, www.mdpi.com/journal/sensors Communication 2011

Fausett, L, 1994. Fundamentals of neural networks: architectures, algorithms, and applications. Pretince Hall. New York.

Hermawan, A. 2006. Jaringan Saraf Tiruan, Teori, dan Aplikasi. Penerbit Andi. Yogyakarta.

Kim, J.H., Park, S.J., Kim, K.T., dan Hwang, S.C. 2003. Stock price prediction using Backpropagation neural network in KOSPI. Proceedings of the International Conference on Artificial Intelligence IC-AI 2003. Las Vegas, United Stated.

Kusumadewi, S. 2004. Membangun Jaringan Saraf Tiruan menggunakan Matlab & Excel Link, Graha Ilmu, Yogyakarta.

Muzakkir, I., Syukur, A., Dewi. I.K. 2014. Peningkatan Akurasi Algoritma Backpropagation Dengan Seleksi Fitur Particle Swarm Optimizational dalam Prediksi Pelanggan Telekomunikasi Yang Hilang. Jurnal Pseudocode, Volume 1 Nomor 1.

Park, T. S., Lee, J. H., & Choi, B., 2009, Optimization for Artificial Neural Network with Adaptive inertial weight of particle swarm optimization. Cognitive Informatics, IEEE International Conference, 481-485.

Prasetyo, H. R., Hadiani, R., Setiono. 2014. Analisis Data Runtun Waktu Debit Menggunakan Jaringan Saraf Tiruan Di Das Wuryantoropada AwlrKecamatan Wuryantoropada. e-Jurnal Matriks Teknik Sipil Vol. 2 No. 2 : 64.

Riedmiller, M. and Braun, H. 1992. RPROP. A fast adaptive learning algorithm. In Proceedings of the 1992 International Symposium on Computer and Information Sciences, Antalya, Turkey, pp.279-285.

Siang, J.J. 2009. Jaringan Saraf Tiruan dan Pemrogramannya Menggunakan MATLAB. Yogyakarta.

Sari, D.R. 2015. Aplikasi Penerapan Metode Neural Network Menggunakan Algoritma Backpropagation Untuk Mengetahui Pembelian Dan Penjualan Bahan Bakar Industri. INFO TEKNIK Volume 16 No. 1. 47-60.

Sulistijanti. W. Peramalan Curah Hujan Wilayah Semarang Barat dengan Algoritma Resilient Backpropagation, Median Volume VI Edisi 20 Mei 2013