

# Analysis of the effect number input and hidden layer variations on the addition kohonen algorithm to backpropagation method

P H Putra<sup>1,\*</sup>, M Zarlis<sup>2</sup>, H Mawengkang<sup>3</sup>

<sup>1</sup>Student, Department of Information Technology, Faculty of Computer Science and Information Technology, Universitas Sumatera Utara, Indonesia

<sup>2</sup>Department of Information Technology, Faculty of Computer Science and Information Technology, Universitas Sumatera Utara, Indonesia

<sup>3</sup>Department of Mathematics, Faculty of Mathematics and Natural Science , Universitas Sumatera Utara, Indonesia

\*purwahasanutra21@gmail.com

**Abstract.** The process of recognizing data patterns using the method of adding kohonen to back propagation is very influential on the amount of input data and the number of hidden layers. The results of the testing with the addition of information on the back propagation algorithm have better epoch results in testing using input data 8 has an epoch value that is better than testing using the number of input data 3,4,5,6,7. The test results using layer 5 hidden have epoch 15 value which is better than the hidden layer 3, 4,6,7,8.

## 1. Introduction

Artificial neural networks provide a very reliable methods to solve non-linear problems. Artificial neural networks are inspired by the human brain where neurons interconnect non-linearly. Neurons are connected to each other through a network [1].

In artificial neural networks there are several methods to accelerate learning including the back propagation method is a gradient reduction method to minimize the squared output error [2]. There are three stages that must be done in network training, namely forward propagation, back-propagation, and changes in weight and bias. This network architecture consists of input layers, hidden layers, and output layers [3].

The process of data pattern recognition using the method of adding kohonen to back propagation is very influential on the amount of input data and the number of hidden layers, therefore this study focuses on variations in the number of inputs and hidden layers in data pattern recognition.

## 2. Research Background

Wahyuni, I. et al. The result of her research showing us the most optimal parameter from backpropagation using a learning rate of 0.4, hidden layer 3, and maximum of epoch 4000. The optimal backpropagation model the error result from the average rmse to the 4 locations dropped by 8.14 from the training process and 8.28 from the testing.

Imran shafi, et al. In this research need investigation eksperiment to knowing the effect of varying the number of neurons and hidden layers in advanced propagation, and back propagation. Varying the number of neurons and hidden layers greatly effect the performance of neural networks [5].

Uday Pratap Singh, et al. In this research, an adaptive neural network like Kohonen neural network (KNN) model reference is used for tracking control of nonlinear system. Proposed

adaptive Kohonen neural network (ADKNN) are used to minimize the error between output and target signal for nonlinear discrete-time systems [6].

### 3. Proposed Method

#### 3.1. Neural Network

Artificial neural networks basically deal with the nodes and connectors. Nodes are analogous to natural neurons those receive signals through synapses situated on dendrites or membrane of the neurons which are connected through Axons as connectors. Synapses are analogous to weights of the inputs. When the strong enough signals, more than a certain threshold, are passed through axons, neurons would be activated. In the same way, other signals may be passed through other synapses, those may instigate other neurons [5].

Artificial Neural Networks (ANN) are information processing paradigms prepared by the biological nervous system, such as information processes in the human brain. The key element of this paradigm is the structure of an information processing system consisting of large interrelated elements (neurons), simultaneous work to solve certain problems.

The way ANN works is like the way humans work, that is learning through examples. The composing layers of ANN are divided into 3, namely the input layer, the hidden layer, and the output layer [1].

#### 3.2. Algorithm Kohonen

Kohonen network, as shown in is a two layer feed-forward network. Learning of KNN is hybrid, since from input to hidden layer unsupervised learning and from hidden to output layer supervised learning is used. Theorem 1 and theorem 2, shows that the basis function of Kohonen network is universal approximator. Training algorithm of KNN is given below [6]:

1. Select initial weights  $W_{ij}$  random values from input vector range and learning rate  $\lambda \in [0,1]$
2. Apply step 3-7, upto stopping criteria is false, stopping criteria may be number of iteration or learning parameter is sufficiently small (say  $\epsilon$ ).
3. Apply euclidean measure from input vector and weight vector, for  $j=1, \dots, m$

$$D(j) = \sum_{i=1}^m \sum_{j=1}^m (x_i - w_{ij})^2 \quad (1)$$

4. Calculate winning node say index  $J$ , so that  $D(j)$  is minimum.
5.  $J$  within a specified neighbourhood of  $j$  and for all  $i$ , calculate new weights as in equation Update new weight using.

$$W_{ij}(\text{new}) = (1 - \lambda)W_{ij}(\text{old}) + \lambda X_i \quad (2)$$

Update new weight  $\lambda$  using

$$\lambda(t+1) = 0.5 * \lambda(t) \quad (3)$$

- 8 Calculate error and test stopping criteria of network

#### 3.3. Algorithm Backpropagation

Backpropagation architecture is one of artificial neural network architecture that can be used to study analyze past data more precisely so that more accurate results are obtained (with minimum errors or errors) [1].

The steps in building a backpropagation algorithm are as follows [1]

- a. Initialize weights (take a small random value).
- b. Forward propagation stage

- 1) Each input unit ( $X_i, i = 1, 2, 3 \dots n$ ) receives the  $x_i$  signal and forwards the signal to all units in the hidden layer.
- 2) Each hidden unit ( $Z_j, j = 1, 2, 3 \dots p$ ) sums up the input signal weight, shown by equation (4).

$$z_{in_j} = v_{0j} + \sum_{i=1}^n x_i v_{ij} \quad (4)$$

And apply the activation function to calculate the output signal, indicated by equation

$$z_j = f(z_{in_j}) \quad (5)$$

The activation function used is the sigmoid function, then sends the signal to all output units.

- 3) Each unit of output ( $Y_k, k = 1, 2, 3 \dots m$ ) adds the weight of the input signal, indicated by equation (6).

$$y_{in_k} = w_{0k} + \sum_{i=1}^p z_i w_{ik} \quad (6)$$

and apply the activation function to calculate the output signal, shown by equation

$$y_k = f(y_{in_k}) \quad (7)$$

c. Back propagation stage

- 1) Each unit of output ( $Y_k, k = 1, 2, 3 \dots m$ ) receives a target pattern that matches the training input pattern, then calculates the error, indicated by equation (8).

$$\delta_k = (t_k - y_k) f'(y_{in_k}) \quad (8)$$

$f'$  is an instance of the activation function.

Then calculate the weight correlation, indicated by equation (9).

$$\Delta w_{jk} = \alpha \delta_k z_j \quad (9)$$

And calculating the bias correction, is shown by equation (10).

$$\Delta w_{0k} = \alpha \delta_k \quad (10)$$

At the same time send to the units in the rightmost layer.

- 2) Each hidden unit ( $Z_j, j = 1, 2, 3, \dots, p$ ) sums up the input delta (from the units in the layer on the right), indicated by equation (11).

$$\delta_{in_j} = \sum_{k=1}^m \delta_k w_{jk} \quad (11)$$

To calculate error information, multiply this value with the derivative of its activation function, indicated by equation (12).

$$\delta_j = \delta_{in_j} f'(z_{in_j}) \quad (12)$$

Then calculate the weight correction, indicated by equation (13).

$$\Delta v_{jk} = \alpha \delta_j x_i \quad (13)$$

After that, also calculate the bias correction, indicated by equation (14).

$$\Delta v_{0j} = \alpha \delta_j \quad (14)$$

d. Stage of changes in weight and bias

- 1) Every unit of output ( $Y_k, k = 1, 2, 3, \dots, m$ ) is carried out changes in weights and biases ( $j = 0, 1, 2, \dots, p$ ), indicated by equation (15).

$$w_{jk}(\text{baru}) = w_{jk}(\text{lama}) + \Delta w_{jk} \quad (15)$$

Each hidden unit ( $Z_j, j = 1, 2, 3, \dots, p$ ) changes the weight and bias ( $i = 0, 1, 2, \dots, n$ ), indicated by equation (16).

$$v_{ij}(\text{baru}) = v_{ij}(\text{lama}) + \Delta v_{ij} \quad (16)$$

#### 4. Results and Analysis

In this study, pattern recognition of temperature data will be carried out to test the data using an artificial neural network addition to the back propagation method. The data used is

Medan city rainfall data taken from 1998-2017. The author wants to find out whether there is a difference between the effect of variations in the number of inputs and hidden layers on the addition of information on the back propagation method to the results of the epoch. The training process is carried out on variant parameters of the number of input and hidden layers. To recognize patterns from temperature data that will be used as training data. To find out whether the addition of the Kohonen algorithm to the back propagation method, several tests were carried out. The first test researchers conducted training with variations in the number of inputs 3,4,5,6,7,8 and the number of hidden 3,4,5,6,7,8 with Learning Rate = 0.01, Max Epoch = 150, Target Error = 0.01.

Before processing the data is normalized first. Normalization of data is done so that the network output matches the activation function used. These data are normalized in intervals [0, 1] because the data used is positive or 0. In addition, the activation function given is sigmoid binary. The sigmoid function is an asymptotic function (never reaches 0 or 1), then the data transformation is carried out at smaller intervals namely [0.1; 0.8], indicated by equation (17).  $a$  is the minimum data,  $b$  is the maximum data,  $x$  is the data to be normalized, and  $x'$  is transformed data. After the data is normalized, training is then conducted to recognize the pattern of data provided. The training is carried out with parameters varying the number of different input and hidden layers. Initialization of weights is also done in starting the training process. The results of the tests are shown in the following table.

**Table 1.** Training Result With Input Data 3

Data Input	3					
Hidden Layer	3	4	5	6	7	8
#	Epoch					
<b>B</b>	105	114	89	90	98	93
<b>B + K</b>	36	17	18	22	17	31

B = Back propagation K = Kohonen

**Table 2.** Training Result With Input Data 4

Data Input	4					
Hidden Layer	3	4	5	6	7	8
#	Epoch					
<b>B</b>	150	150	150	150	150	150
<b>B + K</b>	150	150	150	150	150	150

B = Back propagation K = Kohonen

**Table 3.** Training Result With Input Data 5

Data Input	5					
Hidden Layer	3	4	5	6	7	8
#	Epoch					
<b>B</b>	150	150	150	150	150	150
<b>B + K</b>	150	150	150	150	150	150

B = Back propagation K = Kohonen

**Table 4.** Training Result With Input Data 6

Data Input	6					
Hidden Layer	3	4	5	6	7	8
#	Epoch					
<b>B</b>	150	150	150	150	150	150
<b>B + K</b>	23	16	27	19	18	19

B = Back propagation K = Kohonen

**Table 5.** Training Result With Input Data 7

Data Input	7					
Hidden Layer	3	4	5	6	7	8
#	Epoch					
<b>B</b>	150	150	150	150	150	150
<b>B + K</b>	27	22	23	19	19	24

B = Back propagation K = Kohonen

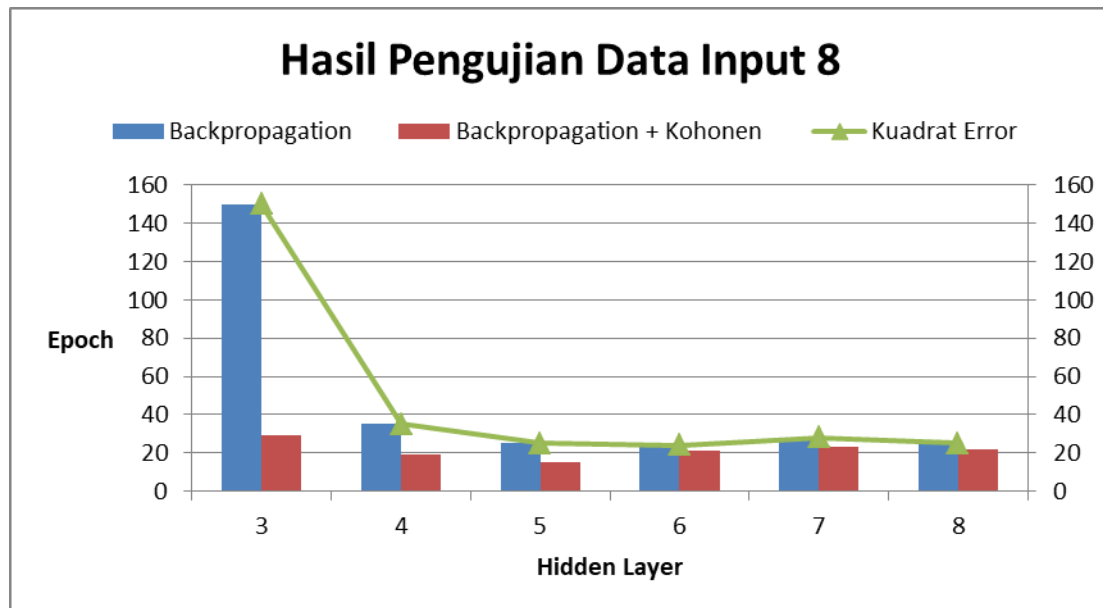
**Table 6.** Training Result With Input Data 8

Data Input	8					
Hidden Layer	3	4	5	6	7	8
#	Epoch					
<b>B</b>	150	35	25	24	28	25
<b>B + K</b>	29	19	15	21	23	22

B = Back propagation K = Kohonen

The test result of 8 is better than 3,4,5,6,7 actually 8 The test results with input data 8 have an epoch value that is better than testing using the amount of input data 3,4,5,6,7. The test results using layer 5 hidden have epoch 15 value which is better than the hidden layer 3,4,6,7,8.

In the test results it can be seen that different target errors will produce a different number of epochs. The smaller the target error, the greater the number of iterations. In the table it can also be seen that the smaller the hidden layer has the higher the epoch value compared to the greater the number of hidden.



**Figure 1.** Error Quadrate Graphic in Data Input 8

In figure 1, it can be seen that the decrease in testing variations in the parameters of the number of hidden layers 3, 4, 5, 6, 7, 8 with input data 8 produces an epoch in the number of layer 5 hidden with the number epoch 15.

## 5. Conclusion

The test results with input data 8 have an epoch value that is better than testing using the amount of input data 3, 4, 5, 6, 7. The test results using layer 5 hidden have epoch 15 value which is better than the hidden layer 3, 4, 6, 7, 8. In the test results it can be seen that different target errors will produce a different number of epochs. The smaller the target error, the greater the number of iterations. In the table it can also be seen that the smaller the hidden layer has the higher the epoch value compared to the greater the number of hidden.

## References

- [1] Sutojo, T., et al, 2010, *Kecerdasan Buatan*, Yogyakarta: Andi Offset.
- [2] Andayani, A., Nababan, E.B., Siregar, B., Muchtar, A.M. & Hamonangan, T. 2017. Optimization Backpropagation Algorithm Based on Nguyen-Widrow Adaptive Weight and Adaptive Learning Rate. *International Conference on Industrial Engineering and Applications*, pp. 363-367.
- [3] Sotirov, S. & Atanassov, K. 2010. Generalized Net Model for Parallel Optimization of Multilayer Perceptron with Momentum Backpropagation Algorithm. *IEEE Xplore Digital Library*, pp. 1-5.
- [4] Wahyuni, I. et al 2017. Modeling Backpropagation Neural Network for Rainfall Prediction in Tengger East Java. *International Conference on Sustainable Information Engineering and Technology*, pp.1-6
- [5] Imran Shafi, et al 2006. Impact of Varying Neurons and Hidden Layers in Neural Network Architecture for a Time Frequency Application.
- [6] Singh U P, Tiwari A, et al. 1989, *Kohonen Neural Network Model References For Nonlinear Discrete Time Systems*, 3rd IEEE International Conference on

"Computational Intelligence and Communication Technology.

- [7] Jaiswal J K, Das R 2017, *Application Of Artificial Neural Networks With Backpropagation Technique In The Financial Data*, IOP Conf. Series: Materials Science and Engineering