International Journal of GEOMATE, Oct., 2017, Vol.13, Issue 38, pp.97-103
Geotec., Const. Mat. & Env., ISSN:2186-2990, Japan, DOI: http://dx.doi.org/10.21660/2017.38.68673

THE COMPARATION OF BACK PROPAGATION METHOD AND KOHONEN METHOD FOR GAS IDENTIFICATION

*Riki Mukhaiyar[1]

[1]ISPAI Research Group, Faculty of Engineering, Universitas Negeri Padang, Indonesia;

*Corresponding Author, Received: 17 Jan. 2017, Revised: 7 March 2017, Accepted: 3 April 2017

ABSTRACT: The identification process of gas flavor is conducted by using the output of gas sensor system to recognize a variety of gas flavor. The identification and analysis process of this system is processed by using an artificial neural network approaches those are back propagation and Kohonen method. According of the experiment's result, the best parameter for back propagation network is the momentum constant (α) = 0.7, the constant of the sigmoid function (β) = 4.5, constant learning (η) = 0.9, and the constant of convergence (ε) = 0001, convergence is achieved more or less in the 19 500 iterations (± 16 seconds). Meanwhile, the best classification for Kohonen network is for the output of 8 knots with an average of 80.7% uniformity (for a maximum of 500 times iteration, approximately ± 3 seconds). Thus, the best network to classify the signal pattern of gas flavor is back propagation network for the parameters (α) = 0.7, a constant sigmoid function (β) = 4.5, a constant learning (η) = 0.9, and a constant convergence (ε) = 0001.

Keywords: Kohonen, Back Propagation, Artificial Neural Network, Supervised Algorithm, Unsupervised Algorithm

1. INTRODUCTION

Materials can be identified through its unique flavor. There are people who are experts in classifying materials by smelling its flavor [1] [2] [3]. Though, some experts have a different opinion regarding to how accurate the rate of the classification process. Therefore, a new solution should be found to improve the flavor's sensibility of the system by implementing an additional instrument such as a smart device. One of the devices is gas sensor which has a number of full or semi conductive polymer-gas sensors in identifying the flavor of a material.

Another thing to obtain a better result is by implementing an analysis of a neural network system [4] [5]. The expected outcome is to recognize and classify gas flavor pattern through its learning mechanism by copying a mapping working of a human brain. In this research, two methods are used and compared i.e. Back Propagation and Kohonen approaches [6]. The two artificial neural networks are trained in order to identify and classify the pattern of a signal mapping of gas sensor.

2. SYSTEM SPECIFICATIONS

The sensor system is an instrument that is expected to be able to replace the human sensitivity of things such as taste, odor, light, smell, flavor, etc. In general, image sensor system can be viewed on the Fig.1 below.

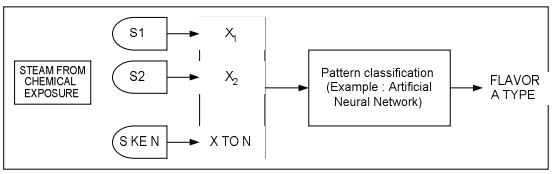

Fig.1. A System of the Flavour's Sensoric

According to the figure, gas sensor system (s_1 to s_n) receives flavor from one material (e.g. chemical material), then each sensor will provide an electrical signal (x_1 to x_n) [3].This signal will be processed by a pattern classification system. Then, the system produces the classes of the information about the flavor.

Based on its structure, the sensor system consists of several parts: some sensors, an interface of the classifier, and a signal processing system. The initial concept of the gas detection system is that the electrical resistance of a sensor will be changed at the moment of a gas molecule absorbed on its surface. It is shown from the changing of its potential difference. However, the changing of the electrical resistance depends on the types of the sensors used and the gas detected as well. This changing information is known as the characteristic of the recognized gas. Furthermore, the interfacing step is needed to make sure that the potential difference data is recognized and adapted well with the signal processing process. The step is important because an input and output of the interfacing process is different, i.e. analog and digital [7] [8].

3. METHOD USED

In this research, two neural networks learning methods, Back Propagation and Kohonen, are used to identify the flavor of several gases. Later, the results of those methods aims are compared and analyzed to find out which implementation is better.

4. 1. Back Propagation Method

In this method, the implementation processing steps used are illustrated as follows:

a. Initialization Weight
Make initial weight, and the initial bias in the form of small random numbers.

b. Calculate the activation
Activation is the output of a node. Activation of the input node is the value of the input to the input node. Activation of the output node and the node between the

$$O_j = F\left(\sum_i W_{ji} O_i - \theta_j\right) \quad (1)$$

With W_{ji} is the weight of the input O_i and θ_j is the threshold of node j. while F is the activation function in the form of sigmoid function

$$F(a) = \frac{1}{1 + e^{-\beta x a}} \quad (2)$$

c. Weight Training
Starting from the output node and forwarded to the weights between input nodes and nodes between, repeatedly. Changes in the weight given by:

$$W_{ji}(t+1) = W_{ji}(t) + \Delta W_{ji}(t) \quad (3)$$

With W_{ji} (t) is the weight from node i to j at the iteration t, is the change of weight ΔW_{ji} given by:

$$\Delta W_{ji}(t) = \eta \delta_j O_i \quad (4)$$

with η is a constant learning (0 <η <1) and δj is the gradient of the error of node j. Sometimes the weight plus the rate of change of momentum:

$$W_{ji}(t+1) = \mathrm{W}_{ji}(t) + \eta \delta_j O_i + \alpha\left(W_{ji}(t) - W_{ji}(t-1)\right) \quad (5)$$

The momentum of factor α value is varied between 0 and 1.

Error gradient is given by:

$$\delta_j = O_j(1 - O_j)(T_j - O_j)$$
$$\delta_j = O_j(1 - O_j)\sum_k \delta_k W_{ji}(t) \quad (6)$$

where T_j is the target output (the expected output) and δ_k is the gradient of the error from node to node k is connected with the j-th.

d. Iteration is repeated continuously until reaching a convergence criterion (or criteria for stopping). One of the convergence criteria is a common limitation of the average squared error (mean square error) with a small value ε.

$$E = \frac{1}{N}\sum_{j=1}(T_j - O_j)^2 \quad (7)$$

where N is the number of trained data.

The input of the system is the signal of eight sensors, so the total numbers of the input nodes are eight (first node for first sensor, second node for second sensor, and so on). Meanwhile, the output results are the class of the gas flavor. Besides, the total numbers of outputs are three (the first node is E with light flavor, the second node is M with mint flavor, and the third node is P with strong flavor). For example, the target output for class E is 1 for first output node, 0 for second node, and 0 for third node, 100. Meanwhile, the target output for class EP is 1 for the first and third node, and 0 for the second output node, 101. The total node of hidden layer is taken six nodes. It means that the grid structure, overall, is 8x6x3.

To obtain the best network for the neural network system, the parameter of this network are varied, such as the constant learning (η), momentum constant (α), the constant of sigmoid function (β), and the convergence constant (ε). The first step to vary the network is by using a quite large convergence constant (0.1). The purpose is to select the quite fast parameter of constant momentum and constant learning (β taken 2.5^8). Moreover, the rate of constant convergence is reduced to 0.001 to get a faster constant momentum and to search a constant learning. After obtaining the two fastest constants, the constant of sigmoid function is modified to obtain the fastest and most accurate training process (the accuracy is the point of this section). The next step is by varying the constant convergence, and finding out the constant relationship with the errors as a result of the obtained network.

4. 2. Kohonen Method

In principle, this method changes the input of

N-dimensional signal into a map discrete with one or two dimensions. Then, the method makes this transformation in accordance of its layout structure (topological ordered fashion). Therefore, this network only needs an input signal so that it will set its weight automatically [9].

There are three basic natures of this network [10]. The first one is approximation of the input space. It means that the network should provide a good approximation of the input space. The second one is the topology order. It means that the network is the spatial location of a node in the output layer corresponds to some domains or characteristics of input patterns. The third one is matching density. This network reflects the variations in the statistics of the input's distributions. Dense regions will be mapped tightly.

The amount of activations or outputs of some output nodes depend on their weight vector. The activations are greater if the vectors weight is closer to the input vector. After computing all activations, all output nodes are going to be matched. The winner is given activation 1, and may modify it weight vector based on the rules of the learning process.

In the early steps, the network is labeled. Two input nodes are taken (for the x-axis and the y-axis). The number of fictitious and output data distributed uniformly and the number of nodes is retrieved with a varied output. On this stage, the network performance could be observed by displaying the output signal weights. The observations of the weight distributions of the output nodes and lay-outing relationships between the output nodes are done. Therefore, maximum iteration and iterative neighbor reductions are varied to obtain the best and fastest systems.

In the real training stage (second stage), eight sensor signals are as the inputs. The numbers of output nodes are varied to see the close relationships within a group (the very close data). Meanwhile, in Kohonen network, the parameters are maintained constant but the numbers of output nodes are varied [11]. This is because the criteria for the completion of this network do not depend on the error, even error is not calculated, but it depends on the numbers of maximum iterations. The numbers of output related to the class of classifications that can be formed [12]. This is related to the classifications of input data that may occur. By varying the numbers of output nodes, the best networks that fit the classification pattern of gas sensor can be analyzed [13].

Total output of the selected node is 5 nodes to 9 nodes. Since there are five groups of classifications (E, M, P, EM, EP), the total output nodes selected is five, while nine output nodes are taken because the layers of 3x3 is in a form of the largest square that still have comparisons within the data classifications with the small enough amount of data (9:30 or 3:10 or about 1:3).

4. RESULTS AND ANALYSIS

4. 1. Back Propagation

1) Results

In order to obtain the best parameter network, varying the parameter of constant learning (η), momentum constant (α), the constant of the sigmoid function (β), and convergence constant (ε) were done. First of all, the convergence constant taken is 0.1. Results of variations are in the following table:

Table 1 Variation Results in Obtaining the Best Parameters

α	β	η	ε	Iteration
0.1	2.5	0.1	0.1	121680
0.2				106920
0.3				93140
0.4				75120
0.5				70980
0.6				49020
0.7				35940
0.8				36980
0.9				15540
1.0				∞

Two parameters of momentum constant (α) which have the smallest iteration were taken. If the iterations had been done for many times (200000-300000) but the network errors were still large in numbers, then it can be concluded that the network does not reach the convergence.

Table 2 Variation Results in Obtaining the Lowest Iteration for β = 2.5; ε = 0.001; and α = 0.7 and0.9

α	β	η	ε	Iteration
0.7	2.5	0.1	0.001	207360
		0.2		∞
		0.3		71640
		0.4		64540
		0.5		42080
		0.6		36420
		0.7		35780
		0.8		24200
		0.9		22900
		1.0		31140
0.9	2.5	0.1	0.001	95080
		0.2		∞
		0.3		24740
		0.4		∞
		0.5		∞
		0.6		∞
		0.7		∞

α	β	η	ε	Iteration
0.7	2.5	0.1	0.001	207360
		0.8		∞
		0.9		∞
		1.0		∞

The next step is obtaining the best parameter by trying out the constant of sigmoid function (β). From the table below it can be seen that the fastest iteration is owned by the arrangement of parameter α = 0.7, η = 0.3, with 24740 iterations. Using this parameter, the constant of sigmoid function is varied where the numbers of errors are also calculated. The numbers of errors obtained if the total number of output nodes which are having output do not reach the target with a tolerance of 0.1. For example, if the first output node is supposed to have one output, then if its output is less than 0.9 then it considered wrong and the wrong number is plus one.

Table 3 Variation Results of Sigmoid Function

A	β	η	ε	Iteration	False
0.7	1.0	0.9	0.001	36880	2
	1.5			33220	2
	2.0			21340	2
	2.5			22900	1
	3.0			24400	2
	3.5			31720	1
	4.0			18520	2
	4.5			19500	1
	5.0			22800	1
	5.5			39440	2
	6.0			∞	-
	6.5			∞	-
	7.0			∞	-
	7.5			∞	-
	8.0			∞	-
	8.5			∞	-
	9.0			∞	-
0.9	1.0	0.3	0.001	256800	4
	1.5			194520	3
	2.0			75300	1
	2.5			24740	2
	3.0			56460	1
	3.5			127340	5
	4.0			∞	-
	4.5			∞	-
	5.0			∞	-
	6.0			∞	-
	7.0			∞	-
	8.0			∞	-
	9.0			∞	-

Two of the best parameters (underlined) of this experiment were taken for the next experiment. In the subsequent experiments, the constant convergences are varied from 0.1 to 0.00001. Like the previous experiments, the numbers of errors are calculated. For this last stage of experiments, the numbers of iterations are unlimited. The results of parameters experiments can be seen in the following table:

Table 4 Variation Results for an Unlimited Number of Iterations

A	β	η	ε	Iteration	False
0.7	4.5	0.9	0.1	10400	13
			0.01	15940	3
			0.001	19500	1
			0.0001	46120	1
			0.00001	413660	1
0.7	5.0	0.9	0.1	10140	10
			0.01	18720	4
			0.001	22800	1
			0.0001	43500	1
			0.00001	295320	1

2) Analysis

In back propagation network, the best parameter is α = 0.7, β = 4.5, η = 0.9, and ε = 0001, where the convergence reached in 19500 of iterations and only one error in the stage of experimenting.

Based on the table of the testing results of constant convergence, it is shown that changing the constant of 0.00001 does not change the convergence to the error. It means that constant of convergence 0.001 is accurate enough for constant learning configuration of 0.9, momentum constant 0.7 and constant of sigmoid function 4.5 and 5.0. Constant convergence of 0.001 is the best constant convergence because the number of error is equal to the smallest constant of convergence with the number of iteration is much smaller, which means it will save learning time.

4. 2. Kohonen

1) Results

At first, the networks were tested with the data which its classification had been known before. The number of maximum iterations was 500, while distance reduction/number of neighbors of iteration was 10 (every 10 iteration, distance reduction minus 1), with the iteration for the number of neighbor distance one to the wins was 250. It means that up to 250 of iteration, the number of neighbor will not become zero. This is then referred to as neighbor iteration 1. The early constant learning was 1.0. The declined trend for all iteration is continuously continued until 0.01. Such a configuration is quite stable which means that the result of classification will not change even if the number of maximum iteration took 5000 iterations and neighbor's iteration 1 took 1500 iteration except for the number of output nodes were 8 nodes.

All data were drilled. The following table contains of number of data with its classification.

In Kohonen network, the output numbers of the classification result were not an exact numbers. It may change in the other trainings.

Table 5 Numbers of Classification

No.	No. Data	Group
1.	1-4, 6, 7, 11, 12, 17, 26, 27, 30	E
2.	13-16, 19	M
3.	18, 20, 22-25, 28	P
4.	8-10, 21, 29	E dan P
5.	5	E dan M

These sensor data were drilled on Kohonen network. The results of the drills for the amount of output varies were observed. The results can be seen in the following tables:

Table 6 Results for Five Nodes Group

Unit	Data to	Data type
1	17-25	EP, E, M, P
2	1-7	E, EM
3	9-16	EP, E, M
4	27, 28	E, P
5	8, 29-30	EP, E

In the output amounted 5 nodes can be described as follows:

- Group 1 represent group P where 6 out of 7 data P were in this group. The uniformity ratio is 6 of data P from 9 data classified in this group (6:9)
- Group 2 represent group E where 5 out of 12 data E were in this group. The uniformity ratio is 6:7
- Group 3 is the group M where 4 out of 5 groups were in this group. The uniformity ration is 4:8
- Group 4 classified two data only; E and P (conclusion grouping cannot be drawn)
- Group 5 is the group EP where 2 out of 5 data EP were in this group. The uniformity ratio is 2:3

Table 7 Results for Six Nodes Group

Unit	Data to	Data type
1	20-25	EP, P
2	3, 5, 17-19, 26	E, P, M, EM
3	9-16	EP, E, M
4	27,28	E, P
5	1, 2, 4, 6, 7	EP, E
6	8, 29, 30	EP, E

In the output amounted 6 nodes can be analyzed as follows:

- Group 1: Group P (5 out of 7 data P were in this group with the uniformity ratio is 5:6)
- Group 2: consist of various groups (combined groups), with the largest group was E. the uniformity ratio was 3:6, while the EM which consists of one word was in this group
- Group 3: Group M (4 out of 5 data). The uniformity ratio was 4:8
- Group 4: classifying two data only; E and P (conclusion grouping cannot be drawn)
- Group 5: Group E (5 out of12 data with the uniformity ratio was 5:5)
- Group 6: Group EP (two out of 5 data, with the uniformity ratio was 2:3)

Table 8 Results for Seven Nodes Group

Unit	Data to	Data type
1	11-16	E, M
2	9, 10	EP
3	1, 2, 4, 6, 7	E
4	3, 5, 17-19	E, EM, P, M
5	27	E
6	8, 19, 30	E, EP
7	20-25	P, EP

In the output amounted 7 nodes can be analyzed as follows:

- Group 1 : Group M (4 out of 5 data M were in this group with the uniformity ratio was 4:6)
- Group 2 : Group EP (2 out of 5 data EP were in this group with the uniformity ratio was 2:2)
- Group 3: Group E (5 out of 12 data E were in this group with the uniformity ratio was 5:5)
- Group 4: Combined group with the largest group was E. The uniformity ratio was 3:6, while EM which consists of one word was in this group
- Group 5: Classifying data E only
- Group 6: Group EP (2 out of 5 data with the uniformity ratio was 2:3)
- Group 7: Group P (6 out of 7 data with the uniformity ratio was 6:7)

Table 9 Results for Eight Nodes Group

Unit	Data to	Data type
1	11-16	E, M
2	9-10	EP
3	1, 2, 4, 6, 7	E
4	17, 18, 26	E, P
5	3, 5, 19	E, EM, M
6	8, 29, 30	E, EP
7	20-25	P, EP
8	27, 28	E, P

In the output amounted 8 nodes can be analyzed as follows:

- Group 1: Group M (4 out of 5 data M were in this group with the uniformity ratio was 4:6)
- Group 2: Group EP (2 out of 12 data EP were in this group with the uniformity ratio was 2:2)
- Group 3: Group E (5 out of 12 data with the uniformity ratio was 5:5)
- Group 4: Group E (2 out of 12 data with the uniformity ratio was 2:3)
- Group 5: Only classifying three data with three different flavors (E, EM, and M with one data of each)
- Group 6: Group EP (2 out of 5 data with the uniformity ratio was 2:3)
- Group 7: Group P (5 out of 7 data with the uniformity ratio was 5:6)
- Group 8: Classifying data E and P (conclusion grouping cannot be drawn)

Table 10 Results for Nine Nodes Group

Unit	Data to	Data type
1	11-16	EM
2	-	-
3	20-25	P, EP
4	9, 10	EP
5	-	-
6	5, 17-19, 26	EM, E, P, M
7	29, 30	E, EP
8	1-4, 6, 7	E
9	27	E

In the output amounted 9 nodes can be analyzed as follows:

- Group 1: Group M (4 out of 5 data M were in this group, with the uniformity ratio was 4:6)
- Group 2 and group 5 did not have data classified
- Group 3: Group P (6 out of 7 data with the uniformity ratio was 6:7)
- Group 4: Group EP (2 out of 5 data with the uniformity ratio was 2:2)
- Group 6 represents the combined group of 2E, IM, IP, and IEM
- Group 7 classifying data E and P (conclusion grouping cannot be drawn)
- Group 8: Group E (6 out of 12 data with the uniformity ratio was 6:6)

The time to do the training for the Kohonen network (for 500 times maximum iteration) was about two to five seconds.

2) Analysis

In the Kohonen network, it is shown that the increasing numbers of classes will be increasingly diverse groups formed. In this network, the uniformity of data classification was noted.

Based on the results and analysis above, it can be seen that the increasing number of output nodes make the classifications more diverse, ranging from four groups, five groups, and six groups. This is in accordance with the logic that the more output nodes, the more classifications done. In nine output nodes, there were two nodes that did not classified. This is because the position of topology did not allow the classification or the ninth nodes took the classification of those two nodes.

5. THE COMPARATION OF BACK PROPAGATION METHOD AND KOHONEN METHOD

If the network using Kohonen Method compared with the network using back propagation method, it seems clear that:

a. The Kohonen method is much faster for the above classification results with only 500 iterations.
b. Based on its accuracy, back propagation method successfully classifies better.

6. CONCLUSIONS

In the back propagation network, the best parameter is by using momentum constant of 0.7, constant of sigmoid function 4.5. The convergence is achieved within 19500 iterations with only one error during the process. The accuracy of this network is 95%.

The network using Kohonen method, the best classification is the network with its output total number is eight nodes with the uniformity average

are 80.7%. In general, with maximum iteration 500 times, the uniformity average is 70.9%. The classification groups are group E (2 groups), group EP (2 groups), group P, and group M. Thus, the best network in classifying signal pattern of gas sensor is the network using back propagation method.

The supervised method is accurate for pattern classification, but the training takes time. On the other hand, the unsupervised method has the advantages in terms of time, due to the availability of the new grouping. Development can be done by combining the network using back propagation method and the network using Kohonen method.

7. REFERENCES

[1] R. Mukhaiyar, S.S. Dlay, and W. L. Woo, "Cancellable Biometric using Matrix Approaches", Newcastle University, Thesis, UK, 2015.

[2] R. Mukhaiyar, S.S. Dlay, W.L. Woo, "Alternative Approach in Generating Cancellable Fingerprint using Matrices Operations", in Proceeding 56th ELMAR (2014 International Conference Electronic in Marine), pp. 1-4, Zadar - Croatia, September 2014.

[3] R. Mukhaiyar, "Core-Point, Ridge-Frequency, and Ridge-Orientation Density Roles in Selecting Regional of Interest of Fingerprint," International Journal of GEOMATE, Vol.12, Issue 30, pp. 146-150, 2017.

[4] E. Sanchez-Sinencio and C. Lau, "Artificial Neural Networks: Paradigms, Applications, and Hardware Implementations", IEEE Press, New York, 1992.

[5] S. Kusuma dewi, "Membangun Jaringan Syaraf Tiruan Menggunakan MATLAB dan EXCEL LINK", Edisi Pertama, Penerbit Graha Ilmu, Yogyakarta, 2004.

[6] R. Mukhaiyar, "Identifikasi Aroma Menggunakan Sistem Jaringan Syaraf Tiruan Metoda Kohonen", in Proceeding Seminar Nasional Rekayasa Sains dan Teknologi, pp. 467-477, Padang, 2010.

[7] T. Masters, "Signal and Image Processing with Neural Networks: A C++ Sourcebook", John Wiley & Sons, Inc, 1994.

[8] A.D. Kulkarni, "Artificial Neural Networks for Image Understanding", Van Nostrand Reinhold, USA, 1994.

[9] A. Roy, M. Matos, K. J. Marfurt, "Automatic Seismic Facies Classification with Kohonen Self Organizing Maps – a Tutorial", Geohorizons, pp. 6-14, 2010.

[10] T. Kohone, "MATLAB Implementations and Applications of the Self-Organizing Map", School of Science, Aalto University, Finland, 2014.

[11] D. Markovic, M. Madic, V. Tomic, S. Stojkovic, "Solving Travelling Salesman Problem by Use of Kohonen Self-Organizing Maps", Acta Technica Corviniensis-Bulletin of Engineering Tome V, pp. 21-24, 2012.

[12] M. Ettaouil, E. Abdelatifi, F. Belhabib, and K. E. Moutaouakil, "Learning Algorithm of Kohonen Network with Selection Phase", WSEAS Transactions on Computers, vol. 11, issues 11, pp.387-396, Nov. 2014.

[13] P. Schneider, "Advanced Methods for Prototype-Based Classification", University of Groningen, University Medical Center Groningen, pp.5-91, 2010.