International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-4, November 2019

# Accuracy of the Neurons Number in the Hidden Layer of the Levenberg-Marquardt Algorithm

**Hindayati Mustafidah, Suwarsito, Silvia Nila Candra Permatasari**

***Abstract*****:** ***Backpropagation, as a learning method in artificial neural networks, is widely used to solve problems in various fields of life, including education. In this field, backpropagation is used to predict the validity of questions, student achievement, and the new student admission system. The performance of the training algorithm is said to be optimal can be seen from the error (MSE) generated by the network. The smaller the error produced, the more optimal the performance of the algorithm. Based on previous studies, we got information that the most optimal training algorithm based on the smallest error was Levenberg–Marquardt with an average MSE = 0.001 in the 5-10-1 model with a level of α = 5%. In this study, we test the Levenberg-Marquardt algorithm on 8, 12, 14, 16, 19 neurons in hidden layers. This algorithm is tested at the learning rate (LR) = 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, and 1. This study uses mixed-method, namely development with quantitative and qualitative testing using ANOVA and correlation analysis. The research uses random data with ten neurons in the input layer and one neuron in the output layer. Based on ANOVA analysis of the five variations in the number of neurons in the hidden layer, the results showed that with α = 5% as previous research, the Levenberg–Marquardt algorithm produced the smallest MSE of 0.00019584038 ± 0.000239300998. The number of neurons in the hidden layer that reaches this MSE is 16 neurons at the level of LR = 0.8.***

***Keywords*****:** ***accuracy, hidden layer, Levenberg-Marquardt, MSE.***

## I. INTRODUCTION

One of the results of the development of intelligent computing-based technology is artificial neural networks (ANN), which are biologically inspired, computational models. One of the findings of the ANN model that is in demand by many people is backpropagation (BP), which has proven for its success in solving computational problems. Some of the results of BP's implementation of research in various fields are [1] and [2] applying BP to predict the level of question validity and accuracy of data pattern recognition, [3] applying BP to predict student achievement values in subjects tested in the National Examination at high school. Besides, [4] apply BP to determine the level of qualifications of prospective students in the new student admission information system.

ANN consists of several processing elements (neurons) and is connected by neurons called weights [5]. In its architecture, many layers of ANN consist of the input layer (IL), hidden layer (HL), and output layer (OL). Each layer is connected to produce the expected output. Using too few neurons in hidden layers will result in something that is underfitting. Likewise, when too many neurons in hidden layers can cause several problems, namely overfitting [6]. In determining the number of neurons used in hidden layers there are several rules, namely the number of neurons in HL should be between the size of IL and OL, or 2/3 of the size of IL plus the size of OL, or smaller (less than) twice the size of IL [7].

Calculation of errors or MSE (Mean Squared Error) is a measurement of how the network can learn so well that when compared to new patterns, it will be easily recognized. We can analogize this MSE value as a variant added to the bias square of a model, where is the model that is evaluated and is the actual model. If there is no bias between the model built with the actual model (unbiased model), then the MSE is commensurate with the variant of the model. The smaller the variant of a model, the more robust the model is in forecasting [8]. MSE is a network performance function that measures performance based on the average of the error squares [9]. This MSE is one of the performance benchmarks of ANN.

There are 12 training algorithms in BP that have different levels of performance. Based on research conducted by [10], [11], [12] of the 12 training algorithms in BP, which have the best performance, is the Levenberg-Marquardt (LM) algorithm. This study used ten neurons in hidden layers. Therefore in this study, the LM algorithm was tested using variations in the number of neurons in HL to get the most optimal performance.

## II. METHOD

The implementation of this study begins with developing a computer program using the MATLAB programming language along with its toolbox [13] to run the LM algorithm. The ANN program design used is presented in Fig. 1 [14] with a network structure, as shown in Fig. 2.

Manuscript published on November 30, 2019.
* Correspondence Author
**Hindayati Mustafidah***, Informatics Engineering, Universitas Muhammadiyah Purwokerto, Purwokerto, Indonesia. Email: fida.mustafidah@ump.ac.id
**Suwarsito**, Geography, Universitas Muhammadiyah Purwokerto, Purwokerto, Indonesia. Email: suwarsito.ito@ump.ac.id
**Silvia Nila Candra Permatasari**, Informatics Engineering, Universitas Muhammadiyah Purwokerto, Purwokerto, Indonesia. Email: silvianila765@gmail.com

Furthermore, the data generated from the LM algorithm are analyzed using statistical tests with test design, as in Fig. 3.

### A. Research Variable

The variables in this study consisted of:

- Value of learning rate (LR): 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1
- Target error: 0.001 ($10^{-3}$)
- Number of neurons in the IL: 10
- Number of neurons in the OL: 1
- Number of neurons in HL: 8, 12, 14, 16, 19

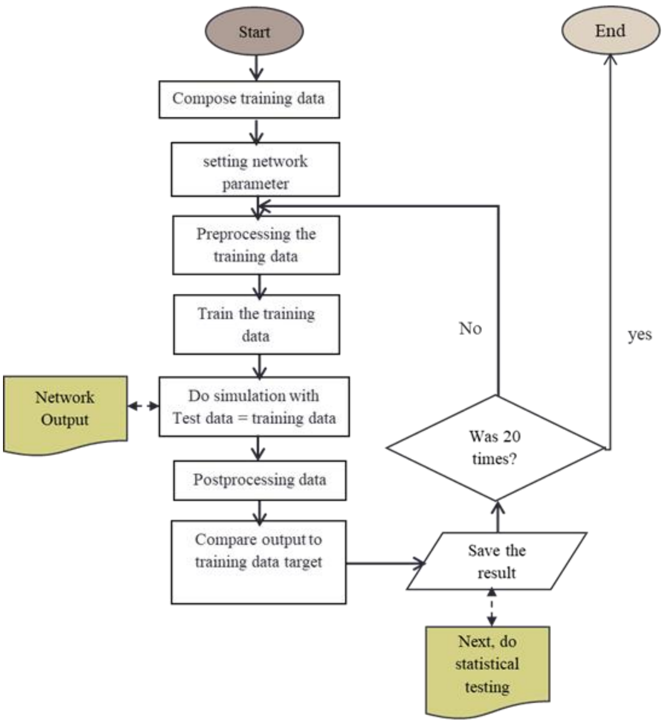


**Fig. 1. Flowchart of ANN computer program [14]**

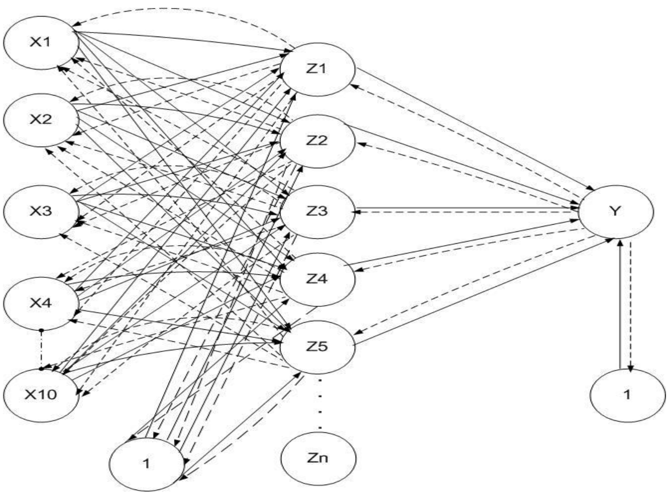


**Fig. 2. Design of ANN structures with 10 IL neurons, n HL neurons, and 1 OL neurons (models 10-n-1; n = 8, 12, 14, 16, 19)**

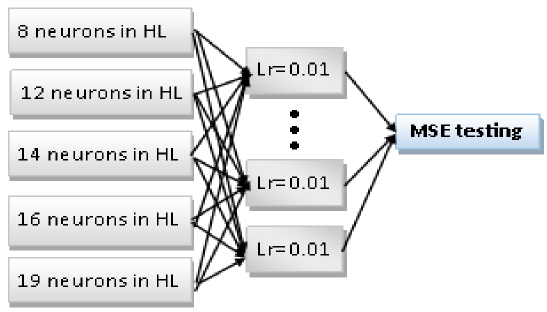


**Fig. 3. ANOVA testing design**

### B. Data Collecting

Data sources of IL neurons and target neurons used are documentation data from research by [15].

### C. Data Analysis

ANN output data were analyzed using the ANOVA test with the following stages [16] as in Fig. 3:

- Hypothesis formulation.
  - $H_0$: no difference in the MSE generated by the LM algorithm based on the number of neurons in the HL in each value of LR.
  - $H_1$: there are differences in MSE generated by the LM algorithm based on the number of neurons in the HL in each value of LR.
- determine $\alpha = 5\%$.
- make a decision: $H_0$ is rejected if sig.< $\alpha$

The subsequent data analysis is to conduct a bivariate correlation test between the LR values used and the average MSE in each number of neurons in the hidden layer using Pearson correlation at $\alpha = 5\%$ with the hypothesis:

- $H_0$: no correlation the LR values used and the average MSE in each number of neurons in the hidden layer
- $H_1$: any correlation the LR values used and the average MSE in each number of neurons in the hidden layer.

The decision will reject $H_0$ if sig.< $\alpha$.

## III. RESULT AND DISCUSSION

### A. Research Data

Data on neurons in IL and targets generated randomly using the MATLAB program as presented in Table I [15]. ANN output data, MSE, is obtained by training the LM algorithm 20 times. This paper is not presenting the MSE data that is a network output.

### B. Data Analysis

After being tested using the ANOVA test, it obtained a significant value (sig.) according to the number of neurons in HL (Table II).

Information is obtained from the data in Table II that the values of sig. is above the value of $\alpha$ (5%), which means the null hypothesis ($H_0$) is accepted. This value shows that there is no difference in MSE generated by the network using the LM algorithm for each number of neurons in HL used. In other words, various variations in the number of neurons in HL do not significantly affect the difference in MSE results.

Nevertheless, it is necessary to find information on the smallest MSE value in each value of LR for each number of neurons in HL used. Table III presented this MSE value.

Based on Table III, the LM algorithm achieved the smallest MSE at the level of LR = 0.8 with 16 neurons in HL. Subsequent tests were carried out on each of the many neurons in HL, namely 8, 12, 14, 16, 19, to obtain the smallest MSE with the results of the analysis, as in Table IV.

Table IV shows the value of sig.=0.000, which means that there are differences in MSE that are produced based on the number of neurons in HL.

This difference can be seen in Table V after the MCA (Multiple Comparison Analysis) tests using the Duncan method.

The relationship between the value of LR and the number of neurons in HL in producing MSE can be known by analyzed using correlation analysis. The smaller the value of LR and the more neurons in HL, the smaller the MSE produced will be. A negative correlation value described this relationship. The results of the correlation analysis between LR and each number of neurons in HL are presented in Table VI with graphs, as presented in Fig. 4.

Table VI shows the smallest Pearson correlation value as MSE produced by 16 neurons in HL with a correlation value of -0.526 at a significance level of 0.079. Although the value of sig. is higher than α, which means that it does not correlate significantly, but based on the graph in Figure 4 shows a tendency to correlate with the increase in the number of neurons in HL used to produce MSE. The research of [17] also supported no-correlation between the MSE produced by the LM algorithm and the value of LR.

**Table- I: ANN data input of 10 neurons IL (Xi) and target data (Y) [15]**

| $X_1$ | $X_2$ | $X_3$ | $X_4$ | $X_5$ | $X_6$ | $X_7$ | $X_8$ | $X_9$ | $X_{10}$ | Y |
|---|---|---|---|---|---|---|---|---|---|---|
| 5.8279 | 2.0907 | 4.1537 | 2.1396 | 6.8333 | 4.5142 | 6.0854 | 0.8408 | 1.2105 | 2.3189 | 4.3979 |
| 4.2350 | 3.7982 | 3.0500 | 6.4349 | 2.1256 | 0.4390 | 0.1576 | 4.5436 | 4.5075 | 2.3931 | 3.4005 |
| 5.1551 | 7.8333 | 8.7437 | 3.2004 | 8.3924 | 0.2719 | 0.1635 | 4.4183 | 7.1588 | 0.4975 | 3.1422 |
| 3.3395 | 6.8085 | 0.1501 | 9.6010 | 6.2878 | 3.1269 | 1.9007 | 3.5325 | 8.9284 | 0.7838 | 3.6508 |
| 4.3291 | 4.6110 | 7.6795 | 7.2663 | 1.3377 | 0.1286 | 5.8692 | 1.5361 | 2.7310 | 6.4082 | 3.9324 |
| 2.2595 | 5.6783 | 9.7084 | 4.1195 | 2.0713 | 3.8397 | 0.5758 | 6.7564 | 2.5477 | 1.9089 | 5.9153 |
| 5.7981 | 7.9421 | 9.9008 | 7.4457 | 6.0720 | 6.8312 | 3.6757 | 6.9921 | 8.6560 | 8.4387 | 1.1975 |
| 7.6037 | 0.5918 | 7.8886 | 2.6795 | 6.2989 | 0.9284 | 6.3145 | 7.2751 | 2.3235 | 1.7390 | 0.3813 |
| 5.2982 | 6.0287 | 4.3866 | 4.3992 | 3.7048 | 0.3534 | 7.1763 | 4.7838 | 8.0487 | 1.7079 | 4.5860 |
| 6.4053 | 0.5027 | 4.9831 | 9.3338 | 5.7515 | 6.1240 | 6.9267 | 5.5484 | 9.0840 | 9.9430 | 8.6987 |

**Table- II: Value of sig. as a result of the ANOVA test for each number of neurons in HL**

| Num. of neurons in HL. | Sig. |
|---|---|
| 8 | 0.371 |
| 12 | 0.423 |
| 14 | 0.152 |
| 16 | 0.313 |
| 19 | 0.071 |

**Table- III: MSE value of the LM algorithm for each LR on several numbers of neurons in HL**

| LR | Neurons in HL | | | | |
|---|---|---|---|---|---|
| | *8* | *12* | *14* | *16* | *19* |
| 0,01 | 0,21429 | 0,00006 | 0,000176 | 0,000213 | 0,000220 |
| 0,05 | 0,58457 | 0,00025 | 0,000161 | 0,000341 | 0,000186 |
| 0,1 | 0,00020 | 0,07493 | 0,000154 | 0,000185 | 0,000179 |
| 0,2 | 0,24472 | 0,00015 | 0,000155 | 0,000161 | 0,000184 |
| 0,3 | 0,01913 | 0,00024 | 0,000233 | 0,000181 | 0,000102 |
| 0,4 | 0,18210 | 0,00015 | 0,000178 | 0,000205 | 0,000150 |
| 0,5 | 0,14088 | 0,00022 | 0,000219 | 0,000213 | 0,000347 |
| 0,6 | 0,00025 | 0,00016 | 0,000398 | 0,000238 | 0,000199 |
| 0,7 | 0,36695 | 0,00030 | 0,000252 | 0,000186 | 0,000177 |
| 0,8 | 0,19957 | 0,00020 | 0,000227 | 0,000121 | 0,000290 |
| 0,9 | 0,19951 | 0,06530 | 0,000277 | 0,000122 | 0,000267 |
| 1 | 0,30725 | 0,00018 | 0,000167 | 0,000183 | 0,000319 |

**Table- IV: ANOVA test results on the number of neurons in HL**

| | Sum of Squares | Df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Between Groups | 7.846 | 4 | 1.961 | 18.613 | 0.000 |
| Within Groups | 125.932 | 1195 | 0.105 | | |
| Total | 133.778 | 1199 | | | |

**Table- V: Duncan test results for each number of neurons in HL**

| | Neuron in Hidden Layer | N | Subset for *alpha* = 0.05 | |
|---|---|---|---|---|
| | | | *1* | *2* |
| Duncan[a] | 16 neurons | 240 | .00019584038 | |
| | 14 neurons | 240 | .00021647232 | |
| | 19 neurons | 240 | .00021833752 | |
| | 12 neurons | 240 | .01184589066 | |
| | 8 neurons | 240 | | .20495162762 |
| | Sig. | | .726 | 1.000 |

**Table- VI: Pearson Correlations between LR and several numbers of neurons in HL**

| | | LR | HL_8 | HL_12 | HL_14 | HL_16 | HL_19 |
|---|---|---|---|---|---|---|---|
| LR | Pearson Correlation | 1 | -,011 | ,013 | ,442 | -,526 | ,554 |
| | Sig. (2-tailed) | | ,974 | ,968 | ,150 | ,079 | ,062 |
| | N | 12 | 12 | 12 | 12 | 12 | 12 |

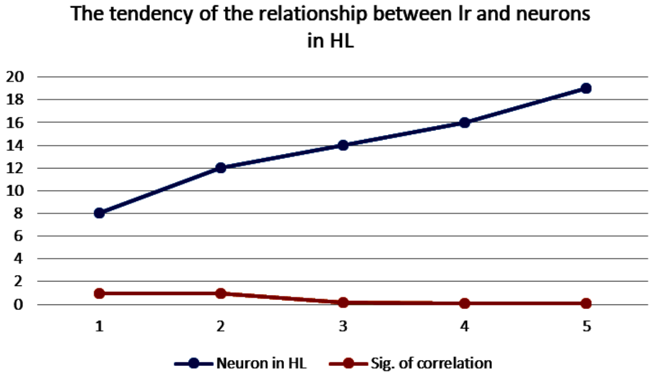


**Fig. 4. Graph of the tendency of the correlation between the value of LR and the number of neurons in HL**

## IV. CONCLUSION

ANOVA test results using α = 5% prove that the LM algorithm achieves optimal performance on the use of 16 neurons in the hidden layer compared to 8, 12, 14, and 19 neurons. This performance is achieved at a learning rate of 0.8 with ten neurons in the input layer and one neuron in the output layer, also known as the 10-16-1 model. MSE generated is 0.00019584038 ± 0.000239300998 of the target error of 0.001. Besides, this study also produces information that there is no correlation between the magnitude of the learning rate and the number of neurons in the hidden layer.

## ACKNOWLEDGMENT

Authors thank DP2M - The Ministry of Research, Technology, and Directorate General of Higher Education (DIKTI) through KOPERTIS Region VI, who has provided funds and facilities in the implementation of this research.

## REFERENCES

1. H. Mustafidah, S. Hartati, R. Wardoyo, and A. Harjoko, “Prediction of Test Items Validity Using Artificial Neural Network,” in *Proceeding International Conference on Education, Technology, and Science (NETS) 2013, “Improving The Quality Of Education To Face The Impact Of Technology”. December 28th, 2013*, 2013.
2. H. Mustafidah, S. Hartati, R. Wardoyo, and A. Harjoko, “Selection of Most Appropriate Backpropagation,” *Int. J. Comput. Trends Technol.*, vol. 14, no. 2, pp. 92–95, 2014.
3. H. Mustafidah, D. K. Hakim, and S. Sugiyanto, “Tingkat Keoptimalan Algoritma Pelatihan pada Jaringan Syaraf Tiruan (Studi Kasus Prediksi Prestasi Belajar Mahasiswa) Optimization Level of Training Algorithms in Neural Network (Case Studies of Student Learning Achievement Predictions),” *JUITA*, vol. II, no. 3, pp. 159–166, 2013.
4. D. C. Febrianto and H. Mustafidah, “Penerapan Jaringan Syaraf Tiruan dengan Metode Pembelajaran Backropagation untuk Mengetahui Tingkat Kualifikasi Calon Siswa pada Sistem Informasi Penerimaan Siswa Baru di MAN 2 Banjarnegara,” *JUITA (Jurnal Inform.*, vol. II, no. 3, pp. 189–197, 2013.
5. S. Shanmuganathan and S. Samarasinghe, *Artificial Neural Network Modelling*, vol. 628. Springer International Publishing, 2016.
6. J. Heaton, *Introduction to Neural Networks with C#*, Second. St. Louis: Heaton Research, Inc., 2008.
7. J. Heaton, *Artificial Intelligence for Humans, Volume 3: Neural Networks and Deep Learning*, 1.0. Chesterfield, USA: Heaton Research Inc., 2015.
8. E. L. Lehmann and G. Casella, *Theory of Point Estimation*, Springer t. Springer, 2003.
9. S. Kusumadewi, *Membangun Jaringan Syaraf Tiruan Menggunakan MATLAB & EXCEL LINK*. Yogyakarta: Graha Ilmu, 2004.
10. H. Mustafidah and S. Suwarsito, “Error Rate Testing of Training Algorithm in Back Propagation Network,” *Int. J. Soft Comput. Eng.*, vol. 5, no. 4, pp. 46 – 50, 2015.
11. H. Mustafidah and S. Suwarsito, “Uji Keoptimalan Algoritma Pelatihan pada Jaringan Syaraf Tiruan,” in *Prosiding Seminar Nasional SENATKOM 2015*, 2015, pp. 243–248.
12. H. Mustafidah and H. Harjono, “Korelasi Tingkat Kesalahan dan Epoh dalam Jaringan Backpropagation,” in *Prosiding SEMNASTIKOM 2017, 3 November 2017, ISBN: 978-602-50434-0-6*, 2017, pp. 55–61.
13. S. L. Lent, *Learning to Program with MATLAB: Building GUI Tools*. Department of Electrical Engineering University of Notre Dame: John Wiley & Sons, Inc., 2013.
14. H. Mustafidah and S. Suwarsito, “Testing Design of Neural Network Parameters in Optimization Training Algorithm,” in *International Conference of Result and Community Services, 6th August 2016*, 2016, p. THN. 139-146.
15. H. Mustafidah and S. Suwarsito, “Model Parameter Jaringan Syaraf Tiruan untuk Pemilihan Algoritma Pelatihan Jaringan Backpropagation yang Paling Optimal,” Purwokerto, Central Java, Indonesia, 2015.
16. T. Taniredja and H. Mustafidah, *Penelitian Kuantitatif (Sebuah Pengantar)*. Bandung: ALFABETA, 2011.
17. H. Mustafidah and S. Suwarsito, “Correlation Analysis Between Error Rate of Output and Learning Rate in Backpropagation Network,” *Adv. Sci. Lett.*, vol. 24, no. 12, pp. 9182–9185, 2018.

## AUTHORS PROFILE

**Hindayati Mustafidah** - graduated from the magister computer of the Faculty of Mathematics and Natural Sciences (MIPA) Universitas Gadjah Mada Yogyakarta in 2002, a lecturer and researcher at the Universitas Muhammadiyah Purwokerto, has published her researches in several national and international journals, also in several national and international seminars, domestic and abroad. The author's interest are intelligent systems (artificial neural networks, fuzzy logic, case-based reasoning, natural language processing, expert systems) with their implementation in various fields including education, fisheries, and health. The courses that the author teaches to students are computer programming, artificial intelligence, research methodology, computational theory, numerical computing, neural networks, fuzzy logic, computer reasoning, and soft computing. In addition to carrying out the main obligations of teaching and researching, the author manages national-scale scientific journals namely JUITA (Journal of Informatics) and JRST (Journal of Science and Technology Research) as chief editor and reviewer. The author is also incorporated in several national and international professional organizations such as INDOCEISS, APTIKOM, IAII, APSI, and IAENG.

**Suwarsito** - graduated from the Masters in Aquaculture, Faculty of Fisheries, Institut Pertanian Bogor – Indonesia in 2004, a lecturer and researcher at the Faculty of Education – Universitas Muhammadiyah Purwokerto. The author has published his research results in several national and international journals, also in several national and international seminars, domestic and abroad. The areas of specialization by the author are fisheries, aquatic science, climatology, marine science, and social geography. The author teaches several courses including statistics, research methodology, hydrology, climatology, and oceanography. In addition to carrying out the main obligations of teaching and researching, the author manages national-scale scientific journals, namely SAIN AKUATIK and SAINTEKS as editors and reviewers.

**Silvia Nila Candra Permatasari** - informatics graduate in the Faculty of Engineering and Science – Universitas Muhammadiyah Purwokerto in 2018, a research assistant who helped conduct research in the Laboratory of Intelligent Systems at the Universitas Muhammadiyah Purwokerto. This laboratory is under the coordination of the Informatics Engineering Study Program - Faculty of Engineering and Science. The author has attended training or seminars to support the work. The work that the author does is to prepare research tools and equipment, document research activities, and assist in data collection and analysis. Some research conducted in the laboratory of intelligent systems at Universitas Muhammadiyah Purwokerto is the development of applications using artificial neural network methods, fuzzy logic, case-based reasoning, and expert systems.