Building of Informatics, Technology and Science (BITS)
Volume 4, No 2, September 2022 Page: 1153−1158
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i2.2130

The Effect of Initialization Weights for Multi Layer Perceptron Performance on Prediction of House Construction Costs

Abdul Rozaq

Teknik, Teknik Informatika, Universitas PGRI Madiun, Madiun, Indonesia
Email: rozaq@unipma.ac.id
Email Penulis Korespondensi: rozaq@unipma.ac.id
Submitted:**18/08/2022**; Accepted:**30/09/2022**; Published: **30/09/2022**

Abstract−The house is one of the primary human needs besides food and clothing. Therefore, the community will always try their hardest to meet primary needs. For the middle and lower class people, it is going to very difficult to build a residential house because the income does not match the increase in house prices. With an artificial neural network, the middle to lower class people can estimate the costs that must be prepared if you want to build a residential house, of course this will be cheaper than using housing developer services. Based on the data that has been obtained, the researcher is then trained and tested using an artificial neural network with 13 data input, 25 hidden layers, a learning rate of 0.75, the number of iterations of 1000, the best test results are MSE value, 0.1, mean accuracy of 97.22 and computation time of 0,028 seconds.

Keywords: MLP; Performance; Prediction; Initialization; ANN

1. INTRODUCTION

The house is one of the primary human needs besides food and clothing. Because the community will always try their hardest to meet these primary needs. For the middle and lower class people, it will be very difficult to build a residential house because the income does not match the increase in house prices.

Market law says that the more demand, the more expensive the price, this also applies to house prices because there is more demand but not followed by additional land so that it will cause housing demand to overcapacity[1].

Based on this case, many people make houses by building themselves without going through a housing developer. In this way the community can save costs. In this way, a new problem will arise if we are not skilled at building houses, of course it will be difficult for us to determine how much it costs to build a house based on the type of building.

From the data that the researcher has obtained, as many as 200 data, with the distribution of training data by 80% and testing data by 20%, the author will calculate and analyze the effect of initializing weights to improve the accuracy performance and MSE multi-layer perceptron on the prediction of house construction costs. Where the weight value in the hidden layer and the weight value on the Output layer will be determined from the start, not do it randomly so that from the predetermined weight value from the beginning, the MSE value will be calculated at the time of training and when doing the test the error value will be calculated and compared between Real data from home developers with predictive data using ANN, from the results of the comparison, the accuracy value of each test data will be obtained, then the overall accuracy of the test data will be obtained.

Research [2] that the results of forecasting using ANN for Madura newspaper sales in 2015 with the best looping were 200 times, the best learning average was 0.6 with training data of 90% while the test data was 10%. the minimum error rate is 0.016.

According to [3] using an artificial neural network can produce an average accuracy of 99% with a mean error value of 0.0001

According to [4], the best results are with an MSE of 0.0009979 with an input value architecture of 4 features, a hidden value of 50 and an output value of 1 where these results are better than using an artificial neural network architecture with an input value of 4 features, there are 2 layers hidden values, namely 50 and 75, and the output value 1 with an MSE of 0.308878.

According to [5] in predicting the contents of the export and import of oil and gas in our country using ANN with 5 models, the best result 12-5-1 can reach 83% accuracy, MSE 0.0281641257 with an error rate used 0.001- 0.05.

In research [6]) that MAPE values below 10% have high accuracy if the forecasting model uses artificial neural networks.

In research [7] with 10 features for food security and food vulnerability in the province of West Papua, the resulting Spearman correlation value is 0.9375 by comparing the results of ANN and the original data.

2. RESEARCH METHODS

2.1 Research Step

a. Identification of problems

The growth in demand for land for residential houses is increasing but not followed by the addition of land so that land prices are getting more expensive, not to mention that the price of building materials continues to increase, this makes it difficult for new families to buy residential houses.

Building of Informatics, Technology and Science (BITS)
Volume 4, No 2, September 2022 Page: 1153−1158
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i2.2130

b. Data Collections

Researchers obtained data from housing developers as much as 200 data from various types of houses where training data were 160 or 80%, test data were 40 data or 20%. There are 13 independent data as input data and dependent data there is 1 data that will be predicted using artificial neural network

Independent data It is a variable that cannot be influenced by other variables but this variable can affect the results of other variables.

This variable is in the form of input data that will be used as input in the artificial neural network, in this study there are 13 data input in the form of data on building materials needed when building a residential house. The input data are: type, amount of cement, amount of sand, amount of bricks, amount of lime, number of splits, amount of iron, number of begel, number of roof beams, number of roof stretchers, number of roof battens, amount of wood and number of roofs.

Independent data Is a variable whose results are highly dependent on other variables. In this study, variable data is in the form of continuous data. This variable is an output in the form of real costs when building a residential house where this real data will be predicted using an artificial neural network so that accuracy will be obtained from the real data with the predicted data.

c. Preprocessing Data

At this step, discarding inconsistent data, completing data that is still blank, cleaning data from incompleteness and making data remain in a consistent state.

d. Training

At this stage the examiner will carry out training by giving an initial weight with random, an initial weight with a value of 0.1 followed by giving an initial weight of 0.3 and finally by giving a weight of 0.5, at this stage also updating the weight using forward and backward propagation.

e. Testing

At this section it will compare the results of the average error value, MSE, accuracy of the training step

2.2 Artificial Neural Network

Artificial Neural Network is an efficient computing system that adopts the analogy of an artificial neural network, in ANN there are many nodes that communicate with each other, each node communicates through a connection link associated with a weight that has information value about the input signal[8] [9] [10]. According to [11] artificial neural network is a system that works to imitate the human brain by using a computer that is always adapting to learning.

2.3 Multi-Layer Perceptron

MLP is an imitation neural network consisting of more than one perceptron. In MLP architecture it consists of an input layer for receiving signals, an output layer which makes decisions or predictions about the input, and in between, a number of hidden layers which are the actual MLP compute engine[12].

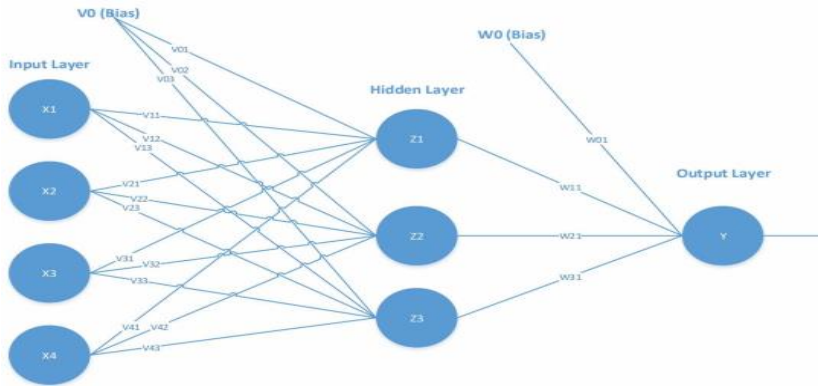

Figure 1. Artificial neural network architecture [13][14] [15]

Figure 1 illustrates a perceptron network with three layers. This network has an input layer (on the left) with four neurons, three hidden layers (in the middle) and an output layer (on the right) with one neuron.

Input Layer is The vector of the predictor variable values is presented to the input layer. In the input layer, the values are distributed to each neuron in the hidden layer. Apart from the predictor variables, there is a constant input called bias which is fed to each of the hidden layers. The bias is multiplied by the weight and added to the amount going to the neuron.

Hidden Layer is the value of each input neuron is multiplied by the weight, in this layer there is a bias which is multiplied by the resulting weight value added to produce a combined value. The output from the hidden layer is distributed to the output layer.

Building of Informatics, Technology and Science (BITS)
Volume 4, No 2, September 2022 Page: 1153−1158
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i2.2130

Output Layer is the value of each hidden layer neuron is multiplied by the weight, and the resulting weight value is added to produce a compound value.

3. RESULTS AND DISCUSSION

The result of the research can be presented in the form of tables, graphs or figures. They can be compiled with written text to build a discussion of the findings, that is about the new, the modification or the established theory.
Experiments were carried out with input data of 13 features, 20 hidden layers, a learning rate of 0.75, an error tolerance limit of 0.0001 and 1000 iterations. All experimental data will be carried out by experiment 1, experiment 2, experiment 3 and experiment 4, the difference is the weight value on the hidden layer and the weight value on the output layer.

a. The first training and testing

At this step the weight value is determined randomly. During the training process, the MSE value is 0.02 that shown figure 2, time required is 659.518 seconds with 200 data and 1000 iterations

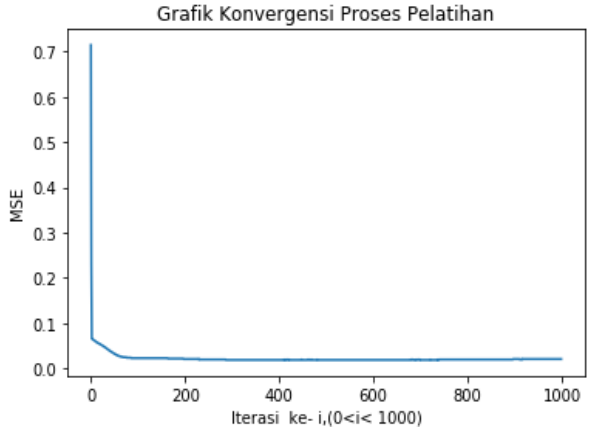

Figure 2. The first training convergence graph

Table 1. The First Test Results

Test Data	ANN Results	Original Data	Error	Accuracy(%)
1	175990122.0	176477812.0	487690.0	99.72
2	175868200.0	173673596.0	2194604.0	98.75
3	175990122.0	181842399.0	5852277.0	96.78
4	176112045.0	173429751.0	2682294.0	98.48
5	175990122.0	172454372.0	3535750.0	97.99
.....	.....	.....	.....	.....
36	176112045.0	170381690.0	5730355.0	96.75
37	175990122.0	167821319.0	8168803.0	95.36
38	176112045.0	175746277.0	365768.0	99.79
39	176112045.0	167211707.0	8900338.0	94.95
40	176112045.0	166358250.0	9753795.0	94.46

Based on the results of table 1, MSE during the testing phase for the first experiment is 0,079 with an average accuracy 96.57 and the time required is 0.048 seconds.

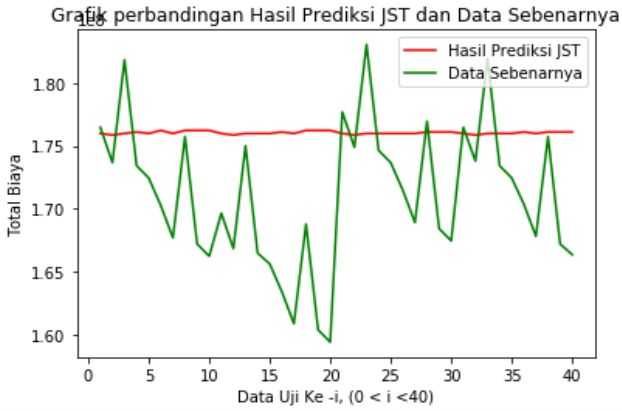

Figure 3. Graph comparison of original data with ANN results of first test

Building of Informatics, Technology and Science (BITS)
Volume 4, No 2, September 2022 Page: 1153−1158
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i2.2130

b. The second training and testing

At this step gives initial weight value is 0.1. During the training process, the MSE value is 0.019 that shown figure 4 and the time required is 650.246 seconds with 200 data and 1000 iterations.

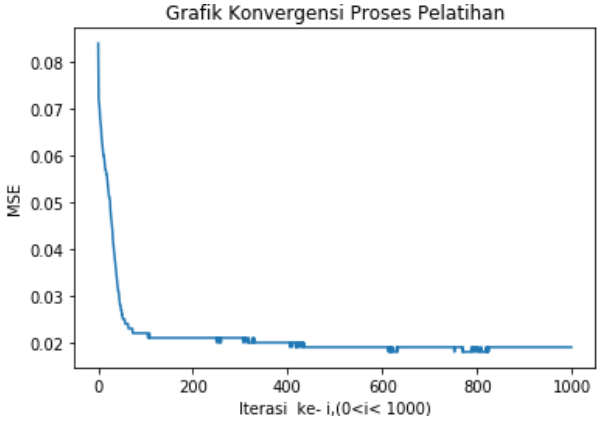

Figure 4. The second training convergence graph

Table 2. The second Test Results

Data Uji	Hasil JST	Data Asli	Error	Akurasi(%)
1	176233967.0	176477812.0	243845.0	99.86
2	176233967.0	173673596.0	2560371.0	98.55
3	176233967.0	181842399.0	5608432.0	96.92
4	176355890.0	173429751.0	2926139.0	98.34
5	176355890.0	172454372.0	3901518.0	97.79
.....	.....	.....	.....	.....
36	176355890.0	170381690.0	4023441.0	97.69
37	176355890.0	167821319.0	6583812.0	96.22
38	174527053.0	175746277.0	1219224.0	99.31
39	174527053.0	167211707.0	7315346.0	95.81
40	174527053.0	166358250.0	8168803.0	95.32

Based on the results of table 2, MSE during the testing phase for the second experiment is 0,105 with an average accuracy 96.87 and the time required is 0.072 seconds.

c. The third training and testing

At this step gives initial weight value is 0.3. During the training process, the MSE value is 0.019 that shown 5 and the time required is 685.15 seconds with 200 data and 1000 iterations

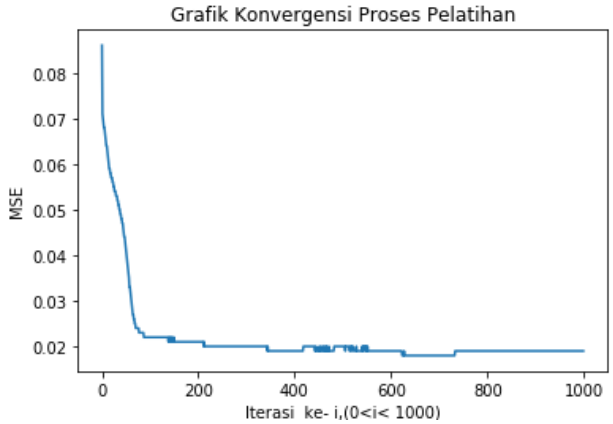

Figure 5. The third training convergence graph

Table 3. The Third Test Results

Data Uji	Hasil JST	Data Asli	Error	Akurasi(%)
1	177453191.0	176477812.0	975379.0	99.45
2	177818959.0	173673596.0	4145363.0	97.67
3	177209347.0	181842399.0	4633052.0	97.45
4	177209347.0	173429751.0	3779596.0	97.87
5	177209347.0	172454372.0	4754975.0	97.32
.....	.....	.....	.....	.....

Building of Informatics, Technology and Science (BITS)
Volume 4, No 2, September 2022 Page: 1153−1158
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i2.2130

36	175990122.0	170381690.0	5608432.0	96.81
37	175990122.0	167821319.0	8656493.0	95.09
38	175746277.0	175746277.0	0	100
39	175746277.0	167211707.0	8534570.0	95.14
40	175746277.0	166358250.0	9388027.0	94.66

Based on the results of table 3, MSE during the testing phase for the third experiment is 0,113 with average accuracy of 96.45 and the time required is 0.032 seconds.

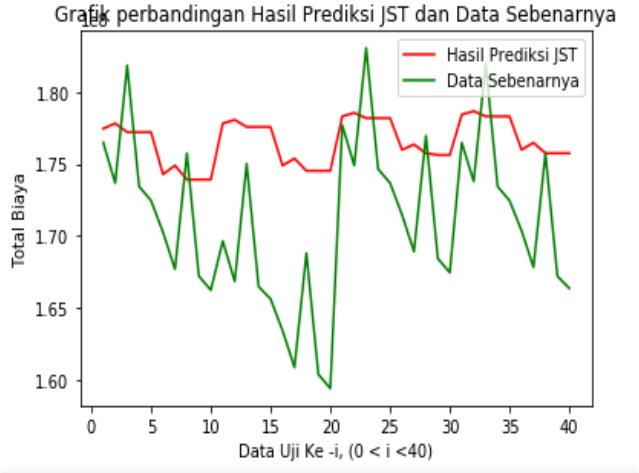

Figure 6. Graph comparison of original data with ANN results of third test

d. The fourth training and testing

At this step gives initial weight value is 0.5. During the training process, the MSE value is 0.21 that shown figure 7 and the time required is 667.816 seconds with 200 data and 1000 iterations.

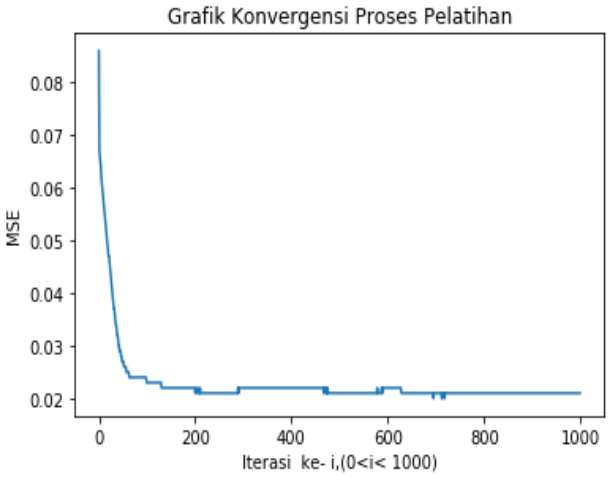

Figure 7. The fourth training convergence graph

Table 4. The Fourth Test Results

Data Uji	Hasil JST	Data Asli	Error	Akurasi(%)
1	175258588.0	176477812.0	1219224.0	99.31
2	175380510.0	173673596.0	1706914.0	99.03
3	175014743.0	181842399.0	6827656.0	96.25
4	175014743.0	173429751.0	1584992.0	99.09
5	175014743.0	172454372.0	2560371.0	98.54
.....	.....	.....	.....	.....
36	173917441.0	170381690.0	3535751.0	97.97
37	173551674.0	167821319.0	5730355.0	96.70
38	174039363.0	175746277.0	1706914.0	99.03
39	174039363.0	167211707.0	6827656.0	96.08
40	174039363.0	166358250.0	7681113.0	95.59

Based on the results of table 4, MSE during the testing phase for the fourth experiment is 0,1 with average accuracy 97.22 and the time required is 0.028 seconds.

Building of Informatics, Technology and Science (BITS)
Volume 4, No 2, September 2022 Page: 1153−1158
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i2.2130

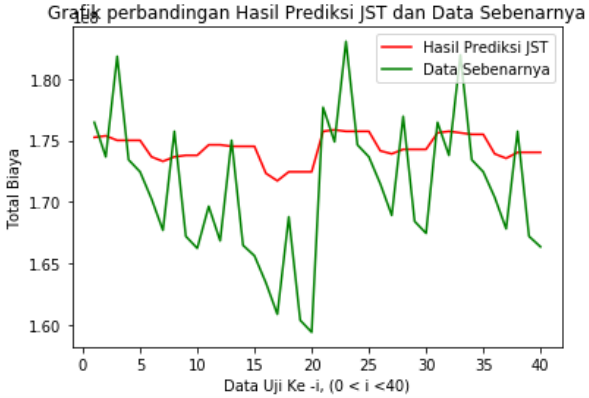

Figure 8. Graph comparison of original data with ANN results of fourth test

4. CONCLUSIONS

Based on the training and testing above, it can be concluded Random weighting is neither always good nor bad results.where is it better than third experiment with accuracy is 96, 57 but worse than second experiment and fourth experiment, it have been given initial weights 0.5 is the best results with MSE values is 0,1, mean accuracy is 97,22 and computation time is 0,028 seconds and this study, the researchers entered initial weight values into the matrix one by one. of course this is not efficient if there are many input features as well as many hidden layers so it is necessary to try random initialization methods based on certain means values.

REFERENCES

[1] M. S. Dr. Zefri, "Analisis Ketersediaan Lahan Untuk Pengembangan Perumahan Di Kecamatan Ciputat Timur Kota Tangerang Selatan," J. Ilm. Plano Krisna, vol. 13, no. 1, pp. 28–39, 2019.
[2] N. P. Sakinah, I. Cholissodin, and A. W. Widodo, "Prediksi Jumlah Permintaan Koran Menggunakan Metode Jaringan Syaraf Tiruan Backpropagation," J. Pengemb. Teknol. Inf. dan Ilmu Komput., vol. 2, no. 7, pp. 2612–2618, 2018.
[3] A. Pujianto, K. Kusrini, and A. Sunyoto, "Perancangan Sistem Pendukung Keputusan Untuk Prediksi Penerima Beasiswa Menggunakan Metode Neural Network Backpropagation," J. Teknol. Inf. dan Ilmu Komput., vol. 5, no. 2, p. 157, 2018.
[4] A. P. Windarto, M. R. Lubis, and S. Solikhun, "Implementasi JST pada Prediksi Total Laba Rugi Komprehensif Bank Umum dan Konvensional dengan Backpropagation," J. Teknol. Inf. dan Ilmu Komput., vol. 5, no. 4, p. 411, 2018.
[5] H. S. A. W. Yuli Andriani, "Prediksi Ekspor Impor Migas Ina," Regist. J. Ilm. Teknol. Sist. Inf., vol. 4, no. 1, pp. 30–40, 2018.
[6] E. Kurniawan, H. Wibawanto, and D. A. Widodo, "Implementasi Metode Backpropogation dengan Inisialisasi Bobot Nguyen Widrow untuk Peramalan Harga Saham," J. Teknol. Inf. dan Ilmu Komput., vol. 6, no. 1, p. 49, 2019.
[7] A. P. R. Pinem, N. Hidayati, and K. Kholidin, "Klasifikasi Prioritas Distrik Terhadap Ketahanan Pangan Menggunakan Metode Jaringan Syaraf Tiruan," Telematika, vol. 11, no. 2, p. 1, 2018.
[8] L. Wall et al., Artificial Intelligence with Python. 2015.
[9] H. Singh, Practical Machine Learning and Image Processing For Facial Recognition, Object Detection, and Pattern Recognition Using Python-Himanshu Singh. 2019.
[10] R. K. Maristte Award, Efficient Learning Machines. APress, 2015.
[11] S. Kusumadewi, Artificial Intelligence (Teknik dan Aplikasinya). Yogyakarta: Graha ilmu, 2003.
[12] Tutorialspoint, Artificial Intelligence With Python. 2016.
[13] H. Wadi, "Jaringan Syaraf Tiruan Backpropagation menggunakan Python GUI." TR Publisher, p. 73, 2020.
[14] Umberto Michelucci, Applied Deep Learning. Switzerland: APress, 2018.
[15] S. F. Mu'afa and N. Ulinnuha, "Perbandingan Metode Single Linkage, Complete Linkage Dan Average Linkage dalam Pengelompokan Kecamatan Berdasarkan Variabel Jenis Ternak Kabupaten …," … J. Ilm. Bid. Teknol. …, vol. 4, no. 2, 2019.