THIS PAGE IS SECURE

Advanced Science Letters

ISSN 1936-6612 (Print)

ADVANCED SCIENCE LETTERS is an international peer-reviewed journal with a very wide-ranging coverage, consolidates research activities in all areas of (1) Physical Sciences, (2) Biological Sciences, (3) Mathematical Sciences, (4) Engineering, (5) Computer and Information Sciences, and (6) Geosciences to publish original short communications, full research papers and timely brief (mini) reviews with authors photo and biography encompassing the basic and applied research and current developments in educational aspects of these scientific areas.

Publisher: American Scientific Publishers

More about this publication?

Volume 24, Number 12, December 2018

Contents   Supplementary Data

research-article

Correlation Analysis Between Error Rate of Output and Learning Rate in Backpropagation Network

Authors: Mustafidah, Hindayati [1] ; Suwarsito [2] ;
Source: Advanced Science Letters, Volume 24, Number 12, December 2018, pp. 9182-9185(4)
Publisher: American Scientific Publishers
DOI: https://doi.org/10.1166/asl.2018.12121

Abstract | References | Citations | Supplementary Data | Article Media | Metrics | Suggestions

Backpropagation network is the most widely used model of neural network. It is based on some information obtained from Levenberg-Marquardt algorithm which is a learning algorithm in backpropagation network that produces the smallest error rate. Calculation error rate is based on one of the network parameters which is the learning rate. In this research, the analysis was performed to measure the correlation between the level of network errors and learning rate. The analysis began with the development of program code to run a learning algorithm that generated error rates which were based on the variation value of 0.01 to 1.0. The results in the form of error rate and learning rate were analyzed using correlation analysis. From the analysis result, it was concluded that no significant correlation found between the level of errors generated by the network and the learning rate at 5% α.

Keywords: Backpropagation; Correlation; Error Rate; Learning Rate; Levenberg-Marquardt

Document Type: Research Article

Affiliations: **1:** Informatics Engineering, Universitas Muhammadiyah Purwokerto, Purwokerto, Central Java, 53182, Indonesia **2:** Geography Education, Universitas Muhammadiyah Purwokerto, Purwokerto, Central Java, 53182, Indonesia

Publication date: 01 December 2018

More about this publication?

RESEARCH ARTICLE

Advanced Science Letters
Vol. 24(12), 9182–9185, 2018

Correlation Analysis between Error Rate of Output and Learning Rate in Backpropagation Network

Hindayati Mustafidah[*], Suwarsito[**]

[*] *Informatics Engineering, Universitas Muhammadiyah Purwokerto*
[**] *Geography Education, Universitas Muhammadiyah Purwokerto*
Purwokerto, Central Java, Indonesia, 53182
* h.mustafidah@ump.ac.id
** ito_warsito@yahoo.co.in

Backpropagation network is the most widely used model of neural network. It is based on some information obtained from Levenberg-Marquardt algorithm which is a learning algorithm in backpropagation network that produces the smallest error rate. Calculation error rate is based on one of the network parameters which is the learning rate. In this research, the analysis was performed to measure the correlation between the level of network errors and learning rate. The analysis began with the development of program code to run a learning algorithm that generated error rates which were based on the variation value of 0.01 to 1.0. The results in the form of error rate and learning rate were analyzed using correlation analysis. From the analysis result, it was concluded that no significant correlation found between the level of errors generated by the network and the learning rate at 5% α.

Keywords: correlation, learning rate, error rate, backpropagation, Levenberg-Marquardt.

Received: 12 February 2017. Accepted: 26 July 2017

1. INTRODUCTION

Backpropagation is the most widely used type of unsupervised learning paradigm in artificial neural networks (ANN), especially in the system development to solve the problem. Systems known having used this backpropagation had been studied to detect intrusion in banking system and to estimate the longitudinal velocity field in open channel junctions.[1-3] In other cases, backpropagation as a multilayer perceptron was used in simulation of open channel bend characteristics and was further used in the prediction of flow parameters in 90° open channel bends.[4-5] The network structure in this paradigm uses more than one layer (multi-layer) to change the weights connected with neurons that exist in the hidden layer. Learning for ANN is a process where the free parameters of ANN is adapted through a process of continuous stimulation by the environment where the network belongs.[6] An ANN learns from its experience. Usual learning process includes three tasks, namely: 1) network output, 2) comparing the output with a desired target, and 3) adjusting the weights and repeating the process.

There are several training algorithm contained in backpropagation method, including *Fletcher-Reeves Update, Polak-Ribiĕre, Powell-Beale Restarts, Scaled Conjugate Gradient, Gradient Descent, Gradient Descent* with *Adaptive Learning Rate, Gradient Descent* with *Momentum and Adaptive Learning Rate, Resilent Backpropagation, BFGS, One Step Secant,* and *Levenberg-Marquardt*.[7] Each of these algorithms has procedures in accordance to the processing characteristics. As a system of learning from experience, ANN does not release results definitively in accordance to the given input which may lead the errors to occur. Calculation of mistake or error is a measure of how well the network can learn the new pattern which is easily recognizable. Error at the output of the network is the difference between actual output (current output) and the desired output. The difference is then usually determined by computation of MSE (Mean Squared Error).

In performing its function, backpropagation network is influenced by several parameters, which are the number of neurons in the input layer, the maximum allowable epoch, the magnitude of the learning rate, and the target error. Each parameter has value variations that affect the output of the network. The research conducted in stages has examined the learning algorithms in backpropagation network to find the most optimum algorithm.[8-14] From twelve learning algorithms in the backpropagation network, Levenberg-Marquardt learning algorithm is the most optimal algorithm that generates smallest error of 0,001001. This research is performed to analyze the correlation between the level of error rate networks and the learning rate (lr) of Levenberg-Marquardt algorithm. The expected benefit of this research includes an overview of the estimated level of error made using the value of the certain of learning rate (lr).

Adv. Sci. Lett. 24(12), 9182–9185, 2018

2. MATERIALS AND METHOD

This study began with the development of a computer program using MATLAB programming language to run a learning algorithm. The program input was a form of random data with 5 neurons, 10 neurons, and 15 neurons structure in the input layer, and one neuron in the output layer. The variables used were as follows:

- Control Variables (values vary): the input data, the maximum epoch (= 10,000), the target error (= 10^{-3}).
- Independent Variable (values vary): learning rate (lr)
- Dependent Variable: error rates generated by Levenberg-Marquardt algorithm.

In order to achieve error, the program used MSE. MSE is a function that measures the network performance based on the average of squared error as the following equation 1.

$$MSE = \frac{\sum_p \sum_j (T_{jp} - X_{jp})^2}{n_p n_o} \qquad (1)$$

T_{jp} = the output value of the neural network
X_{jp} = the desired target value for each output
n_p = the total number of patterns
n_o = the number of outputs

The analysis was aimed to determine the relationship between the level of the resulting error of networks and learning rate of the Levenberg-Marquardt learning algorithm, and was conducted in these following stages:[15]

1) Determining hypothesis
   H_0 : there is no correlation between the level of the resulting error rate of the network and learning rate
   H_1 : there is a correlation between the level of the resulting error rate of the network and learning rate
2) determining the value of alpha (α)
3) determining the testing equipment

   Test equipment used in this case was a F-test as in equation 2.

$$F_{cal} = \frac{MST}{MSE} \sim F_{k-1,\, N-k} \qquad (2)$$

where MST = *Mean Square of Treatment*
MSE = *Mean Square of Error*

4) drawing conclusions

Conclusions drawn by the significant value gained (sig.) with the provisions of Ho was rejected if sig. < α.

3. RESULTS AND DISCUSSION

One of supervised learning paradigm types in artificial neural networks (ANN) is backpropagation as presented in Figure 1.

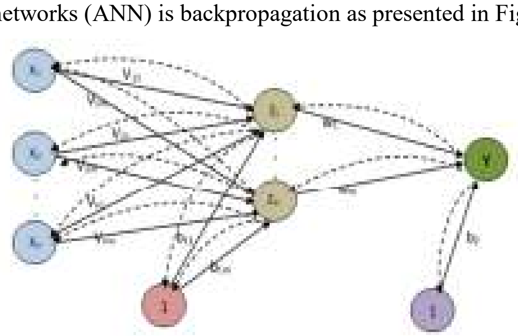

Fig 1. Learning Process in ANN

In Figure 1, the network consists of n units (neurons) in the input layer i.e. $x_1, x_2, ..., x_n$, a hidden layer with m neurons i.e. $z_1, ..., z_m$, and 1 unit of neuron in the output layer y. The weights that connect the neurons in the input layer to the neurons in the hidden layer are v_{11} (weights connecting x_1 to z_1), v_{12} (weights connecting x_1 to z_2), ..., v_{1m} (weights connecting x_1 to z_m), v_{21} (weights x_2 to Z_1), v_{22} (weights connecting x_2 to z_2), and v_{2m} (weights connecting x_2 to z_m). The weights are symbolized by v_{ij} (i.e. the weights that connect the i-th input neuron to the j-th neuron in the hidden layer). The bias weights leading to the first and second neurons of the hidden layer are b_{11} (the bias weight connecting to z_1), b_{1m} (the bias weight connected to z_m). The weights that connect the neurons in the hidden layer are $z_1, ..., z_m$ while in the output layer (y) are $w_1, ..., w_m$. The bias weight of b_2 is the bias weight that leads to the neuron in the output layer. The activation function used in the connection of the input layer and the hidden layer and in the connection of the hidden layers and the output layer is a deferred activation function.

As shown in Figure 1, this research also used 1 output neuron. To produce output, ANN performed a learning process working on experiences. The learning process in ANN is presented in Figure 2.

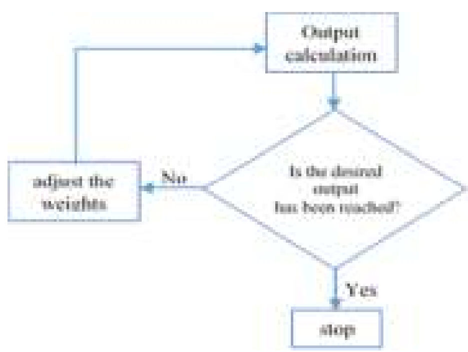

Figure 2. Learning Process in ANN

The input data network was randomized betweem each 5 neurons input, 10 neurons input, and 15 neurons input. In this case, the data of 5 neurons input was presented as in Table 1. The independent variables in this study were the learning rate with the value of 0.01; 0.05; 0.1; 0.2; 0.3; 0.4; 0.5; 0.6; 0.7; 0.8; 0.9; 1.0. The computer program was coded with MATLAB (Figure 3) and compiled 20 times for each learning rate. Program output was in the form of an average of error (MSE) of Levenberg-Marquardt algorithm performance for each variation in the learning rate (lr) as shown in Table 2.

Table 1. Network input and 5 neurons target data

X_1	X_2	X_3	X_4	X_5	Y
9.5013	7.6210	6.1543	4.0571	0.5789	2.0277
2.3114	4.5647	7.9194	9.3547	3.5287	1.9872
6.0684	0.1850	9.2181	9.1690	8.1317	6.0379
4.8598	8.2141	7.3821	4.1027	0.0986	2.7219
8.9130	4.4470	1.7627	8.9365	1.3889	1.9881

```
clc;
disp('=============================================')
for i=1:20, fprintf('\nIterate: %d\n', i)
clear; data_5; //can be replaced with data_10 or data_15
P = P'; T

%Input & Target
net = newff(minmax(P), [10 1], {'tansig' 'purelin'}, 'trainlm');

%first weights
FirstWeightInput = net.IW{1,1}; FirstWeightBiasInput = net.b{1,1};
FirstWeightLayer = net.LW{2,1}; FirstWeightBiasLayer = net.b{2,1};

%set training function
net.trainParam.epochs = 10000; net.trainParam.goal = 1e-3;
net.trainParam.lr = Learning_rate; net.trainParam.mc = 0.3;
net.trainParam.min_grad = 1e-10; net.trainParam.show = 500;

% training function
net = train(net, P, T);

%end weights
EndWeightInput = net.IW{1,1}; EndWeightBiasInput = net.b{1,1};
EndWeightLayer = net.LW{2,1}; EndWeightBiasLayer = net.b{2,1};

%simulation
a = sim(net, P) fprintf('\n'); end
```

Figure 3. Source code of ANN computer program

Table 2. MSE Results of Levenberg-Marquardt learning algorithm for each random data and lr

No.	5 random data		10 random data		15 random data	
	lr	MSE	lr	MSE	lr	MSE
1	0.01	.000198685800	0.01	.000185121492	0.01	.000284530950
2	0.05	.000179383180	0.05	.025779882093	0.05	.000248148088
3	0.1	.000186550202	0.1	.000126952176	0.1	.000168452917
4	0.2	.000147757748	0.2	.000185485482	0.2	.000276679704
5	0.3	.000198685800	0.3	.009731951112	0.3	.000246095458
6	0.4	.000192124870	0.4	.000265669001	0.4	.000249492183
7	0.5	.000203676863	0.5	.009751811074	0.5	.000246095458
8	0.6	.000109168261	0.6	.000185485482	0.6	.000249492183
9	0.7	.000210012926	0.7	.009731951112	0.7	.000163407430
10	0.8	.000198685800	0.8	.000265695786	0.8	.000246095458
11	0.9	.000116184938	0.9	.000126209131	0.9	.000249492183
12	1	.000203676863	1	.000294526060	1	.000163407430

The research data were assumed to be normal and homogenous as they came from the same population. Correlation analysis was performed using F-test with df1 = 1 and df2 = 718 (Table 3). The calculation of the F value was performed using SPSS.

Table 3. Correlations

		lr	MSE
lr	Pearson Correlation	1	-.036
	Sig. (2-tailed)		.336
	N	720	720
MSE	Pearson Correlation	-.036	1
	Sig. (2-tailed)	.336	
	N	720	720

Using Pearson Correlation, a negative correlation was found between the learning rate (lr) and the error generated by the network (MSE) is -0.036. This means that the higher lr will result in smaller resulting error. Otherwise, based on the analysis, the resulting significance value (0.336) was larger than α (5% = 0.05) hence Ho was accepted. This, in turn, means that there was no significant correlation between the error generated by network and learning rate (lr). The value of F-calc obtained after consultation with F-table (0.95, 1, 718) = 3.854443 was 0.928. It was much smaller hence Ho was accepted. This means that there was no significant correlation. According to the Pearson Correlation (r) value of -0.036, the correlation was very small and negative. This means that the higher learning rate will generate smaller error.

Adv. Sci. Lett. 24(12), 9182–9185, 2018

The correlation analysis results were consistent with the regression analysis results using Curve-Fitting as shown in Table 4.

Table 4. Regression Table

Model	F-calc	Sig.	R-value
Linear	.928	.336	0.036
Logaritmic	.485	.486	.026
Quadratic	.465	.628	.036
Exponential	.065	.948	.002

Table 4 showed that the expected **linear** model was best suited for regression analysis in the case of the smallest significant value of the four models. However, the value of F obtained was 0.928, smaller than F-table (0.95, 1, 718) = 3.854443. This means that Ho was accepted and there was no significant influence of the level of learning rate (lr) on the rate of error. The values produced from Table 5 and 6 formulated the regression equation (3).

$$MSE = 0.003 - 0.002 \text{ lr} \quad (3)$$

Table 5. ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Regression	.000	1	.000	.928	.336
Residual	.367	718	.001		
Total	.368	719			

The independent variable was lr.

Table 6. Coefficients

	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
lr	-.002	.003	-.036	-.963	.336
(Constant)	.003	.001		1.956	.051

Significant value of the constant variable was 0.051 while the value of the independent variable coefficient was 0.336. The second significance value was greater than α (5%) hence the regression equation was not significant.

4. CONCLUSIONS

The results of the correlation analysis between the level of the resulting error rate networks and learning rate at the level of 5% α showed no significant correlation. The results of this correlation also affected the degree of influence of the two variables. Based on the regression analysis results, there were no significant effect between the two variables at the level of 5% α.

Acknowledgements: Authors would like to thank DP2M - The Ministry of Research, Technology, and Directorate General of Higher Education (RISTEKDIKTI) through KOPERTIS Region VI and Research Board Universitas Muhammadiyah Purwokerto for providing funds and facilities in the implementation of this research.

References and Notes

1. Harjono & D. Aryanto, Application of Artificial Neural Networks to Predict Student Achievement Study. *SAINTEK* ISSN: 1411-2558, Vol. 5 No. 2. (2009).
2. A.B. Adetunji, A.Q. Ayinde & C.O. Akanbi, Application of Neural Network to Detect Intrusion in Banking System, *AMERICAN JOURNAL OF SCIENTIFIC AND INDUSTRIAL RESEARCH*, 5(2) (2014): 53-59.
3. H. Zaji & H. Bonakdari, Application of artificial neural network and genetic programming models for estimating the longitudinal velocity field in open channel junctions, *Flow Measurement and Instrumentation*, 41(2015): 81–89.
4. A. Gholami, H. Bonakdari, A. H. Zaji, S. A. Fenjan & A.A. Akhtari, Design of modified structure multi-layer perceptron networks based on decision trees for the prediction of flow parameters in 90° open-channel bends, *Engineering Applications of Computational Fluid Mechanics*, 10:1 (2016), 193-208, DOI: 10.1080/19942060.2015.1128358.
5. A. Gholami, H. Bonakdari, A. H. Zaji, S. A. Fenjan & A.A. Akhtari, Simulation of open channel bend characteristics using computational fluid dynamics and artificial neural networks, *Engineering Applications of Computational Fluid Mechanics*, 9:1 (2015), 355-369, DOI: 10.1080/19942060.2015.1033808.
6. A. Desiani & M. Arhami, *Konsep Kecerdasan Buatan*, Andi Offset, Yogyakarta (2006).
7. S. Kusumadewi, *Membangun Jaringan Syaraf Tiruan Menggunakan MATLAB & EXCEL LINK,* Graha Ilmu, Yogyakarta (2004).
8. F. Wibowo, S. Sugiyanto, & H. Mustafidah, Tingkat Ketelitian Pengenalan Pola Data pada Algoritma Pelatihan Perbaikan Metode *Batch Mode* dalam Jaringan Syaraf Tiruan. *Jurnal Ilmiah Nasional "JUITA"* ISSN 2086-9398, Volume II No.4, November (2013), pp. 259 – 264.
9. H. Mustafidah, D. Aryanto, & D.K. Hakim, Uji Optimalisasi Algoritma Pelatihan *Conjugate Gradient* pada Jaringan Syaraf Tiruan. *Proceeding SENATEK*, ISBN: 978-602-14355-0-2 pp. B.9-1, Universitas Muhammadiyah Purwokerto, Indonesia (2013).
10. H. Mustafidah, D.K. Hakim, & S. Sugiyanto, 2013b. Tingkat Keoptimalan Algoritma Pelatihan pada Jaringan Syaraf Tiruan (Studi Kasus Prediksi Prestasi Belajar Mahasiswa). *JUITA* ISSN: 2086-9398 Vol. II Nomor 3, (2013), pp. 159 – 166.
11. H. Mustafidah, S. Hartati, S., R. Wardoyo, & A. Harjoko, Prediction of Test Items Validity Using Artificial Neural Network. *Proceeding International Conference on Education, Technology, and Science (NETS) 2013*, "Improving The Quality Of Education To Face The Impact Of Technology". December 28th, Universitas Muhammadiyah Purwokerto, Indonesia (2013).
12. H. Mustafidah & Suwarsito, Error Rate Testing of Training Algorithm in B Back Propagation Network, *International Journal of Soft Computing and Engineering (IJSCE),* ISSN: 2231-2307, Volume-5 Issue-4, (2015), pp. 46 – 50.
13. H. Mustafidah & Suwarsito, Model Parameter Jaringan Syaraf Tiruan untuk Pemilihan Algoritma Pelatihan Jaringan Backpropagation yang Paling Optimal, *Research Report,* Universitas Muhammadiyah Purwokerto, Purwokerto (2015).
14. H. Mustafidah & Suwarsito, Uji Keoptimalan Algoritma Pelatihan pada Jaringan Syaraf Tiruan, *Prosiding Seminar Nasional SENATKOM 2015,* ISSN: 2460-4690, pp. 243-248. Padang Indonesia (2015).
15. T. Taniredja & H. Mustafidah, *Penelitian Kuantitatif (Sebuah Pengantar).* Alfabeta. Bandung (2011).