Certificate

as awarded to

Windra Swastika

as

a Presenter

at The 4th International Symposium on BioMathematics (SYMOMATH 2016)

at Hasanuddin University, Makassar, Indonesia

on October 7-9, 2016

Organized by:

INSTITUT TEKNOLOGI BANDUNG 1920

HIMPUNAN MATEMATIKA INDONESIA 1976

IBMS Indonesian Biomathematical Society

SYMOMATH 2016

Prof. Dr. Syamsuddin Toaha, M.Sc.
Chair of Symomath 2016

Prof. Dr. Dwia Aries Tina Pulubuhu, MA
Rector of Hasanuddin University

AIP Conference Proceedings ×

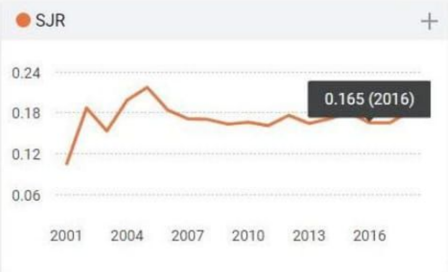
https://www.scimagojr.com/journalsearch.php?q=26916&tip=sid&clean=0

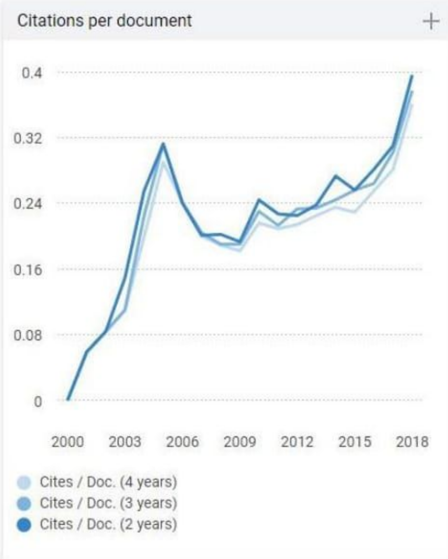
AIP Conference Proceedings

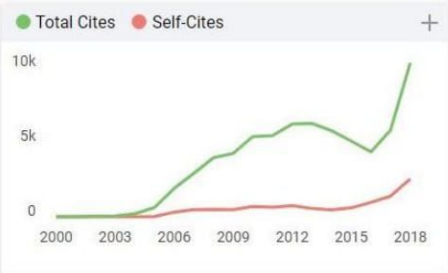
Country	United States - SIR Ranking of United States
Subject Area and Category	Physics and Astronomy Physics and Astronomy (miscellaneous)
Publisher	
Publication type	Conferences and Proceedings
ISSN	00001984, 00002005, 00001983
Coverage	1983-1984, 2005-ongoing
Scope	AIP Conference Proceedings report findings presented at many of the most important scientific meetings around the world. Published proceedings are valuable as topical status reports providing quick access to information before it appears in the traditional journal literature.

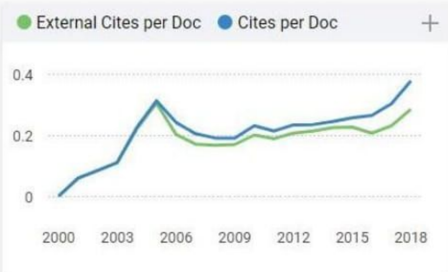
60

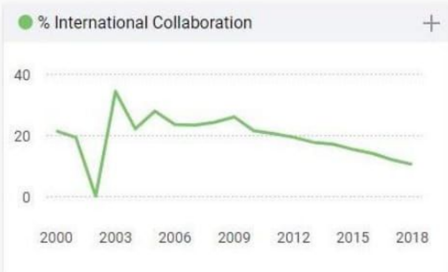
H Index

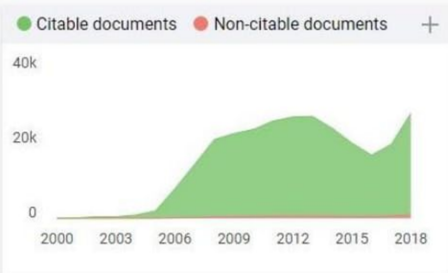


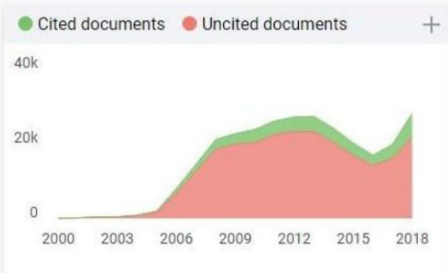












← Show this widget in your own website

Just copy the code below and paste within your html code:

<a href="https://www.scimag

Quickprop method to speed up learning process of Artificial Neural Network in money's nominal value recognition case

Windra Swastika

Citation: AIP Conference Proceedings **1825**, 020022 (2017); doi: 10.1063/1.4978991
View online: http://dx.doi.org/10.1063/1.4978991
View Table of Contents: http://aip.scitation.org/toc/apc/1825/1
Published by the American Institute of Physics

Quickprop Method to Speed up Learning Process of Artificial Neural Network in Money's Nominal Value Recognition Case

Windra Swastika

Fakultas Sains dan Teknologi,
Universitas Ma Chung
Villa Puncak Tidar N1, Malang, Indonesia

windra.swastika@machung.ac.id

Abstract. A money's nominal value recognition system has been developed using Artificial Neural Network (ANN). ANN with Back Propagation has one disadvantage. The learning process is very slow (or never reach the target) in the case of large number of iteration, weight and samples. One way to speed up the learning process is using Quickprop method. Quickprop method is based on Newton's method and able to speed up the learning process by assuming that the weight adjustment (E) is a parabolic function. The goal is to minimize the error gradient (E'). In our system, we use 5 types of money's nominal value, i.e. 1,000 IDR, 2,000 IDR, 5,000 IDR, 10,000 IDR and 50,000 IDR. One of the surface of each nominal were scanned and digitally processed. There are 40 patterns to be used as training set in ANN system. The effectiveness of Quickprop method in the ANN system was validated by 2 factors, (1) number of iterations required to reach error below 0.1; and (2) the accuracy to predict nominal values based on the input. Our results shows that the use of Quickprop method is successfully reduce the learning process compared to Back Propagation method. For 40 input patterns, Quickprop method successfully reached error below 0.1 for only 20 iterations, while Back Propagation method required 2000 iterations. The prediction accuracy for both method is higher than 90%.

INTRODUCTION

In the banking sector, there are many financial processes such as withdrawal, transferring, clearing, and depositing money. With the development of the banking system, it appears the need to automate the processes of the transaction. For the automation of processes such as making cash transactions and transfers, can be done through Automated Teller machine (ATM). However, for the automation process of depositing money, still not common. This automation process, can be done through a system that can recognize the nominal currency.

In other cases, the currency recognition system can also apply to vending machines for drinks or food. In the last few decades, food or drink vending machines recognize the nominal currency by detecting the weight of the coins inserted. But with the smaller coins and paper currency use is increasingly common, then a system that is able to recognize nominal paper currency is required.

Recognition system of paper currency is good for the purpose of automation as well as depositing money in the vending machine, can be supported by computer by performing pattern recognition. In theory, pattern recognition can be regarded as one of the branches of computer science focused on the finding of patterns in data that shows specific information. The data used for pattern recognition can be in the form of images, sounds, text, and moving images (videos). One of the purpose of using pattern recognition techniques is to make the data able to extract the information. The capabilities to provide the information depends on the quality and quantity of the data itself. In the case of pattern recognition nominal value of the currency, the data used is data in the form of images.

There are many approaches to implement pattern recognition. Three widely used approaches are statistical, syntactic and ANN. Statistical approach represents each pattern in a certain features and views a point on a space

Symposium on Biomathematics (SYMOMATH 2016)
AIP Conf. Proc. 1825, 020022-1–020022-6; doi: 10.1063/1.4978991
Published by AIP Publishing. 978-0-7354-1493-8/$30.00

dimension [1]. For syntactic approach [2], the pattern sorted based on structural similarity measure. While the approach of using ANN makes it possible to create systems that are adaptive [3]. ANN is an adaptive system and capable to “learn” from the trained data which is certainly an advantage compared with statistical methods and syntax.

The main problems that occur to recognize a certain pattern and feeding into a process of ANN is how the data acquisition process is done so as to produce a numeric data that are representative and consistent with the given sample. In order to recognize the pattern, ANN requires a learning method. The learning process usually involves three activities, counting the output, compare the output with the desired target, and weights are adjusted accordingly. The process is repeated until certain conditions are met. Learning process begins by defining weights with certain values (or randomly). The weights is then adjusted in the right direction so that the difference between actual output and desired output is minimized [3].

With the growing of banking and industry, the automation for the introduction of nominal currency paper will also be needed for transaction processing can be done more quickly and without human intervention. Given all the above factors, it is necessary to conduct research in the nominal introduction of paper currency using Artificial Neural Networks. In this study will be developed as a learning method in the introduction of nominal currency, the method quickprop, which is expected to have a better learning speed compared to standard Back Propagation methods.

MATERIALS AND METHODS

An image (sample) that will be identified using Artificial Neural Networks, must be processed to extract certain features as input. The sample should be well represented in the form of numerical data. Thus, we need a method that can extract the data characteristics of each sample in a consistent basis.

Image Acquisition and Preprocessing

Samples are recorded in bitmap format file with a size of 100 x 40 pixels. Preprocessing image consists of three stages:

- First stage is grayscale process. Grayscale process used to convert a color image into a black and white image that has a degree of grayish. Basically, this process is done by leveling the pixel values of the three RGB values into one value. Percentages used in this system for a pixel that has the RGB format are 29.9%, 58.7%, and 11.4% for red, green and blue.
- Second stage is histogram equalization. The purpose of the stage is to spread the intensity values of the image to obtain even intensity distribution. Histogram equalization process on this application is calculated with the following steps: G is the number of grayscale levels and n_k is the number of pixels with grayscale level g_k appear in the image, n is a whole number of pixels in the image that is 4000. Then the changes that occur for each grayscale level is shown in the eq (1),

$$s_k = T(g_k) = \sum_{j=0}^{k} \frac{n_j}{n} \quad (1)$$

  where $0 \leq g_k \leq 1$ and k=0, 1, 2, ... G-1.
- Last stage is binarization. It is used to convert grayscale images into black and white image (binary) with very high contrast levels. All pixels lighter than the threshold value will be changed to white, otherwise will be changed to black. In this system, the threshold is set 100.

Methods

A neural network used in the system using multilayer perceptron. Architectural and learning method used is an adaptation of the research in neural network with slight modifications. The composition of the units of this network is shown in Figure 1.

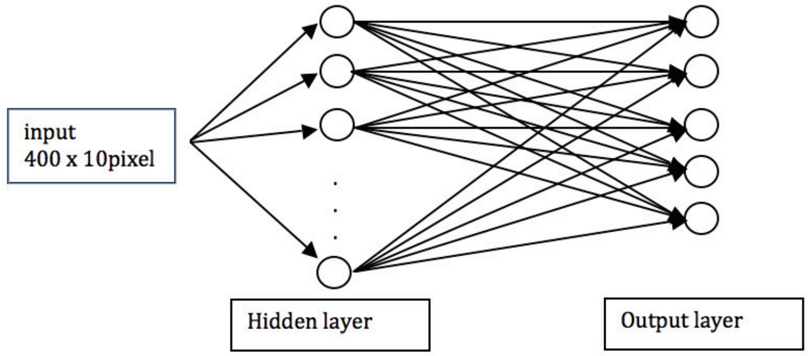


FIGURE 1. ANN architecture for money nominal recognition

Input layer consists of 4000 input unit, which is the result of binarization (0 and 1) from the preprocessing stage. The output layer consists of 5 units of output each output value between 0 and 1. Each unit represents the nominal currency 1,000 IDR, 5,000 IDR, 10,000 IDR, 20,000 IDR and 50,000 IDR. In this study, we use sigmoid activation function as shown in eq. 2.

$$f(x) = \frac{1}{(1+e^{-x})} \tag{2}$$

When x is increased, the result will be closed to 1 and when x is decreased, the result will be closed to 0. In nominal currency recognition system, the number of hidden layer consists of 20 units. There are no specific rules to determine the number of hidden layer in ANN. In our study, the system is divided into two stages:

1. The training process used to set training parameters and generate value weights of ANN.
2. Test the image. Used to determine the nominal currency based on the weight that was gained from the training process.

The process of the detector are as follows:

1. Set the level histogram and do the process grayscale.
2. Noise removal.
3. Image normalization.
4. Recognition process using ANN with Back Propagation learning. The output consists of 5 bits, each of which represents a nominal of existing pattern, namely 1,000 IDR, 5,000 IDR, 10,000 IDR, 20,000 IDR and 50,000 IDR.

Learning Process Improvement

As seen in Figure 1, the architecture of nominal currency recognition system is multilayer feedforward network architecture using the Back Propagation method. There are three layers in the architecture of multilayer feedforward, input, hidden and output layers. Input layer consists of 4,000 units derived from image data that has undergone normalization. The entire unit of input, forwarded to the hidden layer using sigmoid activation function, so that the hidden layer output there will be 20. The result from the hidden layer is forwarded to the output layer that consists of 5 units and compared with the desired target. The difference between the desired target and actual output was propagated to the previous layer to adjust the weights. This method is called the Back Propagation method. Mathematical equations of back propagation method are as follows:

1. Perform the calculation of the input layer to the hidden layer,

$$v_j^{(l)} = \sum_{i=0}^{P} w_{ji}^{(l)} . x_i^{(l-1)} \tag{3}$$

where
j: neuron in the hidden layer
l: hidden layer
x: input of current layer
v: output of current layer
w: weight

p: number of neuron

2. The output of the process to-1, calculated using sigmoid activation function with the general equation:

$$y_j^{(l)} = \frac{1}{1+\exp(-v_j^{(l)})} \tag{4}$$

3. Using the same formula (3) and (4), output of hidden layer is calculated.
4. The difference between the output from step 3 and the desired target is expected to be very small. Calculation error (cost function) using the formula:

$$E = \frac{1}{2}\sum_j^N (d_j - o_j)^2 \tag{5}$$

where
d_j : expected output
o_j: output from ANN

5. Changes in weight (ΔW_{ji}) is proportional to the negative gradient, written as follows:

$$\Delta W_{ji} = -\eta \frac{\partial E}{\partial w_{ji}} \tag{6}$$

Proportional factor η is the learning rate that defines the depth of iteration steps. If the learning rate is big then changes in weight becomes large. On the other hand, small learning rate can result the slow changes of weights. To calculate the gradient ∂E / (∂w_{ji}), chain rule can be used:

$$\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial o_j}\frac{\partial o_j}{\partial v_j}\frac{\partial v_j}{\partial w_{ji}} \tag{7}$$

where
$v_j = \sum_i w_{ji} x_i$
o_j = transfer function from v_j

Partial differential equations (6) is divided into several following equations:

$$\frac{\partial E}{\partial w_{ji}} = \frac{\partial(\frac{1}{2}\sum_j^N (d_j - o_j)^2\ )}{\partial o_j} = -(d_j - o_j) \tag{8}$$

$$\frac{\partial o_j}{\partial v_j} = \frac{\partial f(v_j)}{\partial v_j} = f'(v_j) \tag{9}$$

$$\frac{\partial v_j}{\partial w_{ji}} = \frac{\partial(\sum_i w_{ji} x_i)}{\partial w_{ji}} = x_i \tag{10}$$

So the end result of the above equation can be written:

$$\frac{\partial E}{\partial w_{ji}} = -(d_j - o_j)\ f'(v_j) x_i \tag{11}$$

Since the derivative of sigmoid function f (x) is f '(x) = x (1-x), we can rewrite the eq. 11 to:

$$\frac{\partial E}{\partial w_{ji}} = -(d_j - o_j) v_j (1 - v_j) x_i \tag{12}$$

6. Changes in the of both the input and hidden layer is calculated using the formula:

$$w_{ji}(n+1) = w_{ji}(n) + \Delta w_{ji} \tag{13}$$

Improvements made on the Back Propagation method is the learning method during the weights adjustment. Some methods proposed adding a parameter momentum [4, 5, 6]. The addition of momentum allows to accelerate weight changes during the current error difference with previous error is very small. With the addition of momentum, the formula to change the weight becomes:

$$w_{ji}(n+1) = w_{ji}(n) + \Delta w_{ji} + \alpha w_{ji}(n-1) \tag{14}$$

where α is a constant momentum close to 1.

A method to accelerate the learning is to use a derivative of second order to get the pattern from the surface of any errors found. Unfortunately, the second order derivative calculation would lead to an increase in the time required in the learning process. To avoid this, quickprop is used by assuming that the error function curve surface of each weight is quadratic (parabola that opens upward), and the gradient of the curve of error for a given weight is independent to each other. This concept was first proposed in [7] and successfully implemented in [8, 9, 10].

In this way, Quickprop method is able to overcome the weaknesses in the Back propagation method and enables faster evaluation. The learning process is done exactly the same as standard Back Propagation method, except in step changes in weight. Quickprop method of storing an error in the derivative previous epoch to be used in the estimation error curve. The formula to make changes by using:

$$w_{ji}(n+1) = w_{ji}(n) + \Delta w_{ji} + \frac{\frac{\partial E}{\partial w_{ji}(n)}}{\frac{\partial E}{\partial w_{ji}(n-1)} - \frac{\partial E}{\partial w_{ji}(n)}} \Delta w_{ji}(n-1) \quad (15)$$

RESULTS AND ANALYSIS

In the first experiment, we compare the training time of quickprop method with standard back propagation without and with momentum. The number of input patterns to train ANN are 5, 20, and 40 patterns for each nominal. The threshold error is <0.01 with learning rate 0.1 the results are shown in the Table 1.

TABLE 1. Comparison time required to train ANN.

input patterns	error/learning rate	Training time		
		Standard Back Prop.	Back Prop. With momentum	Quickprop
5	<0.01 / 0.1	640 sec	380 sec	23 sec
20	<0.01 / 0.1	> 240 min	> 120 min	110 sec
40	<0.01 / 0.1	> 300 min	> 240 min	250 sec

In term of time required to train ANN, quickprop outperform the other two methods. In the second experiment, we are interested in the effect of learning rate applied to the ANN system. Two different learning rates, 0.1 and 0.3 were used to compare the accuracy and the number of iteration. The number of patterns use for training data set for this experiment are 5, 20, and 40 for each unit and 10 samples of each nominal were provided as the input after training. The samples represent various condition of money.

Table 2 and 3 show the comparisons of quickprop method with two backprop methods in term of number of iteration, error, and accuracy.

TABLE 2. Comparison of number of iteration, error and accuracy using learning rate 0.1.

	input patterns	# iterations	Error	Accuracy
Standard back propagation	5	8000	0.07	70%
	20	8000	0.07	86%
	40	8000	0.07	94%
Back propagation with momentum	5	1000	0.07	70%
	20	1300	0.07	86%
	40	1500	0.07	92%
Quickprop	5	15	9.42-05	40%
	20	20	6.91E-05	74%
	40	11	8.393E-05	96%

TABLE 3. Comparison of number of iteration, error and accuracy using learning rate 0.3.

	input patterns	# iterations	Error	Accuracy
Standard back propagation	5	2000	0.07	72%
	20	2000	0.07	86%
	40	2000	0.07	96%
Back propagation with monentum	5	590	0.07	72%
	20	1200	0.07	86%
	40	1800	0.07	96%
Quickprop	5	12	8.272E-05	36%
	20	25	9.993E-05	76%
	40	12	9.993E-05	96%

As shown in the tables, higher learning rate of standard back propagation reduced more than half of the number of iteration with the same accuracy. However, compared to back propagation with momentum, the number of iteration using standard back propagation is greater. The number of iteration of back propagation with momentum is also reduced almost half with higher 0.3 learning rate. On the other hand, the number of iteration of quickprop method is already very small in the both learning rates. It is worth noting that the accuracy of quickprop method using only 5 input patters were very low (36%), while the other two methods achieved more than 70%. To achieve higher accuracy, quickprop method requires more input patterns.

CONCLUSION

ANN using standard Back Propagation learning method has several disadvantages during the iterations. The error curve is often stuck in the local minima. This tends to occur on systems that have a ratio of the number of ANN input layer unit with a number of hidden layer unit is quite large. This problem can be reduced by increasing the value of learning rate and the risk of changes in weight becomes very large and the error value cannot reach the minimum. One way to overcome the issue is the addition of constant momentum at the time of updating the weights. This momentum parameter is into the general equation as Learning term, thus the efficiency of the weight adjustment becomes very high. Momentum parameter on the Back propagation method can improve the standard of learning up to 10 times compared with the standard back propagation learning parameter (training set, matrix load, and learning rate).

This study shows that the use of standard back propagation method to recognize nominal value of money has accuracy of 70% and the time to learn more than four hours to as much as 40 patterns. The addition of momentum can speed up by 20% without reducing the detection accuracy. Quickprop method significantly improves the speed of learning to be only 250 seconds with the same degree of accuracy.

REFERENCES

1. Smolinska, Agnieszka, et al. "NMR and pattern recognition methods in metabolomics: from data acquisition to biomarker discovery: a review." *Analytica chimica acta* 750 (2012): 82-97.
2. Flasiński, Mariusz, and Janusz Jurek. "Fundamental methodological issues of syntactic pattern recognition." Pattern analysis and applications 17.3 (2014): 465-480.
3. Yadav, Jitender Singh, Mohit Yadav, and Ankit Jain. "Artificial neural network." *International Journal of Scientific Research and Education* 1.06 (2014).
4. Amato, Filippo, et al. "Artificial neural networks in medical diagnosis." *Journal of applied biomedicine* 11.2 (2013): 47-58.
5. Karayiannis, Nicolaos, and Anastasios N. Venetsanopoulos. *Artificial neural networks: learning algorithms, performance evaluation, and applications*. Vol. 209. Springer Science & Business Media (2013).
6. Graupe, Daniel. *Principles of artificial neural networks*. Vol. 7. World Scientific (2013).
7. Fahlman, Scott E. "Faster-learning variations on back-propagation: An empirical study." (1988): 10-12.
8. Riedmiller, Martin, and Heinrich Braun. "A direct adaptive method for faster backpropagation learning: The RPROP algorithm." *Neural Networks, 1993. IEEE International Conference On*. IEEE, 1993.
9. Ramasubramanian, P., and A. Kannan. "Multi-Agent based Quickprop Neural Network Short-term Forecasting Framework for Database Intrusion Prediction System." (2014).
10. Gazzaz, Nabeel M., et al. "Artificial neural network modeling of the water quality index for Kinta River (Malaysia) using water quality variables as predictors." *Marine Pollution Bulletin* 64.11 (2012): 2409-2420.