stock price

by Dwiarso Utomo

STOCK PRICE PREDICTION USING BACK PROPAGATION NEURAL NETWORK BASED ON GRADIENT DESCENT WITH MOMENTUM AND ADAPTIVE LEARNING RATE

DWIARSO UTOMO

Department of Economic and Business, Department of Computer Science, Dian Nuswantoro University, Semarang, Central Java, Indonesia

Tel: +62 24 3517261;

Email: **dwiarso.utomo@dsn.dinus.ac.id**

PUJIONO

Department of Economic and Business, Department of Computer Science, Dian Nuswantoro University, Semarang, Central Java, Indonesia

MOCH ARIEF SOELEMAN

Department of Economic and Business, Department of Computer Science, Dian Nuswantoro University, Semarang, Central Java, Indonesia

Abstract

Accurate financial predictions are challenging and attractive to individual investors and corporations. Paper proposes a gradient-based back propagation neural network approach to improve optimization in stock price predictions. The use of gradient descent in BPNN method aims to determine the parameter of learning rate, training cycle adaptively so as to get the best value in the process of stock data

training in order to obtain accuracy in prediction. To test BPNN method, mean square error is used to prediction result and data reality. The smallest MSE value shows better results compared to larger MSE value in predictions.

Keywords: **Neural Network Back Propagation; Gradient Descent; Prediction; Stock**

© Dwiarso Utomo, 2017

INTRODUCTION

The capital market is an organized financial system consisting of commercial banks, financial intermediaries and all securities circulating in the community. One of the benefits of capital markets creates an opportunity for people to participate in economic activities, especially in investing. One of the assets for investment is stock. Stock is securities issued by a company. Revenues earned from stockholders, depending on the company that issued the stocks. If the issuer is able to generate large profits then the profits earned by stockholders will also be large. The higher the benefits offered, the higher the risk that will be faced in investing [1]. Therefore it is necessary to predict the current stock price based on yesterday's stock price.

In the stock investment instrument one of the determinants of the rate of return is the gain that is positive between the selling price and the purchase price. Stock price movements generally depend on economic conditions such as monetary policy indicated by the amount of money in circulation, interest rates, fiscal policy or taxes. While affecting the fluctuation of stock prices is the performance of stocks, which became one of the factors of consideration to determine the preferred stock investors. Several decades ago, approaches in predicting stock prices have been applied such as linear regression, time-series analysis, and chaos theory. From some of these approaches there are still some errors in the prediction. The use of machine learning such as neural networks [2-4] then the fuzzy system [5] has been applied to make predictions as the solution of the problem.

In another study, the Adaptive Network Inference System based fuzzy approach has been used to predict stock prices in Istanbul. In the study [5] has been used for three main stages. In the study [3] presented an integrated system with wavelets transform and recurrent neural networks based on bee colony to optimize the prediction of stock prices and their equation.

By knowing stock prices, investors can plan the right strategy to make a profit, but stock prices are fluctuating due to several factors [6]. From the stock movement can be predicted by investors by performing historical analysis and tend to stock prices in the previous period.

Computational stock prediction method can be done by using Back Propagation Neural Network method. The BPNN method is a method that can handle non-linear and time series data. Back propagation Neural Network is a multi-layer perceptron algorithm that has two forward and backward directions, so in the training process there are three layers: input layer, hidden layer, and output layer. As a result of the hidden layer, the error rate on BPNN can be reduced compared to single layer [6]. In this case the hidden layer function to update and adjust the weights so that we get a new weight value that can be directed to approach the desired target. Weight adjustment on the parameters of the BPNN method is important because it affects predicted results.

Amin Mughaddam et al. [7] proposed forecasting the stock index using the artificial neural network (ANN) method on daily transactions. In the ANN approach the author uses a back propagation algorithm to conduct his data training. In another study, Pesaran et al. [8] evaluated the effectiveness of the use of technical indicators such as the average movement of the closing price, the momentum of closing prices on the capital market in Turkey. To illustrate the correlation relationship between technical indication and stock price is investigated using hybrid Artificial Neural Network model, using optimization technique of harmony search approach (HS) and Genetics Algorithm used as the most dominant selection approach in technical indicators.

In another study Akhter et al. [9] proposed a hybrid model for predicting stock returns given the non-linear name of the model. Models comprise the average autoregressive model movement and exponential smoothing called the recurrent neural network. The recurrent method produces highly optimized prediction compared to the linear model. In another study Wei Shen et al. [10] using the Radial basis Neural Network method to conduct training and learning. The author also uses genetic algorithm method and particle swarm optimization to produce optimization using ARIMA dataset. The results show good performance in optimization using artificial fish swarm algorithm compared to some methods such as support vector machine.

It can be concluded that the neural network based algorithm has been known and widely used as a time-series data prediction algorithm. Therefore, this research will use Back propagation Neural Network algorithm optimized using Gradient Descent method as approach to predict stock price.

RELATED WORK

As a prediction system developed in the stock price prediction to help investors in making financial decisions. There are several strategies that explain about gaining profit in stock selling strategies. In most researches it focuses on "lowest price buy", "highest selling price". On the "lowest buy" and "highest selling" strategy of stocks occurs when stocks are at the lowest price and sell shares when prices are highest.

The method used to implement "Low Buy", "Top Selling" in trading strategy is shown in the index or delta symbol in the categorization. The delta symbol is defined as the difference between the closing index value for the next day, and the closing index value for the previous day's day. So this condition depends on the classification and strategy that will be recommended to investors to make stock expenditure, holding stock position or will sell stocks.

Indications on the forecasting process can be classified in the form of multivariate or univariate models. A univariate model of its use at a past value of a time series will be a prediction [7]. The approach has the disadvantage that it is not according to the environmental impact and interaction between several different factors between outputs. In the multivariate model do additional information such as market indications, technical indicators or fundamental factors of the company as input [7].

There are some previous studies on forex predictions using back propagation models that have been done by Joarder Kamruzzaman et al. [11] that accuracy in predicting foreign exchange (Forex) correctly is essential for future investment. Using computational intelligence based on forecasting techniques has proved very successful in making predictions. Joarder Kamruzzaman et al. [11] has developed and explored three Artificial Neural Networks based on model forecasting using Standard Back propagation, Conjugated Scaled Gradient and Back propagation with Baysian Regularization for the Australian Exchange to predict six different currencies against the Australian dollar.

Adetunji Philip, et al. [12] presents the statistical model used for forecasting. In this work, the model used in forecasting is an artificial neural network model of foreign exchange rate forecasting model designed for forecasting foreign exchange rates to correct some problems. Design is divided into two stages: training and forecasting.

At the training stage, back propagation algorithms are used to train foreign exchange rates and training for input estimates. Sigmoid Activation Function (SAF) is used to convert inputs into various standards [0, 1]. The study weight was randomly within the range of [-0.1, 0.1] to obtain output consistent with training.

Awajan et al. [13] proposed hybrid empirical mode with moving model for improve performance in forecasting for financial time series to solved daily stock market.

SAF is depicted using a hyperbolic tangent in order to improve the level of learning and make learning efficient. Feed forward Network is used to improve the efficiency of back propagation. Perceptron Multilayer network is designed for forecasting. The dataset of the FX Converter website is used as input in back propagation for evaluation and forecasting of foreign exchange rates.

METHODOLOGY

This stage will discuss the method that will be used to solve prediction problems based on neural networks. In this approach we propose a neural network method that is utilized using gradient descent. The model we propose as in Figure 1 as follows:

Figure 1: Proposed Method.
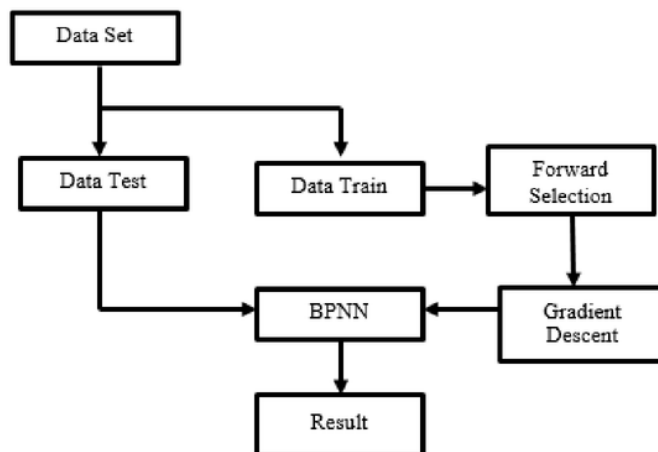
Back Propagation

Back propagation is a decrease in the gradient to minimize the square of output or output errors. There are three stages in network training such as forward propagation or advanced propagation, step propagation step, and the stage of weight change and bias. This network has an architecture that consists of input layer, hidden layer and output layer.

The procedure in the back propagation method can be explained as follows:
1. Initialize network weights at random
2. For each sample data, calculate the output based on the current network weight
3. Perform the process of calculating the error value for each output and hidden node (neutron) in the network. Network relation weights are modified
4. Repeat step 2 so as to achieve the desired condition.

Calculation error in output layer with equation formula:

$$Err_i = O_i(1 - O_i)(T_i - O_i) \qquad (1)$$

Where O_i output of the unit node i output, the T_i is the true value of the training data

node output.

Calculation error on hidden layer with equation formula:

$$Err_i = O_i(1 - O_i)\sum_j Err_j W_{ij} \tag{2}$$

Where O_i is output from the *hidden node* unit i which has output j in the *layer*, Err_j is *error* value in the unit *node* j W_{ij} is the weights between the two *nodes (neurons)*. After the error value of each node (neuron) is calculated, modifications are made to the network weights by the equation:

$$W_{ji} = W_{ij} + l \cdot Err_j \cdot O_i \tag{3}$$

Where *l* is learning rate with value 0 to 1. If the value of the learning rate is small, then the weight changes will be slight in each iteration and vice versa. The value of learning rate decreases during the learning process.

The back propagation neural network description as follows:
1. Initialization of input, bias, epoch, learning rate, error, target, initial weight and bias.
2. Calculate the input value (z, n) on each pair of input elements on the hidden layer with the formula:

$$z_{in_j} = b1_j + \sum_{i=1}^{n} x_1 v_{ij} \tag{4}$$

Parameter b_1 is the input bias, and y is the weight. If we use sigmoid activation, compute the output with the following equation:

$$z_j = f(z_{in_j}) = \frac{1}{1 + e^{z_{in_j}}} \tag{5}$$

Calculate the output signal from the hidden layer to get the output layer output using the equation:

$$y_{in_k} = b2_k + \sum_{i=1}^{m} x_i v_{jk} \tag{6}$$

Calculate the Mean Square error of the output with the equation.

$$MSE = \frac{\sum E^2}{n} \tag{7}$$

In the unit of output, use this formula to correct the weights and bias values δ.

Calculate the return signal (δ_{in}) from the output layer of each unit in the hidden layer.

$$\delta_{in_j} = \sum_{k=1}^{P} \delta_k w_{jk} \qquad (8)$$

In each hidden layer unit, calculate delta_1 to improve the weight and bias values:

$$\delta_j = \delta_{inj} f^1\left(z_{inj}\right) \qquad (9)$$

Improved weights with Δv and bias with Δb by using the formula

$$\Delta v_{ij} = \alpha(\delta_j x_j) \qquad (10)$$

$$\Delta b1_j = \alpha \delta j \qquad (11)$$

For all layers, fix the weights and biases, at the output layer:

$$W_{jk} = W_{jk} + \Delta W_{jk} \qquad (12)$$

In the hidden layer:

$$v_{ij} = v_{ij} + \Delta v_{ij} \qquad (13)$$

Gradient Descent

Back propagation neural network is known as an optimization of heuristic techniques that serves to speed up the training process by improving the adaptive learning rate weights [11].

The rate weight is corrected based on the gradient decrease with the learning rate (α) which is a nonnegative scalar with a value of $0 \geq \alpha \geq 1$. When the learning rate is too high then the algorithm becomes unstable, otherwise if the learning rate is too small then the algorithm will be very long in reaching convergent.

In the Gradient Descent approach, the value of the learning rate will change during the training process to keep the algorithm stable throughout the training. Process. Neural network performance is calculated based on network output value and training error.

In each epoch, new weights are calculated using the existing learning rate. Then calculated the performance of the new network. If the comparison of new neural performance and long nerve performance exceeds the maximum increase in work, then the weight will be ignored, and the value of the learning rate will be increased by multiplying it with the parameter of the addition of learning rate. Calculate the

weight and bias of the new output layer using the eqn. (14). The steps of this heuristic technique are:

$$w_{jk}(new) = w_{jk}(old) + \Delta w_{jk} \tag{14}$$

$$Z_k(new) = bZ_k(old) + \Delta bZ_k \tag{15}$$

Calculate the weight and bias of the hidden layer using the equation:

$$v_{jk}(new) = v_{jk}(old) + \Delta v_{jk} \tag{16}$$

$$b1_j(new) = b1_j(old) + \Delta b1_j \tag{17}$$

Calculate the performance of new neural network (pers2) by using new weights. Compare the performance of new neural networks (pers2) with previous neural network performance (pers).

If $\dfrac{pers2}{pers} > max\_perf\_inc$ then learning rate $(\alpha) = \alpha * lr_{dec}$

If $\dfrac{pers2}{pers} < max\_perf\_inc$ then learning rate $(\alpha) = \alpha * lr_{dec}$

If $\dfrac{pers2}{pers} \leq max\_perf\_inc$ then the new weight is accepted as the weight now.

EXPERIMENTAL RESULT

In this paper experiment, to make the prediction process of stock price used public data obtained from yahoo, finance. The data to be tested for prediction is data from private bank stocks from 2010 to 2016. The stages in the experiment can be described as follows:

Preprocessing

The initial process prior to testing of the proposed method is the preprocessing stage. The preprocessing stage used in this study is data normalization, by using equation of formula.

$$x_{ik} = \frac{x_{ik} - min(x_k)}{max(x_k) - min(x_k)} \tag{18}$$

After preprocessing the data until the normalization process, then the data that has been done normalization formed into a multivariate model or training data that has a function to determine the number of neurons in the input layer.

Neural Network Training Results

After normalization of data, then converted into a form of training data that serves to determine the number of inputs. The following is the experimental model of the neural network method measured using Mean Square Error. Determination of parameters on the neural network is used to find the best model to make predictions, namely by finding the best value of each parameter. Determination of parameters on neural network in this study based on, Epoch or training cycle, Learning Rate, Hidden Layer number, the number of Layer in Hidden Layer and train method used.

Number of Input Layer

In this research the determination of the number of input layer by doing it manually and used the number of input layer as much as 4 inputs (Figure 2).

Figure 2: Neural network model.
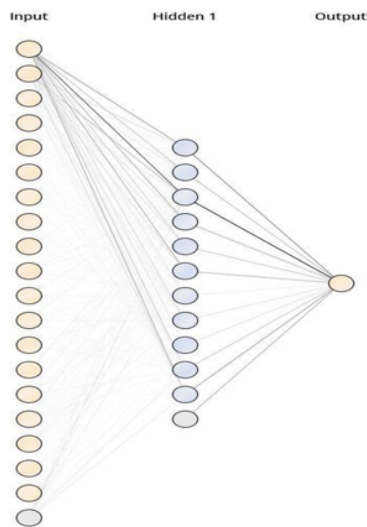
Epoch or Training Cycle

Determining the value of training cycle by doing with try test to insert value with range 100 up to 1000. The following is the result of the experiments that have been conducted for the training cycle determination.

The value of the selected training cycle based on the smallest MSE value generated. In the test results in the table above shows that the training cycle that produces the smallest MSE value is 600.
So the use of training cycle in this study refers to the results of epoch produced by

the Training Method using Resilient Back propagation and Gradient descent back propagation respectively.

Learning Rate

Determine the value of the learning rate by means of a trial to include values ranging from 0.1 to 1, as well as the value of the training cycle from the previous experiment. The following is the result of experiments that have been made for the determination of learning rate.

Hidden Layer

In this experiment, the determination of the number of hidden layers is 1 and determines the neuron size by testing the values of range 1 to 10 and multiples of 10, using the number of training cycles, the learning rate value of the previous experiment. The following are the results of experiments that have been done for the determination of hidden layer and neuron size on the graph:

This section presents the experimental results of the algorithms implemented for prediction stock price. These experiments are conducted on yahoo/finance dataset. Mean Square Error (MSE) is used to measure the performance of our prediction. The lower value of MSE the better is stock prediction. The equation of MSE as follow:
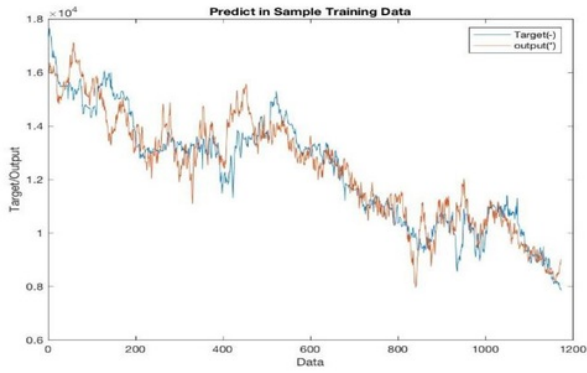
$$MSE(W) = \frac{1}{N}\sum_{i=1}^{N}(y_1 - y_1^{'})^2 \qquad (19)$$

For prediction process on perceptron have done with vector processed from input on model. The model have got from trained process so was getting y' as result class from prediction (Figures 3-10).
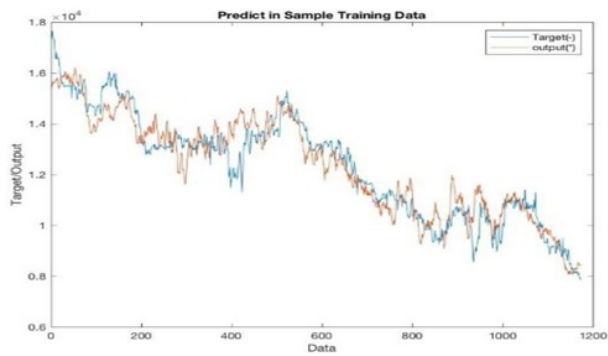
Figure 3: Training results with epoch 100 LR 0.1.
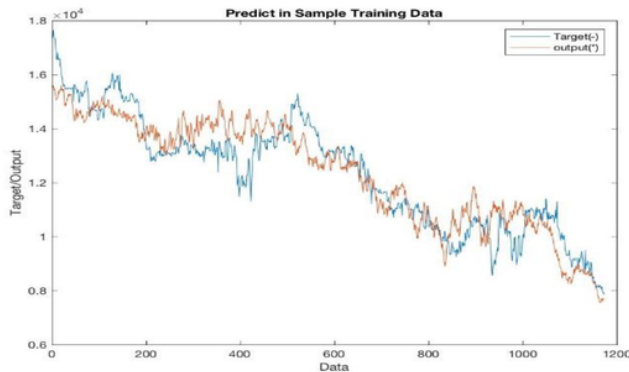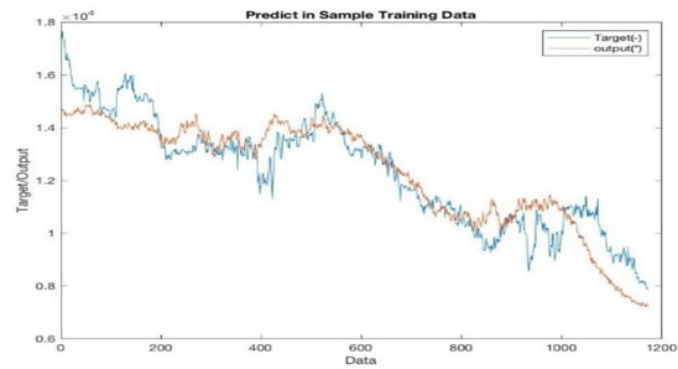
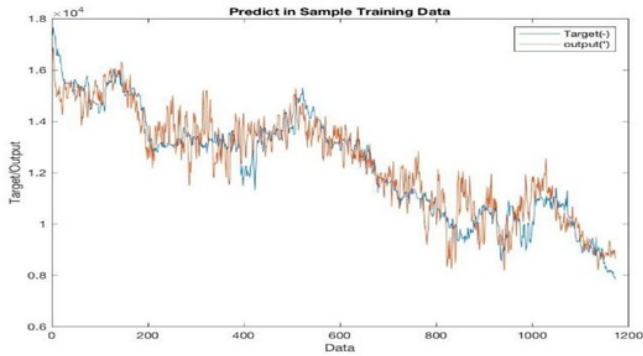**Figure 4:** Training results with epoch 100 LR 0.2.

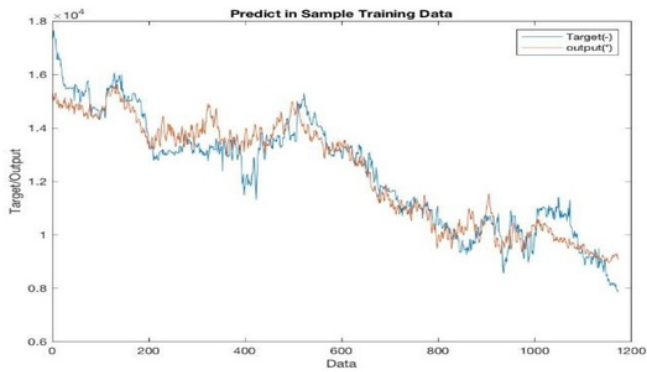**Figure 5:** Training results with epoch 100 LR 0.3.

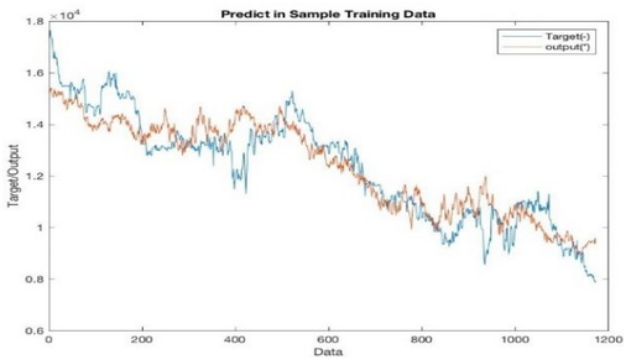**Figure 6:** Training results with epoch 100 LR 0.4.

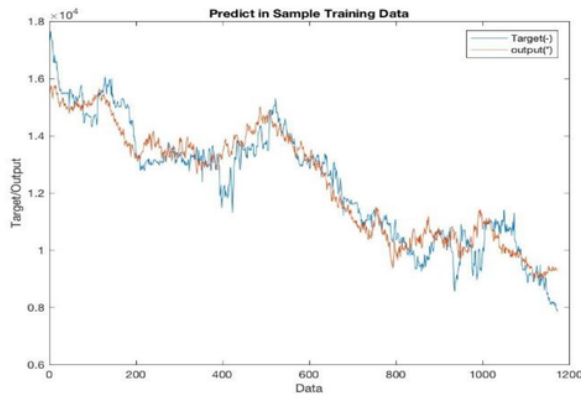**Figure 7:** Training results with epoch 100 LR 0.5.

Figure 8: Training results with epoch 100 LR 0.6.

Figure 9: Training results with epoch 100 LR 0.7.

Figure 10: Training Results with epoch 100 LR 0.8.
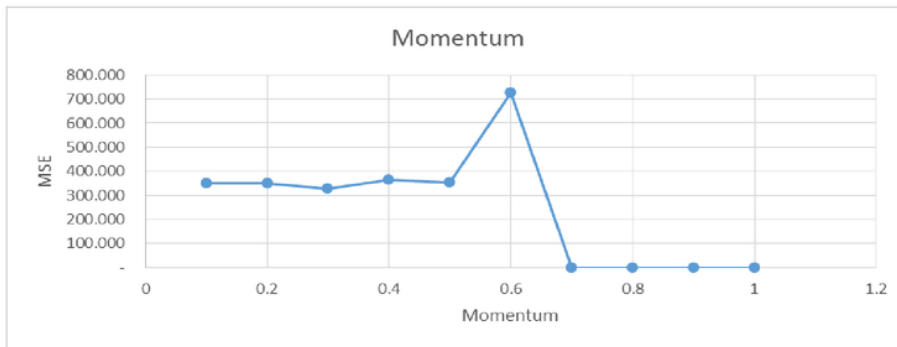

Neural Network Evaluation

The results of the experiments that have been done to determine the best model of the neural network, where the value of training cycle, learning rate, and neuron size based on different training methods. Then we get the best error result between Resilient back propagation and Method Gradient descent back propagation as follows (Table 1).

Table 1: MSE performance on training based on LR and Epoch.

MSE	Leaning Rate	Epoch	Hidden Layer
1209015.485	0.1	100	199
1626347.892	0.2	100	199
4502104.119	0.3	100	199
4499458.548	0.4	100	199
3378148.494	0.5	100	199
2214392.457	0.6	100	199
5742254.904	0.7	100	199
28174695.83	0.8	100	199

The best hidden layer amount is selected based on the smallest MSE value generated. In the test results in the table above shows that the number of hidden layers that produce the smallest MSE value is at learning rate 0.1 on hidden layer 199 (Figures 11-14).
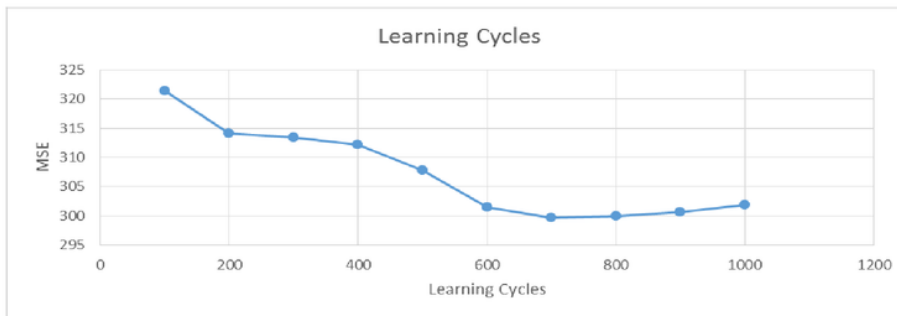
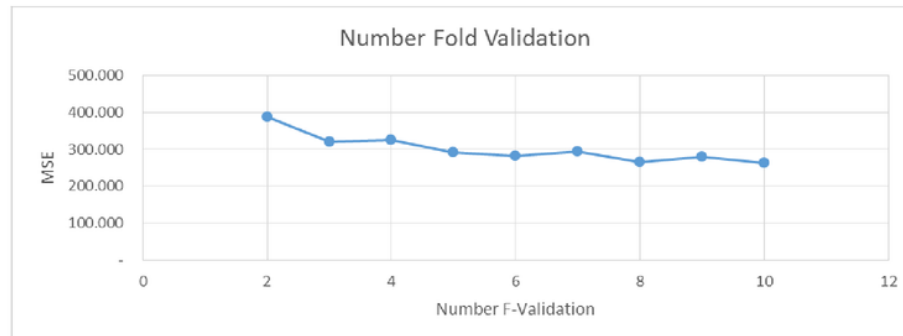**Figure 11:** Best momentum with value 0.3.

Figure 12: Best learning rate.

Figure 13: Best learning cycle.

Figure 14: Number fold validation.

[20]

CONCLUSION AND FUTURE WORK

In this paper we have tested using back propagation neural network algorithm with optimization using gradient descent on stock price data.

From the results of model-based neural network algorithm test, that the use of gradient descent method can find the value of training cycle, learning rate adaptively, so to get the best result in prediction can be determined value automatically, the impact can be more efficient in computing. The best results from predictions are obtained from the smallest MSE values of the computational results of each prediction.

REFERENCES

1. Pai P, Lin C (2005) A Hybrid ARIMA and Support Vector Machine Model in Stock Process Forecasting. Omega, pp: 497-505.
2. Wilamowski B, Chen Y, Malinowski A (1999) Efficient algorithm for training neural networks with one hidden layer. International Joint Conference.
3. Hsieh J, Hsiao F, Yei W (2011) Forecasting stock markets using wavelet transforms and recurrent neural networks : An system based on artificial bee colony algorithm. Applied Soft Computing 11: 2510-2525.
4. Hagan M, Beale H (1999) Neural Network Design. PWS Publishing.
5. Enke D, Grauier M, Mehdiyev N (2011) Stock market prediction with multiple regression, fuzzy type-2 clustering and neural networks. Procedia Computer Science 6: 201-206.
6. Afrianto R, Tjandrasa H, Arieshanti I (2013) Stock Exchange Prediction using Back Propagation Neural Network. Simantec.
7. Moghaddam AH, Hedayati MM, Moerteza E (2016) Stock market index prediction using artificial Neural Network. Journal of Economics, Financial and Administrative Science 21: 89-93.
8. Billah M, Waheed S, Henifa A (2015) Predicting Closing Stock Price using Artificial Neural Network and Adaptive Neuro Fuzzy Inference System. International Journal of Computer Applications.

9. Rather AM (2014) A Hybrid Intelligent Method of Predictions Stock Returns. Advances in Artificial Neural Network, pp: 1-7.

10. Wei S, Xiaopen G, Chao W, Desheng W (2011) Forecasting stock indices using radial basis function neural networks optimized by artificial fish swarm algorithm. Knowledge-Based Systems, 24: 378-385.

11. Kamruzzaman F, Sarker R (2009) Comparing ANN Based Models with ARIMA for Prediction of Forex Rates 22: 2-11.

12. Philip A (2011) Artificial Neural Network Model for Forecasting Foreign Exchange Rate. Information Technology Journal 1: 110-118.

13. Ahmad A, Ismail MT, Al W (2017) A Hybrid approach EMD-MA for short-term forecasting of daily stock market time series data. Journal of Internet Banking and Commerce.

stock price

5 "Neural Information Processing", Springer Nature America, Inc, 2006
Publication

1%

6 Moch Arief Soeleman, Mochamad Hariadi, Mauridhi Hery Purnomo. "Adaptive threshold for background subtraction in moving object detection using Fuzzy C-Means clustering", TENCON 2012 IEEE Region 10 Conference, 2012
Publication

1%

7 financedocbox.com
Internet Source

1%

8 Fasli, M.. "Learning approaches for developing successful seller strategies in dynamic supply chain management", Information Sciences, 20110815
Publication

1%

9 Pretesh B. Patel. "Neural Networks, Fuzzy Inference Systems and Adaptive-Neuro Fuzzy Inference Systems for Financial Decision Making", Lecture Notes in Computer Science, 2006
Publication

1%

10 M. A. Babak, A. R. Sharafat, S. K. Setarehdan. "An adaptive backpropagation neural network for arrhythmia classification using R-R interval signal", Neural Network World, 2012

1%

Publication

11 Submitted to College of Engineering Trivandrum
    Student Paper   1%

12 Junta Zeniarja, Ramadhan Rakhmat Sani, Ardytha Luthfiarta, Heru Agus Susanto et al. "Search Engine for Kids with Document Filtering and Ranking Using Naive Bayes Classifier", 2018 International Seminar on Application for Technology of Information and Communication, 2018
    Publication   <1%

13 Zhu, X.. "Predicting stock index increments by neural networks: The role of trading volume under different horizons", Expert Systems With Applications, 200805
    Publication   <1%

14 Bayu Kanigoro, Afan Galih Salman. "Recurrent gradient descent adaptive learning rate and momentum neural network for rainfall forecasting", 2016 International Seminar on Application for Technology of Information and Communication (ISemantic), 2016
    Publication   <1%

15 vdocuments.mx
    Internet Source   <1%

www.hrpub.org

16 Internet Source <1%

17 www.canberra.edu.au
Internet Source <1%

18 Andik Setyono. "Study and design framework for e-Dengue teleconsultation system", 2011 2nd International Conference on Instrumentation Communications Information Technology and Biomedical Engineering, 11/2011
Publication <1%

19 Chauhan, Seema Shrivastava, R.K.. "Reference evapotranspiration forecasting using different artificial neural networks algorithms.(Tech", Canadian Journal of Civil Engineering, Sept 2009 Issue
Publication <1%

20 Lecture Notes in Computer Science, 2008.
Publication <1%

21 Dong Tianyang, Zhang Jian, Gao Sibin, Shen Ying, Fan Jing. "Single-Tree Detection in High-Resolution Remote-Sensing Images Based on a Cascade Neural Network", ISPRS International Journal of Geo-Information, 2018
Publication <1%

22 D. Habermann, A. Hata, D. Wolf, F.S. Osorio.

"Artificial Neural Nets Object Recognition for 3D Point Clouds", 2013 Brazilian Conference on Intelligent Systems, 2013
Publication

<1%

23 Patel, Jigar, Sahil Shah, Priyank Thakkar, and K Kotecha. "Predicting stock market index using fusion of machine learning techniques", Expert Systems with Applications, 2015.
Publication

<1%

24 Dongli Zhang, Xinping Guan, Yinggan Tang, Yong Tang. "Modified Artificial Bee Colony Algorithms for Numerical Optimization", 2011 3rd International Workshop on Intelligent Systems and Applications, 2011
Publication

<1%

25 Mustain Billah, Sajjad Waheed, Abu Hanifa. "Stock market prediction using an improved training algorithm of neural network", 2016 2nd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE), 2016
Publication

<1%

26 V. Pathak. "Conjoint analysis for quantification of relative importance of various factors affecting BPANN classification of urban environment", International Journal of Remote Sensing, 11/1/2006
Publication

<1%

27 Wei, Yicheng, Junzo Watada, and Witold Pedrycz. "A Fuzzy Support Vector Machine with Qualitative Regression Preset", 2012 Sixth International Conference on Genetic and Evolutionary Computing, 2012.
Publication

<1%

28 ourspace.uregina.ca
Internet Source

<1%

29 Lianguang Mo. "The Engineering Cost Evaluation Based on IPSO", 2015 8th International Conference on Intelligent Computation Technology and Automation (ICICTA), 2015
Publication

<1%

30 "Artificial Neural Networks and Machine Learning – ICANN 2016", Springer Nature, 2016
Publication

<1%

31 "Engineering Applications of Neural Networks", Springer Nature America, Inc, 2013
Publication

<1%

32 Su-Lin Pang, Yan-Ming Wang, Yuan-Huai Bai. "Credit scoring model based on neural network", Proceedings. International Conference on Machine Learning and Cybernetics, 2002
Publication

<1%

33 Ahmed Radhwan, Mahmoud Kamel, Mohammed Y. Dahab, AboulElla Hassanien. "chapter 85 Forecasting Exchange Rates", IGI Global, 2016

Publication

<1 %

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		