Stock Market Prediction by Using Artificial Neural Network

Yunus YETIS[1], Halid KAPLAN[2], and Mo JAMSHIDI[3], Fellow IEEE

Department of Electrical and Computer Engineering
University of Texas at San Antonio
San Antonio, Texas, USA
[1] yunusyetis68@hotmail.com ,[2] halid88@gmail.com , [3]moj@wacong.com

Abstract— **A neural networks based model have been used in predicting of the stock market. One of the methods, as an intelligent data mining, is artificial neural network (ANN). In this paper represents how to predict a NASDAQ's stock value using ANNs with a given input parameters of share market. We used real exchange rate value of NASDAQ Stock Market index. This paper makes use generalized feed forward networks. The network was trained using input data of stock market price in between 2012 and 2013. It shows a good performance for NASDAQ stock market prediction.**

Keywords— **Stock Market, Prediction, Artificial Neural Networks**

I. INTRODUCTION

Prediction of stock market price is one of the most important issues in finance. Many researchers have been given their idea how to forecast the market price in order to make gain using different techniques, such as technical analysis, statistical analysis, with different methods [1]

Nowadays, artificial neural networks (ANNs) have been applied in order to predict exchange index prediction. ANN is one of data mining techniques that are learning capability of the human brain. Data patterns may perform dynamics and unpredictable because of complex financial data used. Several researches efforts have been made to improve efficiency of share values [2].

ANNs have been used in stock market prediction during the decade. Kimoto had used one of the first projects that was the forecasting of Tokyo stock market index by using ANNs [3].

Mizuno and friends had applied the Tokyo stock exchange to forecast buying and selling signals with an overall forecasting rate of 63% by using ANN [4]. Sexton and friends started of learning at random points that indicate in the training process [5]. Phua and friends had applied ANNs with the genetic algorithm to the stock market value of Singapore and predicted the market value with an accuracy of 81% [6].

In this paper, we review the ANNs for the important problem of prediction of the stock market and create the ANN to predict the stock market.

II. ARTIFICIAL NEURAL NETWORK

Artificial neural networks (ANNs) are an information processing system that was first inspired by generalizations of mathematical of human neuron (***Fig-1***).

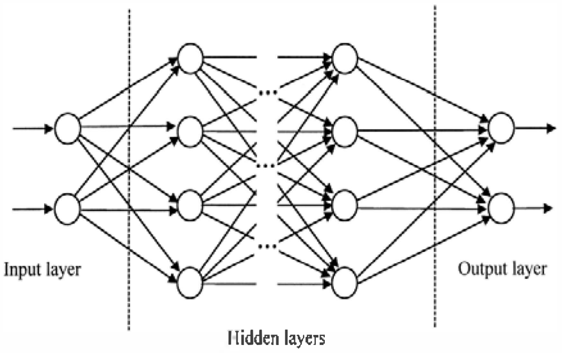

Fig-1: *Architecture of a feed forward MLP*

Each neuron receives some signals from other neurons or outside. Above figure has three layers of neurons, where one input layer is present. Every neuron employs activation function that fires when total input is more than a given threshold.

In this paper, we focus that Multi Layer Perception (MLP) networks are layered feed-forward networks typically trained with backpropagation [7].

MLP neural networks select one of the examples of training; make a forward and a backward pass. The first advantage of MLP networks is that eases an approximation of any input or output map. The first disadvantage is that they train very slowly and require lots of training data. It should be dramatically decreased the learning speed and increased number of neurons and layers of the networks.

III. METHODOLOGY

The work was divided into following units:

1. Training Process
2. Experimental Results
3. Comparison

A. Training Process

Training process is used by the free parameters which are weights and optimal values of the networks. Learning models, which are used for MLP, are characterized by particular input patterns and the changes in connection weights. In these models, the input layers distribute input signals to the optimal values of the network. Weight of the connections passes through of signals. Hidden layers and output layer include a vector of processing elements with an activation function.

One of the networks which are weight and biases is ready for training. The network is adjusted based on a comparison of the output and the target during the training (***Fig-2***).

Fig-2: *Network based on a comparison of the output the target (Source: Ball and Tissot, 2006 [13])*

The training process requires a set of examples of proper network behaviour and target outputs. During the training, the network of the weights and biases are repeated to minimize the network performance function. The Mean Square Error (MSE) was performance function used during training of feedforward neural network. MSE is the average squared error between the network outputs and the target outputs [9].

Training is the process of backpropagation errors though the system from the output layer towards. Backpropagation is one of the important ways for the training process because hidden units have not training target value that can be used it, so they must be trained based on errors from previous layers. Output is the only layer which has a target value in order to compare. Training occurs until the errors in the weights. Training has been kept on until reach all differences between updated weights and old weights in the previous epoch are below the threshold. MLP (Multilayer Perception) is the most common feed-forward networks (***Fig-3***). MLP has three types of layers: an input layer, an output layer and a hidden layer.

TABLE - I

(Source from NASDAQ, 2013)

Date	Open	High	Low	Close	Volume	Adj Close
4/15/2013	3277.58	3283.4	3213.46	3216.49	1779320000	3216.49
4/12/2013	3292.39	3296.5	3271.02	3294.95	1471180000	3294.95
4/11/2013	3289.59	3306.95	3287.74	3300.16	1829170000	3300.16
4/10/2013	3246.06	3299.16	3245.8	3297.25	1769870000	3297.25
4/9/2013	3229.81	3249.95	3215.02	3237.86	1498130000	3237.86
4/8/2013	3207.15	3222.26	3195.57	3222.25	1323520000	3222.25
4/5/2013	3174	3206.21	3168.88	3203.86	1594090000	3203.86
4/4/2013	3219.11	3226.24	3206.02	3224.98	1475720000	3224.98
4/3/2013	3257.38	3260.15	3210.39	3218.6	1813910000	3218.6
4/2/2013	3252.55	3267.93	3245.41	3254.86	1580800000	3254.86
4/1/2013	3268.63	3270.23	3230.57	3239.17	1481360000	3239.17
3/28/2013	3257.32	3270.3	3253.21	3267.52	1636800000	3267.52
3/27/2013	3230.76	3258.26	3227.02	3256.52	1420130000	3256.52
3/26/2013	3249.95	3252.93	3239.92	3252.48	1444500000	3252.48
3/25/2013	3255.85	3263.63	3222.48	3235.3	1666010000	3235.3
3/22/2013	3235.3	3247.94	3230.86	3245	1681360000	3245
3/21/2013	3228.17	3237.57	3215.69	3222.6	1692260000	3222.6
3/20/2013	3251.91	3257.99	3240.9	3254.19	1599120000	3254.19
3/19/2013	3246.7	3252.6	3205.42	3229.1	1690680000	3229.1
3/18/2013	3215.71	3249.37	3211.1	3237.59	1550510000	3237.59
3/15/2013	3260.46	3260.62	3242.65	3249.07	2305230000	3249.07
3/14/2013	3253	3258.93	3250.24	3258.93	1651650000	3258.93
3/13/2013	3243.04	3251.45	3230.62	3245.12	1577280000	3245.12
3/12/2013	3244.85	3249.78	3229.92	3242.32	1673740000	3242.32
3/11/2013	3237.74	3252.87	3233.67	3252.87	1628500000	3252.87
3/8/2013	3245.85	3248.7	3227.89	3244.37	1611700000	3244.37
3/7/2013	3224.5	3235.1	3221.47	3232.09	1675640000	3232.09
3/6/2013	3233.31	3233.44	3217.67	3222.36	1764020000	3222.36
3/5/2013	3200.38	3227.31	3200.27	3224.13	1891510000	3224.13
3/4/2013	3159.46	3182.27	3154.79	3182.03	1718290000	3182.03
3/1/2013	3143.54	3171.5	3129.4	3169.74	1870250000	3169.74
2/28/2013	3161.43	3182.6	3159.72	3160.19	2022530000	3160.19
2/27/2013	3129.72	3177.8	3127.27	3162.26	1727260000	3162.26
2/26/2013	3126.23	3135.57	3105.36	3129.65	1847750000	3129.65
2/25/2013	3180.59	3186.25	3116.25	3116.25	1930990000	3116.25
2/22/2013	3149.09	3161.82	3139.55	3161.82	1581500000	3161.82
2/21/2013	3154.88	3155.19	3118.62	3131.49	2052630000	3131.49
2/20/2013	3211.99	3213.25	3163.95	3164.41	2001800000	3164.41
2/19/2013	3197.46	3213.6	3194.92	3213.59	1843840000	3213.59
2/15/2013	3202.84	3206.21	3184.03	3192.03	1858670000	3192.03
2/14/2013	3182.74	3202.33	3182.39	3198.66	1924900000	3198.66
2/13/2013	3195.34	3205.52	3187.06	3196.88	1822450000	3196.88
2/12/2013	3190.73	3196.92	3184.84	3186.49	1786800000	3186.49
2/11/2013	3192.53	3194.01	3182.19	3192	1551370000	3192
2/8/2013	3178.06	3196.89	3177.18	3193.87	1816480000	3193.87
2/7/2013	3167.44	3170.42	3135.98	3165.13	1955960000	3165.13
2/6/2013	3159.38	3174.82	3157.35	3168.48	2002740000	3168.48
2/5/2013	3140.9	3178.52	3136.82	3171.58	2150080000	3171.58
2/4/2013	3161.72	3169.63	3130.57	3131.17	1874750000	3131.17
2/1/2013	3162.94	3183.14	3154.91	3179.1	2012930000	3179.1
1/31/2013	3140.67	3154.18	3136.82	3142.13	2190840000	3142.13
1/30/2013	3157.43	3164.06	3135.83	3142.31	2014350000	3142.31
1/29/2013	3149.62	3156.94	3133.11	3153.66	2050670000	3153.66
1/28/2013	3152.17	3161.83	3144.89	3154.3	1935590000	3154.3
1/25/2013	3140.65	3156.2	3135.86	3149.71	1920250000	3149.71
1/24/2013	3125.67	3153.56	3124.45	3130.38	2046990000	3130.38
1/23/2013	3155.82	3161.06	3149.74	3153.67	1698190000	3153.67
1/22/2013	3135.63	3143.18	3121.54	3143.18	1790730000	3143.18
1/18/2013	3127.91	3134.73	3119.2	3134.71	1860070000	3134.71
1/17/2013	3130.49	3144.05	3125.79	3136	1766510000	3136
1/16/2013	3110.72	3124.65	3106.79	3117.54	1692380000	3117.54
1/15/2013	3101.06	3112.29	3093.32	3110.78	1852870000	3110.78
1/14/2013	3113.65	3123.48	3104.08	3117.5	1876050000	3117.5
1/11/2013	3122.12	3126.59	3114.1	3125.64	1772600000	3125.64
1/10/2013	3125.64	3127.72	3098.47	3121.76	1754240000	3121.76

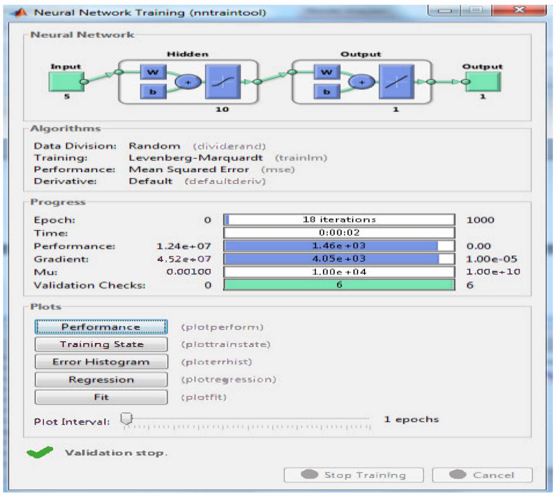

Fig-3: Neural Network Training

With these setting, the input vectors and target vectors will be randomly divided into three sets as follows (***Fig-4***):

70% will be used for training.

15% will be used to validate that the network is generalizing and to stopping training before overfitting.

The last 15% will be used as a completely independent test of network generalization(***Fig-5***).

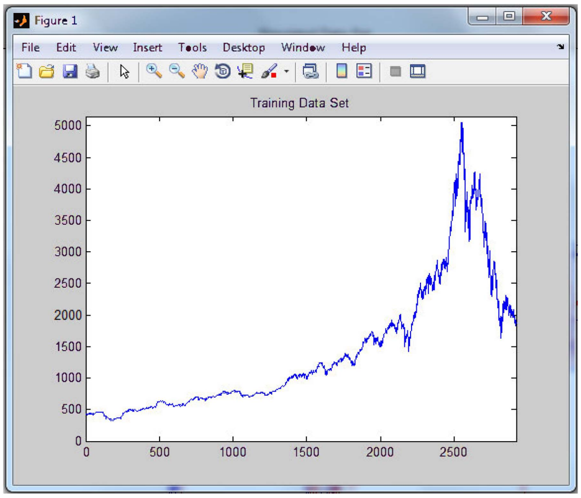

Fig-4: *Training Data Set*

Fig-5: *Randomly divided three sets of input and target vector*

B. Experimental Results

The dataset of NASDAQ daily stock price has been used for an experiment [10]. We used five input variables for ANN such as opening price, the highest price, the lowest price, volume of stock, and adjusted close price in the day to get the output at closing price of the day. Table of variables is shown above in *table-1*.

This observation is taken when other parameters are:

- ✓ Hidden Neurons:***10***
- ✓ Learning Rate(alpha):***0.4***
- ✓ Momentum Constant (mom): ***0.75***
- ✓ Max Epochs (epochs): ***1000***

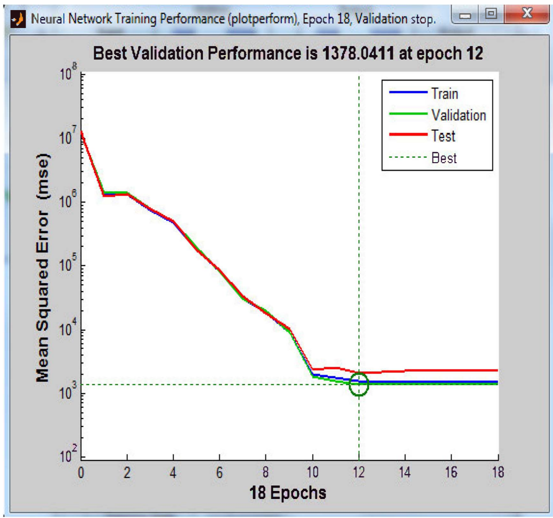

Fig-6: *Performance curve for training*

Mean Squared Error is the average squared difference between outputs and targets. Lower values are better. [12]. If the test curve had increased significantly before the validation curve increased, it means it is possible that some overfitting might have occurred [12]. According to the above information which explains diagram's principle of operation, the result is reasonable because of the following considerations which are the train set error, test curve had increased before the validation curve increased (***Fig-6***).

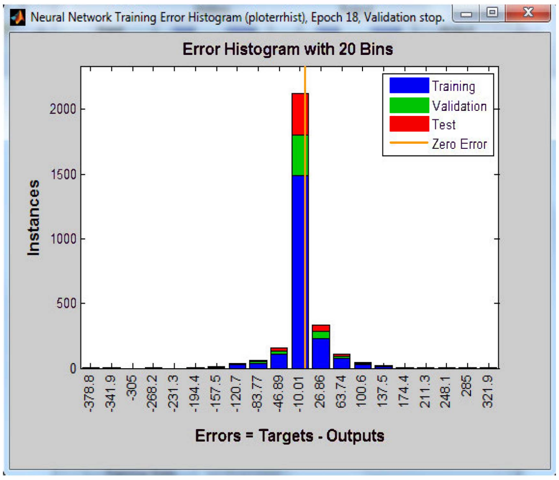

Fig-7: Error Histogram

Error histogram obtains additional verification of network performance. It can be clearly seen that errors are between -120 and +100 (***Fig-7***). Data sets are represented hundreds of thousands, so these errors are negligible considering to error is smaller than approximately %0.02 of targets.

Mean Square Error (MSE) means that average MSE is different between outputs and targets (***Fig-8***). Lower values are better. Zero means that there is no error.

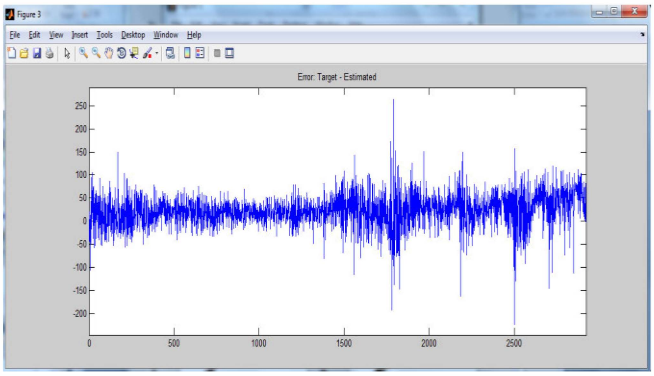

Fig-8: Errors=Target-Estimated

One of the important issues is that making the artificial neural network implement better relates to input normalization. Each input variable should be preprocessed. Mean value, average of the training set is small compared to its standard deviation. Index rage is between -1 and +1 [11]. We are able to use simple formula which is *Index(x) = (Index(x) – Min (Index)) / (Max (Index) – Min (Index))* [11]. It can be clearly seen the regression plot of the training set. Each of the figures corresponds to the target from the output array. Regression values (correlation coefficients) are very close to 1. It means that the correlation between the outputs and the target is very high.

Regression is used to validate the network performance. The following regression plots display the network outputs with respect to targets for training, validation, and test sets. For a perfect fit, the data should fall along a 45 degree line, where the network outputs are equal to the targets. For this problem, the fit is reasonably good for all data sets, with R values in each case of 0.99 or above (***Fig-9***).

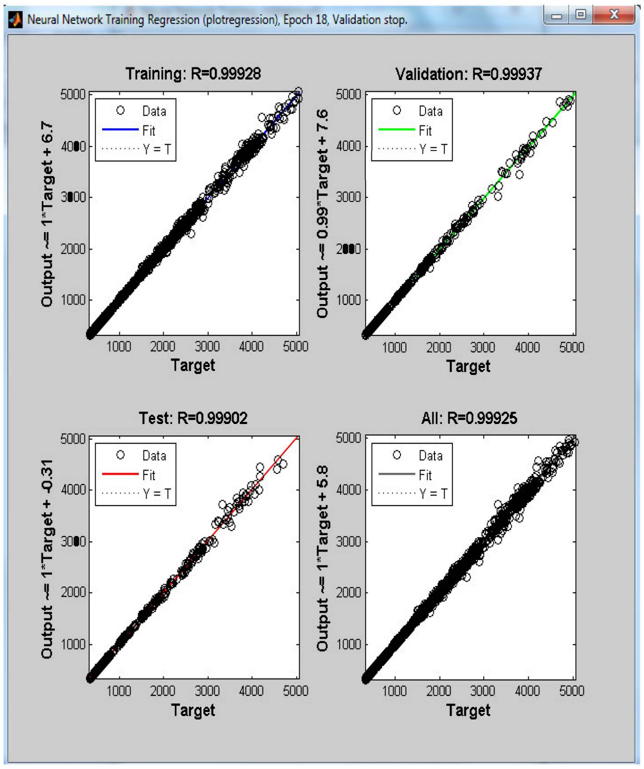

Fig-9: *Regression plot for training*

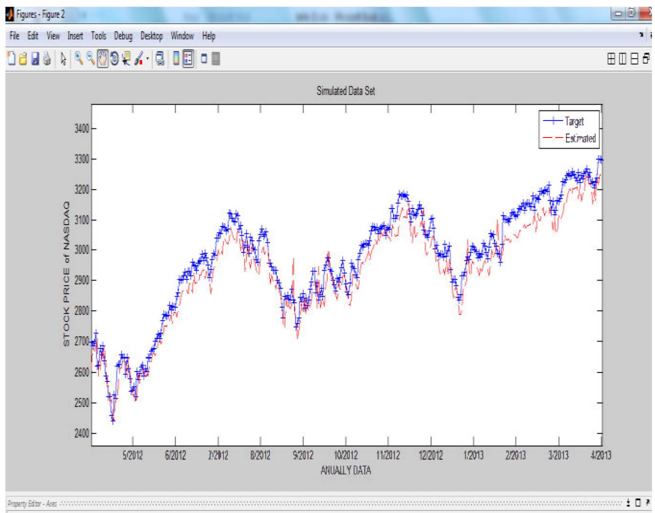

Fig-10: *Target and Estimated data of annually NASDAQ data*

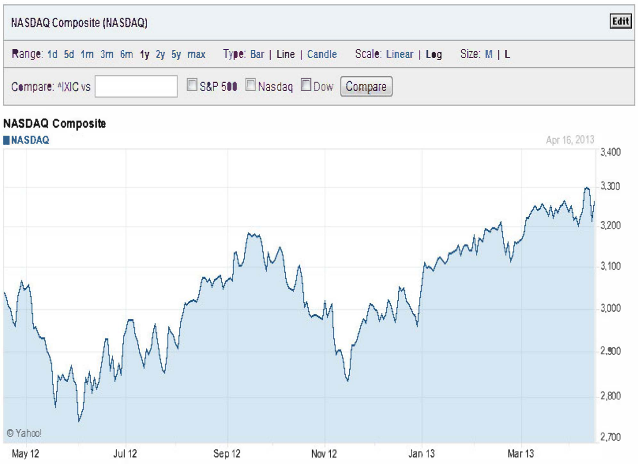

Fig-11: *Real Data of NASDAQ Stock Price (Source: finance.yahoo.com)*

IV. CONCLUSION

In order to create the prediction model, the implementation process should be included different steps like data collection, data pre-processing, classification and model evaluation. The experiment conducted in this paper uses simple and efficient approach to stock prediction using backpropagation with Feed Forward Network.

The performance of the network recorded was 99% in case of training with five inputs, 10 neurons in the input layer and one in the output layer. The best validation performance (MSE) is 1378.0411. This means that difference between predicted output and actual output, root of MSE (1378.04) is 37.12. While taking into consideration the stock closing prices are bigger than 3000, the error is smaller than 2%.

This model can be very beneficial for individual and corporate investors, financial analysts, and users of financial news. They can provide the future behavior and movement of stock prices; take correct actions immediately and act properly in their trading to gain more profit and prevent loss.

In future work, we intend to determine the critical impact of specific fundamental analysis variables on quality of stock price prediction.

REFERENCES

[1] M.T. Philip, K. Poul, S.O. Choy, K. Reggie, S.C. Ng, J. Mark, T.Jonathan, K. Kai, and W. Tak-Lam, *"Design and Implementation of NN5 for Hong Stock Price Forecasting"*, Journal of Engineering Application of Artificial Intelligence, vol. 20, 2007, pp.453-461.

[2] X. Wu, M. Fund and A. Flitman, *"Forecasting Stock Performance using Intelligent Hybrid Systems"*, Springerlink, 2001, pp. 447-456.

[3] T. Kimoto, K. Asawaka, M. Yoda, and M. Takeoka, *"Stock Market Prediction System with Modular Neural Network"*, Processing of International Joint Conference on Neural Networks, 1990, pp. 1-6.

[4] H. Mizono, M. Kosaka, H. Yajma, and N. Komoda, *"Application of Neural Network to Technical Anaysis of Stock Market Prediction"*, Studies in Informatics and Control, 1998, Vol.7, No.3, pp.111-120.

[5] S. R. Sexton, R. E. Dorsey and J. D. Johnson, *"Toward Global Optimization of Neural Network A Comparison of the Genetic Algorithm and Backpropagation"*, 1998, Decision Support Systems vol.22, pp.117-185.

[6] Phua, P. K. H. Ming, W. Lin, *"Neural Network With Genetic Algorithms For Stocks Prediction"*, Fifth Conferences of the Association of Asian-Pacific Operations Research Societies, 2000, 5th-7th July, Singapore

[7] D. E. Rumelhart, G. E. Hinton, R. J. Wiliams, *"Learning Internal Representation by Error Propagation Parallel Distributed Processing Explorations in the Microstructures of Cognition"*, McClelland J. L. (eds.), 1:318-362, MIT Press, Cambridge, 1986.

[8] Robyn Ball and Philippe Tissot, "*Demonstration of Artificial Neural Network in Matlab*", Division of Nearhsore Research, Texas A&M University, 2006

[9] S. Neenwi, P. O. Asagba, L. G. Kabari, *" Predicting the Nigerian Stock Market Using Artificial Neural Network"*, European Journal of Computer Science and Information, Vol.1, No.1, pp.30-39, June 2013

[10] Yahoo Historical Prices of NASDAQ, Retrieved May 5, 2013 *"http //finance.yahoo.com/q/hp?s=%5EIXIC&a=00&b=2&c=2013&d=03&e=15&f=2013&g=d"*

[11] Ciumac Sergiu (25 May 2011) *"Financial Predictor via Neural Network"*. Retrieved from http //www.codeproject.com

[12] Mathworks Website, November 2013 . http //www.mathworks.com/help/nnet/ug/analyze-neural-network-performance-after-training.html