Applying Multiple Linear Regression and Neural Network to Predict Bank Performance

Nor Mazlina Abu Bakar

Faculty of Business Management and Accountancy

University of Darul Iman Malaysia

21300 Kuala Terengganu, Malaysia

Tel: 60-9-665-3767   E-mail: normazlina@udm.edu.my

Izah Mohd Tahir (Corresponding author)

Faculty of Business Management and Accountancy

University of Darul Iman Malaysia

21300 Kuala Terengganu, Malaysia

Tel: 60-9-665-3863   E-mail: izah@udm.edu.my

Abstract

Globalization and technological advancement has created a highly competitive market in the banking and finance industry. Performance of the industry depends heavily on the accuracy of the decisions made at managerial level. This study uses multiple linear regression technique and feed forward artificial neural network in predicting bank performance. The study aims to predict bank performance using multiple linear regression and neural network. The study then evaluates the performance of the two techniques with a goal to find a powerful tool in predicting the bank performance. Data of thirteen banks for the period 2001-2006 was used in the study. ROA was used as a measure of bank performance, and hence is a dependent variable for the multiple linear regressions. Seven variables including liquidity, credit risk, cost to income ratio, size, concentration ratio, inflation and GDP were used as independent variables. Under supervised learning, the dependent variable, ROA was used as the target output for the artificial neural network. Seven inputs corresponding to seven predictor variables were used for pattern recognition at the training phase. Experimental results from the multiple linear regression show that two variables: credit risk and cost to income ratio are significant in determining the bank performance.   Two variables were found to explain about 60.9 percent of the total variation in the data with a mean square error (MSE) of 0.330. The artificial neural network was found to give optimal results by using thirteen hidden neurons. Testing results show that the seven inputs explain about 66.9 percent of the total variation in the data with a very low MSE of 0.00687. Performance of both methods is measured by mean square prediction error (MSPR) at the validation stage. The MSPR value for neural network is lower than the MPSR value for multiple linear regression (0.0061 against 0.6190). The study concludes that artificial neural network is the more powerful tool in predicting bank performance.

Keywords: Bank performance, Multiple linear regression, Neural Network, Malaysia

1. Introduction

Performance of the banking and finance industry plays a significant role in determining financial stability of any country. Furthermore, globalization and technological advancement has created a highly competitive market. This affects all organizations regardless of business emphasis. Banks are of no exception. They have to compete not only among the local banks, but also among the foreign banks. The situation requires the needs for the decision makers in this industry to be able to make an accurate decision. Mathematical and statistical tools can assist the decision makers to be able to make accurate predictions and face challenges ahead.

In the literature, the most common statistical technique used in predicting bank performance is by using multiple linear regression. The procedure is found to be very useful in determining bank profitability and consequently the performance of banks. In this study, two different statistical techniques will be used to predict bank performance. Multiple linear regression and artificial neural network techniques will be applied to Malaysian banks data.

Determinants of bank performance and the variables used are based from past literature and the results using both methods are obtained. The predictive abilities of the two techniques are then compared to find the more powerful tool in predicting bank performance.

The objectives of this paper are to predict Malaysian bank performance using Multiple Linear Regression and Artificial Neural Network and to evaluate which of this method is more powerful in predicting bank performance.

2. Literature Review

Researchers in banking and finance have indicated that bank performance is related to internal and external factors. The internal factors relate to banks' characteristics and external factors are described as the economic and legal environment (Athanasoglou, Brissimis & Delis, 2008). Multiple linear regression is a very common statistical technique used in finding the determinants of bank performance, for example Athanasoglou, Brissimis & Delis (2008), Haron (2004) and Sanusi & Mohamed (2007). The analysis of multiple linear regression often produced low coefficient of multiple determination, or R^2 values and the presence of outliers is seen to be a very common problem (Midi & Imon, 2006).

The performance measures are represented by return on assets (ROA), return on equity (ROE) and return on deposits (ROD) from balance sheets (Sanusi & Mohamed, 2007). In a study on panel data in finding determinants of Islamic bank profitability, Haron (2004) found that internal factors such as liquidity, total expenditures, funds invested and profit sharing ratio have a significant effect on bank profitability. Interest rate, market share and bank size, described as external effects, are also found to have the same effect in determining bank profitability.

A similar study in finding determinants of bank profitability, Sanusi & Mohamed (2007) found that bank's characteristics and the financial structure of a country are significant variables affecting bank profitability. They also compared the results of fixed effects and random effects on the proposed model and observed low adjusted R^2 values, indicating a low proportion of variation in profitability explained by the significant independent variables.

Athanasoglou, Brissimis & Delis (2008), investigated the effect of bank-specific and industry-specific and macroeconomics determinants on bank profitability in Greece. Two variables are found to have significant effect; labour productivity growth (positive effect) and operating expenses (negative effect). Variables used by Athanasoglou, Brissimis & Delis (2008), are adapted in this study to perform multiple linear regression on the Malaysian banks.

As with Artificial Neural Network, Vellido (1999) listed a variety of research that has used this method. In banking and finance, artificial neural network has been used to predict banks and firms bankruptcy, predict credit card performance, credit evaluation and also detect insurance fraud. Aiken (1999) used artificial neural network to forecast inflation and concluded that neural network is able to fairly accurately forecast the consumer price index of a country.

The future of the artificial neural network in finance is discussed by Brunell & Folarin (1997). They have looked at the promising performance of artificial neural network in debt risk assessment and its ability to improve on loan assessment. They found that the artificial neural network has helped bank managers to evaluate good or bad credit risks by estimating the likelihood that a firm's or borrower's ability to require additional capital through borrowing. The high performance of artificial neural network in many areas of banking and finance has led to the application of the artificial neural network to predict bank performance in this study.

The performance of artificial neural network has been compared with many other traditional statistical techniques. For example, artificial neural network is compared with multiple linear regression (Nguyen & Cripps, 2001 and Arulsudar, Subramaniam & Murthy, 2005), discriminant analysis and logistic regression (Leshno & Spector, 1996), decision trees and logistic regression (Delen, Walker & Kadam, 2004), stepwise regression and ridge regression (Chokmani,Quarda, Hamilton, Hosni & Hugo, 2008), logistic regression (Zhang, Hu, Patuwo & Indro, 1997). The artificial neural network has outperformed the traditional methods in all of these studies. Specifically, the artificial neural network is found to have better performance than multiple regression analysis when moderate to large data sample size is used (Nguyen & Cripps, 2001).

Comparison of artificial neural network and multiple linear regression has also been done in various fields of study. Artificial neural network is extensively being applied in predicting bankruptcy. Leshno & Spector (1996) have compared artificial neural network with multivariate discriminant analysis and logistic regression in their study on bankruptcy using a limited number of firms. Prediction capabilities of artificial neural network are found to be more accurate than the classical discriminant analysis and logistic regression. They also concluded that an ample number of examples must be provided for neural network to perform at its optimum. Another study in predicting bankruptcy is by Boritz & Kennedy (1995) who examined different types of artificial neural network and compared against other bankruptcy prediction techniques such as discriminant analysis, logit and probit techniques. Performance of the artificial neural network is found to be affected by the choice of variables. Although the artificial neural network has outperformed the traditional methods, the later has advantages of being easy to understand and use.

Nguyen & Cripps (2001) examined the performance of various artificial neural network architectures. Standard back propagation is found to perform better than other neural network architectures. The network performance is also found to improve with training size.

The applications of neural network in various fields of study have showed positive and promising results. Multiple linear regression is a very popular method but the method is non-robust, in which influential outliers can effect regression results significantly. Researchers in the field of robust statistics indicate that real data may contain about 1 to 10% outliers (Midi & Imon, 2006). The predictive ability and robustness of artificial neural network is an eye-catcher. Therefore, in this study, multiple linear regression and artificial neural network are used to predict bank performance and results of both methods are then compared. The results can then be of importance to predict bank performance in Malaysia.

3. Data Description and Methodology

A sample data set consisting of 13 banks for the period of 2001 – 2006 was randomly selected from a list of Malaysian banks obtained from Bank Negara Malaysia.   Data for all variables, except for GDP and CPI, were collected from the BANKSCOPE database.   Data for chosen variables were selected, calculated and transferred into an Excel spreadsheet. Data for Gross Domestic Product (GDP) and Consumer Price Index (CPI) were obtained from the Bank Negara Malaysia official website.

Predictor variables found to be significant in the banking and finance literature were adapted into the study. Return on assets (ROA) was used as a measure of bank performance and seven predictor variables were chosen to be analyzed. The chosen variables are listed in **Table 1**.

3.1   Multiple Linear Regression Model

Multiple linear regression analysis is a technique for modelling the linear relationship between two or more variables. It is one of the most widely used of all statistical methods. In banking and finance literature, regression analysis is a very common method used to find the determinants of bank performance.

The general linear regression model, with normal error terms, simply in terms of X variables is shown in Equation 1.

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + ... + \beta_{p-1} X_{i,p-1} + \varepsilon_i \tag{1}$$

where, $\beta_0, \beta_1, ..., \beta_{p-1}$ are parameters, $X_{i1}, ..., \beta_{p-1}$ are known constants,

ε_i are independent $N(0, \sigma^2)$

$i = 1, ..., n$

In building the multiple linear regression model for bank performance, the 96 data collected were randomly separated into 2 sub-samples – the training (86 data) and testing samples (10 data). Training data set is used for the model building, and the testing data set is used for the model validation at the end of analysis.

Kutner, Nachtsheim & Peter (2004) has recommended 30% of sample size as adequate size for model validation. In this study, only 10% of the sample size is used as the testing sample, due to the limited available number of samples.   In order to test the relationship between bank performance and its determinants, the following multiple regression equation is proposed for the bank data in Equation 2.

$$\begin{aligned} ROA_i &= \beta_0 + \beta_1 LIQ_i + \beta_2 LLOSS_i + \beta_3 SIZE_i + \beta_4 COSTINC_i + \beta_5 CONC_i \\ &\quad + \beta_6 GDP_i + \beta_7 CPI_i + \varepsilon_i \\ & i = 1, ..., n \end{aligned} \tag{2}$$

where

ROA_i	=	Return on average assets which serve as performance indicator
LIQ_i	=	Loan-to-assets ratio as a measure of liquidity
$LLOSS_i$	=	Loan loss provision-to-loans ratio as a measure of credit risk
$SIZE_i$	=	Size of a bank based on its total assets
$COSTINC_i$	=	Cost income ratio which measures bank efficiency
$CONC_i$	=	A concentration ratio, calculated by taking the largest 3 banks divided by total assets of the banking sector
GDP_i	=	Gross Domestic Product
CPI_i	=	Consumer Price Index
ε_i	=	error term

The underlying assumptions of linearity, normality, constant variance and independence of error terms must be satisfied in order to get a more valid model. Diagnostics for the underlying assumptions must be done and remedial measures can then be taken accordingly.

3.2 Artificial Neural Network

The development and abundance of high speed computers has made artificial neural network to become an increasingly popular research subject. The method has been applied to many areas including banking and finance. Artificial neural network has been applied to areas such as determining bank bankruptcy, time series, loan assessment etc.

The idea of neural network originated from the most fascinating organ in the human body, the brain. The human brain consists of billions of basic units called neurons. The basic neuron unit is illustrated in **Figure 1**. The neuron consists of dendrites, a cell body and an axon connecting to axon terminals.

Information or inputs received by the brain are transferred into the cell body through dendrites. The cell body acts as the processing unit, where all the learned information is then transferred into outputs and passed down by the axon. The muscles or other parts of the body receive the outputs via the axon terminals for actions. This concept was first studied in 1943 by McCulloch and Pitts to form a mathematical model.

Figure 2 illustrates a one hidden layer feed forward network with inputs $x_1, ..., x_i$ and output y_k. Each input is given its own synaptic weight. The weights are then transferred into the hidden layer, which consists of a number of hidden neurons. Each neuron performs a weighted summation of the inputs and then passes a nonlinear activation function. The output of the network is given by Equation 3.

$$y_k = \varphi(\sum w_{ki} v_k) = \varphi(v) \quad \textbf{(3)}$$

$y_k = $ network's output

$w_{ki} = $ synaptic weight between output of neuron k and input of neuron i

$v_k = \sum w_{ki} v_k = $ activation potential of neuron i (net input)

$\varphi(v) = $ activation function

The nonlinear activation function in the neuron is usually chosen to be a smooth step function. In this study, the sigmoid function $\varphi(v) = \dfrac{1}{1 + e^{-v}}$ is used.

Researchers in the field of robust statistics indicated that real data are never freed from outliers. About 10% of the real data set is expected to contain outliers. Preliminary checking was done to identify any gross outliers in the bank data. A simple method, the boxplot was chosen to identify outliers and get a rough idea of the symmetry of the data.

4. Results

4.1 Results of Multiple Linear Regression

The first order regression model was considered using all the seven predictor variables. Stepwise regression was performed on the modelling dataset and the obtained results are shown in **Table 2**.

From the regression output, it was found that only two predictor variables are significant in affecting the bank performance, LLOSS and COSTINC. The significant estimated parameters were found to be $\beta_0 = 3.881$, $\beta_2 = -0.199$ and $\beta_4 = -0.061$. The mean square error for the residuals was found to be 0.585. The total variation explained by the two significant variables, LLOSS and COSTINC on the bank performance ROA is about 57.2%, as shown by the R^2 value.

Table 3 shows the residual statistics obtained from the bank dataset. The Mahalanobis distance shows a very large maximum value of 24.662. The maximum studentized residual value is also much larger than the cut-off value 3. This indicates the existence of at least one outlying value in the regression model. Further investigation was made to identify the outlying value(s).

Remedial measures must be taken to solve the matter concerning influential outliers. Robust regression approaches such as Least Absolute Residuals, Reweighted Least Square and Least Median of Squares (LMS) regression can be employed in the presence of outliers. The robust method gives a superior result than the ordinary least square. Mean Square Error (MSE) from the ordinary least square may be inflated when influential outlier exists in the dataset.

In this study, the Least Median of Squares Regression is applied to the data set to compare with the results obtained by eliminating the outlier. Results of LMS regression is presented in **Table 4**.

The fitted robust regression function is $\hat{Y} = 3.4234 - 0.4197 LLOSS - 0.0288 COSTINC$

with a low proportion of variation in response explained by the model (28.73%). The residual scale estimate is calculated as 0.4295 with degree of freedom 78. In the presence of influential outlier(s), the robust regression method always outperformed the traditional multiple linear regression. In the absence of outlier, the two performed almost equally well, but the ordinary least square method is much preferred due to extensive calculations involved in the robust method.

4.2 Results of Artificial Neural Network

The 96 data bank data collected in the study were randomly assigned into three different sub-samples as given in **Table 5**. An ample number of data is needed for the training data set. Only 10% of the data are used for testing and validation purposes, due to the limited number of available sample.

Feedforward neural network or multilayer perceptron with one hidden layer and seven inputs, corresponding to the seven variables are suggested to be used in the study. Experiments are done to determine the best number of neurons in the hidden layer and to evaluate the performance of the neural network in predicting bank performance. Under supervised learning, the desired output (ROA) for each input is given to the network. The network will then adjust weights in the hidden neuron so as to minimize the error obtained between the network's output and the desired output.

Results for different numbers of neurons are presented in **Table 6**. The number of neurons with the lowest mean square error (MSE) for the testing data is chosen to be the best number of hidden nodes for the seven inputs. Neural network with 13 hidden nodes are identified to perform at its best during the training and testing. The lowest mean square error (MSE) of 0.00687 is obtained. R^2 value shows a satisfactory value at 0.66868.

The performance of the network during the training phase is very high for different numbers of neurons. The testing data set gives the lowest mean square error (MSE) value when the network contains 13 hidden neurons. The testing data gives a low R^2 value of about 0.669. The value indicates the proportion of total variation explained by the results is about 66.9%. Results are said to be satisfactory if the R^2 value is more than 0.80. Nevertheless, results obtained from the network has a higher R^2 value than the multiple linear regression model

4.3 Comparisons of Performance

Performance of the two methods under study, namely multiple linear regression and artificial neural network, in predicting bank performance is measured by using the mean square prediction error and the R^2 value which indicates the proportion of variation explained by the validation model. Model validation is done by using the 10 hold-out sample.

During the validation stage, the linear relationship gives mean predicted error (MSPR) of 0.6199. The network shows a very good performance with R^2 value of 0.99616 with very low means square predicted error (MSPR) of 0.0061. This indicates a very high predictive performance of the neural network with total variation explained by the factors by 99.616%. Performance of the two statistical methods can be summarized in Table 7.

From **Table 7**, the results indicate that the artificial neural network outperformed the multiple linear regression model. This finding is similar to Nguyen & Cripps (2001) with different type of data. The predictive ability of the artificial neural network is very high and gives a highly accurate prediction as a result of pattern recognition or generalization made by the network. Although the predictive ability of the artificial neural network is very accurate, it somehow lacks explanation on the parameters used. Some scholars describe artificial neural network as a 'black box'. Multiple linear regression gives an easy and simple explanation on the estimated parameters. This makes the method still very useful.

The multiple linear regression requires underlying assumptions of linearity, normality, etc. that may be violated when using real data. Remedial measures are then needed to be taken on the violation of these assumptions. Artificial neural network does not require these assumptions. It only requires a good number of data for the network to be able to recognize the pattern formed by the data. A single influential outlier can affect the estimated regression function but artificial neural network is very robust with noisy or unexpected data.

The artificial neural network has another disadvantage in which training time and determination of the optimal number of neurons can take quite some time and can be very complicated and exhaustive. The multiple linear regression method is very simple to apply by using available statistical computer software. The future of artificial neural network is promising but further research is needed on the development of computer software that can help reduce the long training time.

5. Conclusion

From the study, we may conclude that multiple linear regression can be used as a simple tool to study the linear relationship between the dependent variable and independent variables. The method provides two significant explanatory variables to bank performance and explains the effect of the contributing factors in a simple, understood manner. The method somehow has its limitations for its underlying assumptions are always violated when using real data. The presence of outliers also produces biased estimators of the parameters.

Violations of the underlying assumptions are often accompanied by remedial measures. Data transformation, robust regression and ridge regression are among the remedial measures to be taken. Thus, this requires the needs to understand further statistical techniques, which is out of the scope of this study.

The artificial neural network gives highly accurate results from the inputs. The method increases its performance with a large number of examples. An optimal number of neurons also need to be determined because the network tends to memorize with too many neurons but it can hardly make any generalization if too few are used. The method does not require any distributional assumptions and it is robust to outliers and unexpected data in the inputs. The artificial neural network outperformed the multiple regression in predicting bank performance but somehow, the method gives no explanation on the estimation of the parameters. Decision makers are provided with the information on the estimated parameters from the results of multiple linear regression. The prediction of the method is only made on the mean performance and thus gives a higher MSPR value.

A similar study can be performed using a larger dataset. As suggested by Kutner, Nachtsheim & Peter (2004), the validation model should consist about 30% of the dataset. Furthermore, three different sub-samples are required for the training, testing and validation in artificial neural network. The effect of different years and different banks should also be taken into considerations. Other predictor variables such as bank ownership, bank labour growth, macro or microeconomic factors are to be included which may explain the total variation in predicting bank performance.

References

Aiken, M. (1999). Using a Neural Network to Forecast Inflation, Industrial Management & Data Systems, Vol. 99/7:.296-301.

Arulsudar, N., Subramaniam, N. & Murthy, R.S.R (2005). Comparison of Artificial Neural Network and Multiple Linear Regression in the Optimization of Formulation Parameters of Leuprolide Acetate Loaded Liposomes, Journal of Pharmacy and Pharmaceutical Sciences, 8(2):.243-258.

Athanasoglou P. P., Brissimis, S. N. & Delis M.D. (2008). Bank-specific, Industry-specific and Macroeconomic Determinants of Bank Profitability, International Financial Markets, Institutions & Money, Vol. 18:.121-136.

Bank Negara Malaysia at www.bnm.gov.my.

Boritz, J. E., & Kennedy, D.B. (1995). Effectiveness of Neural Network Types for Prediction of Business Failure, Expert Systems with Applications, Vol. 9, No. 4, p.503-512.

Brunell, P.R. & Folarin, B.O. (1997). Impact of Neural Networks in Finance, Neural Computation & Application, Vol. 6:.193-200.

Chokmani, K. T. Quarda, J. V., Hamilton, S., Hosni,G. M., & Hugo, G. (2008). Comparison of Ice-Affected Streamflow Estimates Computed Using Artificial Neural Networks and Multiple Regression Techniques, Journal of Hydrology, Vol. 349:.383-396.

Delen, D., G. Walker, G., & Kadam, A. (2005). Predicting Breast Cancer Survivability: A Comparison of Three Data Mining Methods, Artificial Intelligence in Medicine, Vol. 34:.113-117.

Haron, S. (2004). Determinants of Islamic Bank Profitability, Global Journal of Finance & Economics, USA, 1 (1): 11-33.

Kutner, M., Nachtsheim C. J., & Peter, J. (2004). Applied Linear Regression Models, McGraw Hill/Irwin Series, 4th Edition.

Leshno, M., & Spector, Y. (1996). Neural Network Prediction Analysis: The Bankruptcy Case, Neurocomputing, Vol. 10:.125-147.

Nguyen, N., & Cripps, A. (2001). Predicting Housing Value: A Comparison of Multiple Linear Regression Analysis and Artificial Neural Networks, Journal of Real Estate Research, 22 (3): 313-336.

Midi, H., & Imon, A. H. M. R. (2006). The Use and Abuse of Statistics, Paper presented at the National Statistics Conference, Putrajaya International Convention Centre, Putrajaya, Malaysia, 4th-5th September 2006.

Sanusi, N.A., & Mohammed, N. (2007). Profitability of an Islamic Bank: Panel Evidence from Malaysia, Readings in Islamic Economics & Finance, Chapter 6:.97-116.

Vellido, A. (1999). Neural Networks in Business: A Survey of Applications (1992-1998), Expert Systems with Applications, Vol. 17: 51 – 70.

Zhang, G., Hu, M. Y., Patuwo B. E. & Indro, D. C. (1999). Artificial Neural Networks in Bankruptcy Prediction: General Framework and Cross-Validation Analysis, European Journal of Operational Research, 116:.16-32.

Table 1. Dependent & Independent Variables to Predict Bank Performance

Variable	Measure	Notation
Dependent Variable		
Profitability/ Performance	Net profits/Assets	ROA
Independent Variables		
1.　Liquidity	Loans/Assets	LIQ
2.　Credit Risk	Loan Loss Provisions/Loans	LLOSS
3.　Cost Income Ratio	Management Cost/Assets	COSTINC
4.　Size	ln (Real Assets)	SIZE
5.　Concentration Ratio	3-Bank Concentration Ratio	CONC
6.　Inflation	Current Period Inflation Rate (Consumer Price Index)	CPI
7.　Economic Activity	Gross Domestic Product	GDP

Table 2. Stepwise Regression Results

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	0.670	0.449	0.442	0.86267	
2	0.756	0.572	0.561	0.76496	0.991

ANOVA Results

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	50.891	1	50.891	68.383	.000
	Residual	62.513	84	0.744		
	Total	113.403	85			
2	Regression	64.835	2	32.417	55.399	.000
	Residual	48.569	83	0.585		
	Total	113.403	85			

Coefficients

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	T	Sig.
1	(Constant)	3.873	0.326		11.867	0.000
	COSTINC	-0.064	0.008	-0.670	-8.269	0.000
2	(Constant)	3.881	0.289		13.410	0.000
	COSTINC	-0.061	0.007	-0.637	-8.834	0.000
	LLOSS	-0.199	0.041	-0.352	-4.881	0.000

Table 3. Residuals Statistics

	Minimum	Maximum	Mean	Standard Deviation
Residual	-1.531	4.416	0.000	0.756
Standardized Residual	-2.001	5.773	0.000	0.988
Studentized Residual	-2.319	6.068	0.004	1.041
Deleted Residual	-2.056	4.879	0.006	0.843
Stud. Deleted Residual	-2.384	8.085	0.028	1.198
Mahalanobis Distance	0.010	24.662	1.977	4.046
Cook's Distance	0.000	1.298	0.043	0.205

Table 4. Results of LMS Regression

Coefficients:	Value	Std. Error	t-value	Pr(>|t|)
(Intercept)	3.4234	9.1990	0.6992	0.4865
LIQ	-0.3136	0.4006	-0.7829	0.4361
LLOSS	-0.4197	0.0606	-6.9222	0.0000
SIZE	0.0773	0.0363	2.1272	0.0366
COSTINC	-0.0288	0.0048	-6.0058	0.0000
CONC	-2.8235	5.1090	-0.5526	0.5821
CPI	0.0799	0.1574	0.5078	0.6130
GDP	-1.4113	2.5623	-0.5508	0.5833

Table 5. Sub-Samples in Artificial Neural Network

Type of Data	Percentage	Frequency
Training Data	80%	80
Testing Data	10%	6
Validation Data	10%	6

Table 6.Mean Square Error and R^2 value at Selected Number of Neurons

Neurons	Training Data		Test data	
	MSE	R^2	MSE	R^2
5	0.02040	0.81364	0.17329	0.85239
10	0.00799	0.94691	0.01500	0.83672
11	0.00505	0.96707	0.02444	0.65187
12	0.00515	0.96438	0.20735	0.59641
13	**0.00358**	**0.97767**	**0.00687**	**0.66868**
14	0.00529	0.96644	0.00964	0.79567
15	0.00297	0.96683	0.04524	0.60258
20	0.00428	0.97182	0.00738	0.87044
25	0.00672	0.96680	0.03358	0.64716

Table 7. Comparison of Performance

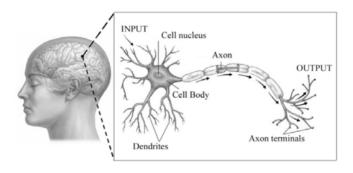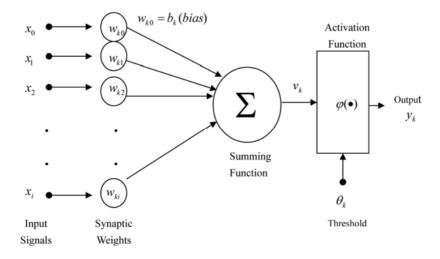| Method	MSPR
Multiple Linear Regression	0.6190
Artificial Neural Network	0.0061

Figure 1. The Structure of the Neuron

Figure 2. The Mathematical Model of the Neural Network