African Journal of Business Management Vol. 6(2), pp. 501-513,18 January, 2012
Available online at http://www.academicjournals.org/AJBM
DOI: 10.5897/AJBM11.1221
ISSN 1993-8233 ©2012 Academic Journals

Full Length Research Paper

Comparative analysis of stocks returns' predictable factors in Tehran Stocks Exchange and New York Stocks Exchange using artificial neural network and linear regression

Mahmood Yahyazadehfar*, Shahabeddin Shams and Alireza Garshasebi

Faculty of Economics and Administrative Sciences, University of Mazandaran, Babolsar, Iran.

Accepted 1 August, 2011

Stocks return prediction is one of the most important discussions of financial analysts in recent decades. The majority of researchers including Balvers et al. (1990), Breen et al. (1990), Campbell (1987), Fama (1977), Ferson (1989), Keim and Stambaugh (1986) and Schwert (1990) showed evidences that stocks return can be predicted by means of historical information such as periodic data of financial and economic variables. Although most studies indicated that the relationship between available information and stocks return is based on linear assumptions, there are no evidences that this relationship is completely linear. Therefore, it is probable that non-linear models have more ability to predict stocks price fluctuations and their returns. The present paper predicted the stocks return with artificial neural network (ANN) and linear regression in Tehran Stocks Exchange and New York Stocks Exchange (NYSE) using financial and macroeconomic variables. Furthermore, the results of two markets were compared. As it is known, stocks return depends on many variables. Thus, in this research, variables were screened and irrelevant or redundant variables were removed using principal component analysis (PCA). Then, in order to predict the stocks return, feed-forward and linear regression model were applied. Based on the mean square error (MSE), the results showed that the estimated error of ANN in two markets is less than the estimated error of linear regression. In addition, ANN and linear regression model will make a better prediction in NYSE.

Key words: Artificial neural network, feed-forward, financial market, principal component analysis, regression analysis.

INTRODUCTION

Stock return forecasting is an important financial topic. This is because once the prediction of returns is successful, monetary rewards will be substantial. However, predicting stock returns is very difficult since it depends on several known and unknown factors, and frequently, the data used for prediction is noisy, uncertain, and incomplete.

Numerous empirical studies have investigated the predictability of stock/market returns using macroeconomic and financial variables. Keim and Stambaugh (1986) find that several observable variables from bond and stock markets explain a substantial portion of future stock return movements. Pontiff and Schall (1998) showed that the book-to-market ratio is able to predict market returns. Lewellen (2004) demonstrates the predictive power of financial ratios such as dividend yield, book-to-market ratio and earnings-price ratio. Rapach et al. (2005) examines a large set of macro variables and presents evidence that stock returns can be predicted using macro variables. Abugri (2008) investigated whether the dynamics in key macroeconomic indicators like exchange rates, interest rates, industrial production and money supply significantly explain market returns. Chang (2009)

*Corresponding author. E-mail: m.yahyazadeh@umz.ac.ir. Tel: +989111119814.

examined the effects of macroeconomic variables including interest rate, dividend yield, and default premium on stock return movements using various regime switching GJR-GARCH models. He showed that macroeconomic factors can affect the stock return dynamics through two different channels. The effects of these economic variables on returns are related to stock market volatilities. In addition, he found that interest rate and dividend yield play an important role in predicting conditional variance. But these factors do not play any role in predicting transition probabilities. Most studies attempt to capture the relationship between the available data and the stock/market return using linear assumptions. Practically, there is no evidence that the relationship between stock returns and the financial and macroeconomic variables is perfectly linear. This is because there is significant residual variance of the predicted return from the actual return. Therefore, it is possible that nonlinear models will be more useful in explaining the residual variance and will produce more reliable predictions of stock return.

Artificial neural networks (ANN) and statistical tools are different methods that can be used to predict financial indexes. ANNs incorporate a large number of parameters which allows learning the intrinsic non-linear relationship presented in time-series, enhancing their forecasting possibilities (Haykin, 2001; Specht, 1990). ANNs have been successfully applied to predict important financial and market indexes, like for example, Standart and Pool500 (SP and 500), Nikei 225 Index, and others (Chen, 1994; Enke and Thawornwong, 2005; Huarng and Yu, 2006; Huang et al., 2007; Refenes et al., 1994; Yu and Huarng, 2008).

Olson and Mossman (2003) indicated that back propagation neural networks outperform the best regression alternatives for both point estimation and in classifying firms expected to have either high or low returns. de Faria et al. (2009) compared the performance forecasting using artificial neural networks and the adaptive exponential smoothing method. Their results showed that both methods produce similar results regarding the prediction of the index returns.

RESEARCH METHODOLOGY

Principal component analysis

The selection of input data is an important decision that can greatly affect the model performance. There are hundreds of financial and macroeconomic variables available for analysis. However, many of the variables may be irrelevant or redundant to the prediction of stock returns. Leaving out relevant variables or including irrelevant variables in the input data may be detrimental, causing confusion to the ANNs. The added volume of irrelevant or redundant variables can also slow down the network learning process. Principal component analysis (PCA) is a variable reduction procedure and is widely used in data processing and dimensionality reduction. It is useful when you have obtained data on a number of variables (possibly a large number of variables) and believe that there is some redundancy in those variables. PCA minimizes the sum of the squared perpendicular distances to the axis of the principle component (PC) while least squares regression minimizes the sum of the squared distances perpendicular to the x axis (Truxillo, 2003).

Principal component analysis is one of the methods of multivariate analysis and has been used widely with large multidimensional data sets. The use of PCA allows the number of variables in a multivariate data set to be reduced, while retaining as much as possible of the variation present in the data set. This reduction is achieved by taking p variables X1, X2. . .Xp and finding the combinations of these to produce principle components (PCs) PC1, PC2, . . .,PCp, which are uncorrelated. These PCs are also termed eigenvectors. The lack of correlation is a useful property as it means that the PCs are measuring different “dimensions” in the data set. Nevertheless, PCs are ordered so that PC1 exhibits the greatest amount of the variation, PC2 exhibits the second greatest amount of the variation; PC3 exhibits the third greatest amount of the variation, and so on. That is var (PC1) ≥ var (PC2) ≥ var (PC3) ≥. . .≥ var (PCp), where var (PCi) expresses the variance of PCi in the data set being considered. Var (PCi) is also called the eigenvalue of PCi. When using PCA, it is hoped that the eigenvalues of most of the PCs will be so low as to be virtually negligible. Where this is the case, the variation in the data set can be adequately described by means of a few PCs where the eigenvalues are not negligible. Accordingly, some degree of economy is accomplished as the variation in the original number of variables (X variables) can be described using a smaller number of the new variables (PCs).

Artificial neural network

One of the major application areas of artificial neural networks (ANNs) is forecasting. Recently, there is an increasing interest in using ANNs in financial and economic forecasting (Atsalakis and Valavanis, 2009; Cao and Perry, 2009; Chang et al., 2009; Zhu et al., 2008; Tsang et al., 2007; Shachmurove, 2005; Zhang and Beradi, 2007).

The success of ANNs applications can be attributed to their unique features and powerful pattern recognition capability. ANNs have been successfully applied to a wide range of forecasting problems in almost all areas of business, industry, and science. For example, in financial applications, ANNs have been used for predicting bankruptcy or business failure, exchange rate, interest rate, futures price, stock return, trading volume, capital market index, initial public offering price, property value and many others.

A neural network is a massively parallel system comprising highly interconnected, interacting processing elements (often termed units, nodes or neurons) that are based on neurobiological models. ANNs process information through the interactions of a large number of simple processing elements. Knowledge is not stored within individual processing units but is represented by the weight between units (Cheng and Titterington, 1994). A neural network is characterized by the pattern of connections among the various network layers, the numbers of neurons in each layer, the learning algorithm, and the neuron activation functions. In general, a neural network is a set of connected input and output units where each connection has a weight associated with it. Figure 1 shows the structure of simple neural network.

Feed-forward neural network

Over the last several decades, many types of ANNs models have been developed, each aimed at solving different problems. But by far the most widely and successfully used for forecasting has been the feed-forward type neural network. Feed-forward NNs are the most commonly used networks for a variety of applications in

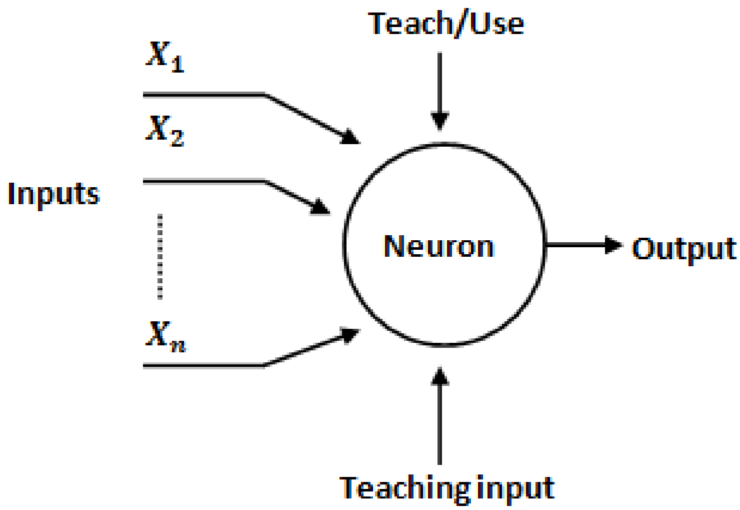

Figure 1. A simple neural network structure.

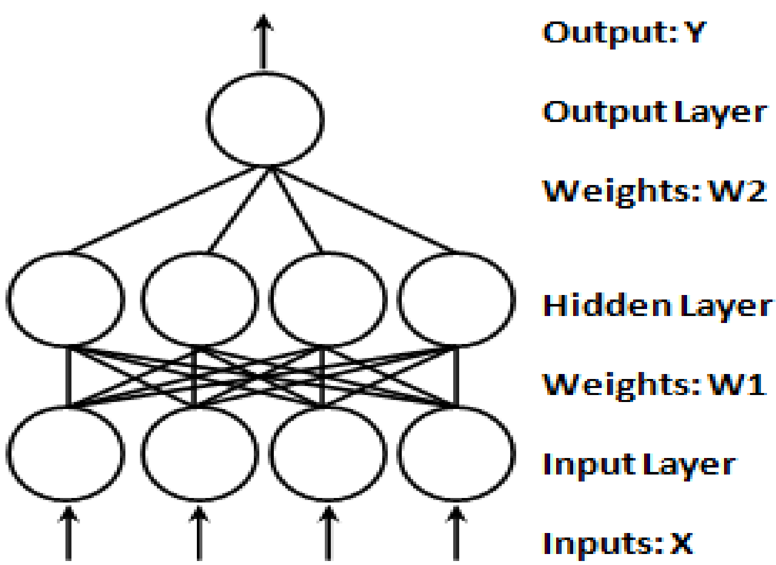

Figure 2. A typical feed-forward neural network.

finance and accounting (Coakley and Brown, 2000) and have been widely used for financial forecasting due to their ability to correctly classify and predict the dependent variable. Figure 2 shows the architecture of a three-layer feed-forward NN that consists of neurons (circles) organized in three layers: Input layer, hidden layer and output layer.

The neurons in the input nodes correspond to the independent or predictor variables (x) that are believed to be useful for forecasting the dependent variable (Y) which corresponds to the output neuron. Neurons in the hidden layer are connected to both input and output neurons and are keys to learning the pattern in the data and mapping the relationship from input variables to the output variable. With nonlinear transfer functions, hidden neurons can process complex information received from input neurons and then send processed information to the output layer for further processing to generate forecasts.

Back propagation learning algorithm

Artificial neural networks can be classified into several categories based on supervised and unsupervised learning methods and Feed-Forward and Feed-Back recall architectures. A back propagation neural network (BPNN) is a neural network that uses a supervised learning method and Feed-Forward architecture. A BPNN is one of the most frequently utilized NNs techniques for classification and prediction and is considered an advanced multiple regression analysis that can accommodate complex and non-linear data relationships (Jost, 1993). The output of a BPNN is compared with the target output and an error is calculated for each training iteration. This error is then back propagated to the NN and utilized to adjust the weights, thereby minimizing the mean squared error between the network's prediction output and the target output.

Consequently, the BPNN model yields predictive output that is similar to the target output.

The most popular neural network training algorithm for financial forecasting is the BPNN (Atsalakis and Valavanis, 2009; Cao and Parry, 2009; Chang et al., 2009; Lee and Chen, 2002; Lee and Chiu, 2002; McNelis, 2004; Vellido et al., 1999; Yudong and Lenan, 2009; Zhang et al., 1998), which has a simple architecture but a powerful problem-solving ability. BPNN is essentially a gradient steepest descent training algorithm. For the gradient descent algorithm, the step size, called the learning rate, must first be specified. The learning rate is crucial for BPNN since smaller learning rates tend to slow down the learning process before convergence while larger ones may cause network oscillation and unable to converge (Chi-Jie and Lu, 2010).

EMPIRICAL STUDY

Datasets and performance criteria

In the literature, there is an apparent uncertainty in selecting the input variables that are used to forecast stock returns. In fact, the inputs selected to model the ANNs are often different even when the same test case for prediction was examined. Based on the literature review, stock return is affected by different factors. Since there are no specific methods for selecting some of the effective variables in the stock return, identifying these variables is an effective issue for the researcher's progress. Table 1 indicates some of the studies that have applied various input variables.

Regarding the data access restriction, 17 macroeconomic and 16 financial variables for New York stock exchange and 15 macroeconomic variables and 16 financial variables for Tehran stock market were selected. Table 2 indicates the applied variables for each market. As shown in Table 2, applied financial variables for stock return prediction in New York Stock Exchange as well as Tehran stock market are the same. However, applied macroeconomic variables in predicting stock return in two above-mentioned markets are different due to lack of some of variables in Tehran Stock market.

Statistical populations include listed companies in Tehran and New York stock exchanges over 2007 to 2009. Listed companies in Standard and Poor's 100 for New York stock exchange and 100 listed companies in Tehran Stock exchange are selected as the samples of statistics. The monthly data is used in time series for the 3-year period during 2007 to 2009. In order to compare the obtained result in each technique, mean square error (MSE) prediction error is used. Also, T-test is used for comparing the obtained results.

Obtaining relevant variables and eliminating irrelevant variables using PCA technique

In order to identify relevant variables and redundant or irrelevant variables, PCA technique is used. Firstly, appropriate data is used in conducting analysis; Kaiser-Meyer-Olkin (KMO) and Barttlet's test are used as well. Obtained results are indicated in Table 3.

Cerney and Kaiser believed that value of KMO should be more than 0.6 and significant level of Barttlet's test should also be lower than 0.05. Table 3 indicates the related results of appropriate data. In order to use components analysis technique, LISREL 14 is used. Tables 4 to 7 show the results of principal components analysis technique for macroeconomic and financial variables for each market. Eigen values are identified in the second columns of the tables as the principle variables. By considering the models including three to seven variables, the best prediction for deemed period was established (Olson and Mossman, 2003). In this stage, 7 financial and 4 macroeconomic variables (totally 11 variables) for New York stock exchange and 4 financial and 4 macroeconomic variables (totally 8 variables) for Tehran stock exchange are identified. Also, among these identified variables, each of the variables that are extracted explains at least eight percent of the total variance alone. Therefore, the descriptive power of earning per share (EPS), quick ratio (QR), debt ratio (DTA) and exchange rate (ER) variables in New York stock Exchange and variable of unemployment rate (UR) in Tehran Stock Exchange that are less than eight percent of total variance are eliminated. Table 8 indicates the remaining principle variables for Tehran and New York stock exchange.

Modeling and predicting stock return using multiple linear regression technique

Multiple linear regression (MLR) is a method used to model the linear relationship between a dependent variable and one or more independent variables. MLR is based on least squares: the model is fit such that the sum-of-squares of differences of observed and predicted values is minimized. MLR is probably the most widely used method in predicting stock or market returns. The MLR model is reviewed:

$$Y_i = \beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \beta_3 X_{i,3} + \ldots + \beta_k X_{i,k} + \varepsilon_i$$

$X_{i,k}$: value of k^{th} predictor in year *I*; β_0: regression constant; β_k: coefficient on k^{th} the predictor; k: total number of predictors; Y_i: Predictand in year I; ε_i: error term.

For being assured of the data normality, Kolmogorov-smirnov (KS) test is used. With regard to 5% significant level, the result indicates that all research data are normal (Table 9). After conducting KS test, analysis of regression for each market is done in which the results are shown in Tables 10 and 11. Value of "t" must be more than 1.98 and the significant level be lower than 0.05. Therefore, DPS and M_1 variables in New York Stock

Table 1. Summary of input variables in other researches.

Article	Input variable
Austin et al. (1997)	S and P 500 div, stock/bond spread, commercial bank prime rate, discount rate, advance/decline, up/down volume, Arms index, short sales ratio
Brownstone (1996)	FTSE index, exchange rate, interest rate, futures market, etc.
Cogger et al. (1997)	Index return, day of week variable, holiday variable
Darrat et al. (2000)	Stock prices
Deboeck et al. (2000)	Earning per share, relative strength, % stocks held by funds, % stocks owned, shares outstanding, market capitalization, debt-to-equity
Dorsey et al. (1998)	S and P 500 return, S and P 500 index, DJIA index, NYSE volume, DJ 20 Bond, DJUA, standard deviation
Dropsy (1996)	Excess return, government spending ratio, money growth rate, interest rate, trade balance to GDP ratio, inflation rate, currency depreciation
Hochreiter et al. (1997)	Industrial production/money supply, business sentiment, foreign order, tendency ratio, interest rate, div rate, DAX index, RSI, MACD
Thawornwong and Enke(2004)	Nominal Standard and Poor's 500 index, Nominal dividends per share, Annualized average of bid and ask yields, Monthly holding period return, Nominal stock returns, Excess stock returns, Dividend yield, 3-month T-bill rate, 6-month T-bill rate, 1-year T-bill rate, 5-year T-bill constant maturity rate, 10-year T-bill constant maturity rate, 1-month certificate of deposit rate, 3-month certificate of deposit rate, 6-month certificate of deposit rate, Moody's seasoned Aaa corporate bond yield, Moody's seasoned Baa corporate bond yield, Producer Price Index, Industrial Production Index, Consumer Price Index, M1 Money Stock,
Rapach et al. (2005)	relative money market rate, relative 3-month Treasury bill rate, relative long-term government bond yield, term spread, inflation rate, industrial production growth, narrow money growth, broad money growth, change in the unemployment rate, real stock return.
Lam (2004)	Current Assets/Current Liabilities, Net Sales/Total Assets, Net Income/Net Sales, (Long-Term Debt + Short-Term Debt)/Total Assets, Total Sources of Fund/Total Uses of Fund, Research Expense, Pretax Income/Net Sales, Current Assets/Common Shareholders' Equity, Common Shares Traded, Capital Expenditure, Earnings per Share, Dividend per Share, Depreciation Expense, Tax Deferral and Investment Credit, Market Capitalization, Relative Strength Index, Federal Budget/Gross Domestic Product, Government Spending/Gross Domestic Product, Money Supply 1, Money Supply 2, Short-Term Interest Rate, Consumer Price Index, Trade Balance/Gross Domestic Product, Effective Exchange Rate, Purchase Price of Crude Oil
Pierdzioch et al. (2008)	industrial production, orders inflow, The 3-month treasury bill rate, A January dummy, The IFO overall business climate indicator (IFO; WGIFOMXLE), The level of the VDAX-NEW index
Chen (2009)	Nominal returns, Term spreads (3M-10Y), Term spreads (3M-5Y), Inflation rates, Industrial production growth, M1 growth, M2 growth, Changes in unemployment rates, Changes in federal funds rates, Changes in exchange rates, Changes in public debt
Abugri (2008)	3-month Treasury-bill, MSCI world index, nominal exchange rate, the money Supply (M1), industrial production index, nominal interest rate
JörgDöpke et al. (2008)	3-month Treasury-bill, term spread, oil price, January dummy, dummy variable, unemployment rate

Table 2. The selected variables.

Macroeconomic variable		Financial variable
New York S.E	**Tehran S.E**	**Both market**
OP	OP	CR
IP	IP	QR
CPI	CPI	IT
PPI	PPI	GM
T_{12}	CD_6	PM
T_{60}	CD_{12}	ITA
T_{120}	CD_{60}	DTE
M_1	M_1	DPS
M_2	M_2	EPS
ITR	ITR	ROE
IFR	IFR	NTS
ER	ER	OITA
AAA	PB	STA
BAA	UR	DTA
CD_6	GDP	OITS
UR		OITR
GDP		

The definitions of all the variables are given in the Appendix.

Table 3. Results of Kaiser-Meyer-Olkin and Bartlett's test.

Variable	Market	KMO	Bartlett	Sig.
Financial	New York	0.622	2020.298	0.000
	Tehran	0.607	1912.295	0.002
Macroeconomic	New York	0.602	1922.303	0.000
	Tehran	0.712	2310.015	0.005

Exchange and EPS variable in Tehran stock Exchange are not used in the model.

Modeling and predicting stock return using neural network

Unlike traditional statistical models, known as Box–Jenkins ARIMA (Box and Jenkins, 1970), ANNs are data-driven and non-parametric models that do not require strong model assumptions and can map any nonlinear function without a priori assumption about the properties of the data.

After obtaining the principle forecasting variables, they are then used in building the BPNN forecasting model. For building BPNN forecasting model, the neural network toolbox of MATLAB software is adapted in this study. Since one hidden layer network is sufficient to model any complex system with desired accuracy (Chauvin and Rumelhart, 1995), the designed BPNN model in this study will have only one hidden layer. The performance of BPNN is mainly affected by the setting of network topology, that is, the number of nodes in each layer and learning rates. There are no general rules for the choice of network topology. The selection is usually based on the trial-and-error (or called cross-validation) method. In this study, the optimal network topology of the BPNN models is determined by the trial-and-error method.

In the modeling of the BPNN model, the input layer has seven nodes as seven forecasting variables are used. Since there are no general rules for the choice of the number of nodes in the hidden layer, the number of hidden nodes to be tested was set to 11, 12, 13 and 14. And the network has only one output node, the forecasted stock return. As lower learning rates tended to give the best network results (Chauvin and Rumelhart, 1995), learning rates 0.01, 0.02, 0.03, 0.04 and 0.05 are tested during the training process. The network topology

Table 4. Results of PCA technique for financial variables in New York S.E.

Variable	Eigen value	Proportion	Cumulative
CR	3.647	22.794	22.794
ROE	2.030	12.686	35.480
DPS	1.776	11.099	46.579
OITA	1.360	8.499	55.078
EPS	1.208	7.550	62.628
QR	1.077	6.731	69.359
DTA	1.017	6.357	75.716
STA	0.954	5.961	81.677
ITA	0.814	5.087	86.764
STA	0.620	3.554	90.318
NTS	0.510	3.189	93.507
DTE	0.431	2.692	96.199
OITR	0.346	1.539	97.738
PM	0.222	1.358	99.123
GM	0.210	0.487	99.610
IT	0.199	0.390	100.00

Table 5. Results of PCA technique for macroeconomic variables in New York S.E.

Variable	Eigen value	Proportion	Cumulative
IFR	2.115	20.150	20.150
IP	1.725	11.010	31.16
M_1	1.310	8.125	39.285
ER	1.010	7.205	46.490
CPI	0.995	7.105	53.595
T_{12}	0.950	6.721	60.316
GDP	0.805	6.590	66.906
OP	0.765	6.450	73.356
T_{60}	0.710	6.112	79.468
T_{120}	0.640	5.995	85.463
PPI	0.510	3.110	88.573
ITR	0.431	3.005	91.578
M_2	0.286	2.910	94.488
AAA	0.255	1.850	96.338
BAA	0.780	1.805	98.143
CD_6	0.620	1.127	99.270
UR	0.322	0.730	100.00

with the minimum testing MSE is considered as the optimal network. The testing results of the BPNN models with combinations of different hidden nodes and learning rates for Ney York stock exchange and Tehran stock exchange are summarized in Tables 12 and 13. Table 12 shows that the {7-12-1} topology with a learning rate of 0.04 gives the best forecasting result (minimum testing RMSE) and hence is the best topology setup for New York stock exchange BPNN forecasting model in forecasting stock return. Here, {7-12-1} represents the seven nodes in the input layer, 12 nodes in the hidden layer and one node in the output layer. As depicted in Table 13, the {7-11-1} topology with a learning rate of 0.02 is the best topology setup for the Tehran stock exchange BPNN forecasting model.

Comparing results of tow prediction techniques in both markets

Here, we will focus on using the outcome results from

Table 6. Results of PCA technique for financial variables in Tehran S.E.

Variable	Eigen value	Proportion	Cumulative
QR	3.985	27.210	27.210
PM	2.992	16.105	43.315
EPS	1.210	11.262	54.577
OITR	1.005	8.006	62.583
DPS	0.995	7.112	69.695
CR	0.980	6.514	76.209
DTE	0.976	6.211	82.420
IT	0.944	4.801	87.221
GM	0.819	4.103	91.324
OITA	0.598	3.005	94.329
NTS	0.567	2.010	96.339
DTA	0.401	1.025	97.364
OITS	0.318	0.950	98.314
ROE	0.252	0.601	98.915
ITA	0.225	0.575	99.490
STA	0.179	0.510	100.00

Table 7. Results of PCA technique for macroeconomic variables in Tehran S.E.

Variable	Eigen values	Proportion	Cumulative
IFR	2.210	21.209	21.209
M_1	1.542	14.855	36.064
PPI	1.325	8.105	44.169
UR	1.112	7.950	52.119
M_2	0.986	7.421	59.540
CD_6	0.931	6.865	66.405
CD_{60}	0.825	6.950	72.995
ITR	0.810	5.875	78.870
GDP	0.733	5.426	84.296
IP	0.612	4.325	88.621
OP	0.435	4.950	93.571
ER	0.273	3.765	97.336
CPI	0.250	1.295	98.631
CD_{12}	0.625	0.905	99.536
PB	0.315	0.464	100.00

each prediction technique and using MSE prediction error and comparing prediction power of each technique in each market. For this purpose, T-test is used. Table 14 shows values of MSE and significant of T-test in each market. It should be mentioned that both techniques in New York Stock Exchange work as predictors better than Tehran Stock Exchange. Moreover, neural network pattern in each market has much more better function compared to linear regression.

Conclusions

This paper compared the prediction power of neural network (NN) forecasting model with the multiple linear regressions (MRL) method in predicting stock return in New York stock exchange and Tehran stock exchange using prediction error. First we have focused on the selecting stocks returns' predictable factors for both stocks exchange markets among numerous financial and macroeconomic variables. Then, 33 macroeconomic and financial variables for New York stock exchange and 31 macroeconomic and financial variables for Tehran stock market were selected among various factors effective in stocks return. Furthermore, using the principal components analysis (PCA) techniques, relevant variables were identified and irrelevant and redundant variables were

Table 8. Principle variables.

New York S.E	Tehran S.E
CR	QR
ROE	PM
DPS	EPS
OITS	OITR
IFR	IFR
M_1	M_1
IP	PPI

Table 9. Results of Kolmogorov-Smirnov's test.

Financial market	Variables	Asymp. Sig.
New York S.E	CR	0.000
	ROE	0.000
	DPS	0.001
	OITS	0.009
	IFR	0.000
	M_1	0.001
	IP	0.013
Tehran S.E	QR	0.000
	PM	0.000
	EPS	0.000
	OITR	0.001
	IFR	0.001
	M_1	0.005
	PPI	0.013

Table 10. Regression analysis in New York S.E.

Coefficients[a]					
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1(Constant)	-0.471	0.147		-3.2	0.002
CR	-0.241	0.092	- 0.135	-2.631	0.009
ROE	2.701	0.606	0.497	4.445	0.000
DPS	1.21	0.312	0.305	0.615	0.056
OITS	- 0.658	0.289	- 0.172	-2.279	0.023
IFR	0.441	0.131	0.122	3.372	0.001
IP	0.837	0.131	0.378	6.376	0.000
M_1	- 0.490	0.211	- 0.115	-1.025	0.066

Table 11. Model selection results of the BPNN forecasting model for New York S.E.

Number of nodes in the hidden layer	L.R	Training MSE	Testing MSE
	0.01	0.045575	0.041318
	0.02	0.045436	0.041214
11	0.03	0.045412	0.041127
	0.04	0.044590	0.040417
	0.05	0.044327	0.040514
	0.01	0.045369	0.041211
	0.02	0.044563	0.040617
12	0.03	0.044580	0.040419
	0.04	0.044435	0.040210
	0.05	0.045005	0.041105
	0.01	0.045417	0.041850
	0.02	0.045325	0.041630
13	0.03	0.045211	0.041215
	0.04	0.044114	0.040313
	0.05	0.044014	0.040420
	0.01	0.045619	0.041411
	0.02	0.045413	0.041325
14	0.03	0.044618	0.040255
	0.04	0.044435	0.040315
	0.05	0.045410	0.041310

Table 12. Model selection results of the BPNN forecasting model for Tehran S.E.

Number of nodes in the hidden layer	L.R	Training MSE	Testing MSE
	0.01	0.095820	0.091915
	0.02	0.094005	0.090205
11	0.03	0.095315	0.091535
	0.04	0.094810	0.090905
	0.05	0.094320	0.090225
	0.01	0.095150	0.091315
	0.02	0.094920	0.090825
12	0.03	0.094705	0.090610
	0.04	0.094210	0.090333
	0.05	0.094475	0.090211
	0.01	0.095945	0.091470
	0.02	0.095610	0.091606
13	0.03	0.094605	0.090744
	0.04	0.094655	0.090413
	0.05	0.095015	0.091695
	0.01	0.095115	0.091544
	0.02	0.094730	0.090804
14	0.03	0.094844	0.090415
	0.04	0.095150	0.091355
	0.05	0.095610	0.091412

Table 13. Regression analysis in Tehran S.E.

Model	Coefficients[a] Unstandardized coefficients B	Unstandardized coefficients Std. Error	Standardized coefficients Beta	t	Sig.
1 (Constant)	-12.710	27.315		-2.200	0.045
QR	2.158	0.955	2.420	3.190	0.001
PM	5.150	3.260	0.140	2.366	0.042
EPS	- 0.220	0.115	- 0.035	-0.640	0.075
OITS	0.042	0.256	0.045	2.015	0.005
IFR	- 0.022	0.001	- 0.032	-2.277	0.047
M_1	0.235	0.009	0.190	2.995	0.001
PPI	0.490	0.211	0.115	1.985	0.005

Table 14. Comparing prediction techniques in both markets.

Financial market	MSE ANN	MSE Regression	sig
New York	23.776	0.040210	0.000
Tehran	26.891	0.090205	0.000

eliminated. After identifying the principle variables, modeling and predicting stock return in each market is performed. Considering mean square error (MSE) prediction error, the results indicated that error of predicting neural network in both stock markets is less than linear regression techniques. Moreover the results showed that both linear regression techniques and neural networks pattern in New York stock exchange predict superior to Tehran stock exchange.

REFERENCES

Abugri BA (2008). Empirical relationship between macroeconomic volatility and stock returns: Evidence from Latin American markets. Int. Rev. Financ. Anal., 17: 396–410

Amihud Y (2002). Liquidity and Stock Returns: Cross-Section and Time Series Effects, J. Financ. Mark., 5: 31-56.

Archarya VV, Pedersen LH (2005). Asset pricing with Liquidity risk. J. Financ. Econ., 77: 375-410.

Atsalakis GS, Valavanis KP (2009). Surveying stock market forecasting techniques – Part II: Soft computing methods. Expert Syst. Appl., 36(3): 5932–5941.

Balkin SD, Ord JK (2000). Automatic neural network modeling for univariate time series. Int. J. Forecast., 16: 509-515.

Box GEP, Jenkins GM (1970).Time aeries analysis, forecasting and control. San Francisco: Holden-Day.

Cao Q, Parry ME (2009). Neural network earnings per share forecasting models: A comparison of backward propagation and the genetic algorithm. DecIS. Support Syst., 47(1): 32–41.

Chang KL (2009). This study applies various regime switching GJR-GARCH models to analyze the effects of macroeconomic variables (interest rate, dividend yield, and default premium). Econ. Model., 26: 1283–1299.

Chang PC, Liu CH, Lin JL, Fan CY, Ng CSP (2009). A neural network with a case based dynamic window for stock trading prediction. Expert Syst. Appl., 36(3): 6889–6898.

Chauvin Y, Rumelhart DE (1995). Backpropagation: Theory, architectures, and applications. New Jersey: Lawrence Erlbaum Associates.

Chen CH (1994). Neural networks for financial market prediction. Proceedings of the IEEE Int. Conf. Neural Netw., 2: 1199–1202.

Cheng B, Titterington DM (1994). Neural network: A review from a statistical perspective (with discussion). Stat. Sci., 9: 2–54.

Coakley JR, Brown CE (2000). Artificial neural networks in accounting and finance: Modeling issues. Int. J. Intel. Syst. Account. Financ. Manage., 9: 119–144.

De Faria A, Marcelo P, Gonzalez JL, Cavalcante JTP, Albuquerque MP (2009).Predicting the Brazilian stock market through neural networks and adaptive exponential smoothing methods. Expert Syst. Appl., 36: 12506-12509.

Enke D, Thawornwong S (2005). The use of data mining and neural networks for forecasting stock market returns. Expert Syst. Appl., 29: 927–940.

Haykin S (2001). Neural networks, principles and practice. 119: 1293 - 1302.

Huang W, Lai KK, Nakamori Y, Wang SY, Yu L (2007). Neural networks in finance and economics forecasting. Int. J. Inform. Technol. Decis. Mak., 6: 113-140.

Huarng K, Yu TH (2006).The application of neural networks to forecast fuzzy time series. Phys. A, 363: 481-491.

Jost A (1993). Neural networks: a logical progression in credit and marketing decision system. Credit World, 81(4): 26–33.

Keim D, Stambaugh R (1986). Predicting returns in the stock and bond markets. J. Financ. Econ., 17: 357-390.

Lewellen J (2004). Predicting returns with financial ratios. J. Financ. Econ., 74 (2): 209–235.

Lu CJ (2010). Integrating independent component analysis-based denoisingscheme with neural network for stock price prediction. Dec. Support Syst., 47 (2009): 115-125.

Lu CJ, Lee TS, Chiu CC (2009). Financial time series forecasting using independent component analysis and support vector regression. Decis. Support Syst., 47: 115-125.

Olson D, Mossman C (2003). Neural network forecasts of Canadian stock returns using accounting ratios. Int. J. Forecast., 19: 453-465.

Pontiff J, Schall LD (1998). Book-to-market ratios as predictors of market returns. J. Financ. Econ., 49 (2): 141-160.

Rapach DE, Wohar ME, Rangvid J (2005). Macrovariables and international stock return predictability. Int. J. Forecast., 21 (1): 137–166.
Refenes AN, Zapranis A, Francis G (1994). Stock performance modeling using neural networks: A comparative study with regression models. Neural Netw., 7(2): 375–388.
Truxillo C (2003). Multivariate Statistical Methods: Practical Research Applications Course Notes. Cary NC: SAS Institute.
Tsang PM, Kwok P, Choy SO, Kwan R, Ng SC, Mak J (2007). Design and implementation of NN5 for Hong Kong stock price forecasting. Eng. Appl. Artif. Intel., 20(4): 453–461.
Yu THK, Huarng KH (2008). A bivariate fuzzy time series model to forecast the TAIEX. Expert Syst. Appl., 34: 2945–2952.
Zhu X, Wang H, Xu L, Li H (2008). Predicting stock index increments by neural networks: The role of trading volume under different horizons. Expert Syst. Appl., 34(4): 3043–3054.

APPENDIX

Definition of variables

OP: Oil price.
IP: Industrial production.
CPI: Consumer price index.
PPI: producer price index.
T_{12}: T-bill rate (1-year).
T_{60}: T-bill constant maturity rate (5-year).
T_{120}: T-bill constant maturity rate (10-year).
M_1: M1 Money Stock.
M_2: M2 Money Stock.
ITR: Interest rate.
IFR: Inflation rate.
ER: Exchange rate.
AAA: Moody's seasoned Aaa corporate bond yield, averages of business days.
BAA: Moody's seasoned Baa corporate bond yield, averages of business days.
CD_6: 6-month Certificate of deposit rate.
CD_{12}: 1-year Certificate of deposit rate.
CD_{60}: 5-year Certificate of deposit rate.
UR: Unemployment rate.
GDP: Gross domestic product.
CR: Current ratio.
QR: Quick ratio.
IT: Inventory turnover.
GM: Gross margin.
PM: Profit margin.
ITA: Inventory to total assets.
DTE: Long term debt to equity.
DPS: Dividends per share.
EPS: Earnings per share.
ROE: Return on equity.
NTS: Net income to total assets.
OITA: Operating Income to total assets.
STA: Total sells to total assets.
DTA: Total debts to total assets.
OITS: Operating income to total sells.
OITR: Operating income to total revenue.
PB: Participation Bonds; Participation bonds issued by the government, governmental companies, Iran central bank, municipalities and the companies registered with securities and exchange organization in Iran.