

Forecasting Stock prices of banks using artificial neural networks (GMDH)

Alireza Mousavi^a, Ali Karshenasan^{b*}

^a MA Accounting, Saveh Branch, Islamic Azad University, Saveh, Iran

^b Assistant Professor of Economics, University of Gonbad e-Kavoos, Iran

Abstract

According to the importance of shares and investment forecasts for those in the capital market and insufficient analysis of financial information regardless of other variables and parameters as well as the economic data, behavioral and social sciences of time series which are effective in banks stock price, investors are encouraged toward additional tools and available soft wares. In this study modeling and forecasting stock prices of banks using artificial neural networks (GMDH) is accomplished. Obtained results showed that performance of the models in processing simultaneous effects of turnover, value of transactions and number of transactions on bank stock price is excellent.

Keywords: banks stock prices forecasting, GMDH neural network, turnover, value of transactions, and number of transactions.

1. Introduction

Always a concern in the capital market is to earn the highest return and the lowest risk for investment, so investors having available tools and models and software, well as the information that companies provide business and by analysis of makes it possible to create the lowest risk and highest returns. So in this field profits from investing in the capital market is very important and given the capital market environment and the impact of variables such as political and economic, information and other behavioral sciences factors and so on in stock price makes it very difficult for investors. So in this non-linear and chaotic space most of the investors use all available financial software as well as the tools, templates and models to increase their share of profits in the capital market. One of the recently models of attention for prediction is smart adaptive models. These models are capable of considering input factors, the model structure and parameters of the model according to temporal changes and due to the flexibility, scalability and power of high learning in order to pursue structural interpretation, can act better in appropriate prediction. Most researchers using a variety of methods such as the Multilayer Perceptron Neural Network (Heydari and Korlooi, 2010), Fuzzy networks and genetic algorithms (Monajemmi et al., 2009) and ANFIS with ARIMA (Fahimifard et al. 2011) had predicted stock price which each method had its

own advantages and disadvantages. In the present study banks stock price is modeled using artificial neural networks (GMDH) in order to evaluating this method capability for precise prediction.

Literature Review

Banks activities in a society is very important in economic growth and in the long term help the development of the productive capacity of an economy, because bank is the main sources of investment. Performance of financial markets, particularly the banking system in recent years has gotten great importance that many economic experts, want using available tools seeks to explain the link between economic growth and financial markets (Khadem, 2013). Nowadays according to increased bank's activities competitiveness of private banks, need for more attention to the profitability of resources is undeniable (Zarif and Jahankhani, 1995). Profit and profit growth, both theoretically and empirically are considered as determinants of stock returns (Easton & Harris, 1991; Ali & Zarowin, 1992; Easton, et al, 1992; Ohlson & Juettner-Nauroth, 2005). Profit growth regards as a criterion to determine the profit situation of companies in the future (Jackson, 1996). In this approach accounting earnings is considered with respect to the stock market price that is representing expectations and forecasts of future markets and future profitability of the company (Estiwart, 1991). Accounting earnings is the most traditional performance evaluation criteria that has importance for investors, shareholders, managers, creditors and securities analysts. Accounting earnings which is calculated based on assumption of an obligation to many users of accounting information is one of the most important criteria to measure performance (Worthington and West, 2004). The credit risk occurs at equipment and completion of the integrated banking system network and single window in providing banking and credit services and integrated system of payments and credit rating of customers and bank transfers and bank customer data network security (Khajavi and Allahyari, 1999). However, value of assets is a function of the cash flow which that assets creates. It is possible that cash flows be determined in different ways that this matter result in different models of evaluation. Due to the complexity and importance of evaluation of equity in the firm valuation there are techniques and different approaches to evaluation.

Heydari and Kordloii (1999) had showed that the possibility of neural networks maneuver on network variation is too much but in regression models only by variables changes forecasting will change. Monajemi et al. (1999) stated that hybrid fuzzy neural networks model and genetic algorithms had much better forecasts and have faster and stronger strengthen the ability to predict stock prices. Aghaii et al. (1999) proved that in private banks of Iran, customer profitability measurement system is underdeveloped. Rostamiyan and Hajibabii (1999) showed that liquidity risk during the years under review had downtrend. In Saeedi and Pageh (2011) the results of regression showed that nominal interest rate is effective variable on stock returns of financial institutions and there is an inverse relationship between changes in nominal interest rates and stock returns of financial institutions. Etemadi and Saeedi (2012) suggest that investment growth will increase the intensity of the current stock value and profits for companies that have high profitability. Karimzadeh et al. (2012) stated that inflation and exchange rates had the negative impact and bank deposit interest rate and GDP had a positive effect on the index of bank stocks. Shavvalpour and Ashari (2012) there is a significant negative relationship concluded that, between risk and profit banks. Bastos (2010) evaluated the performance of feed-forward neural networks to model and forecast recovery rates of defaulted bank loans. In order to guarantee that the predictions are mapped into the unit interval, the neural networks are implemented with a logistic activation function in the output neuron. The statistical relevance of explanatory variables is assessed using the bootstrap technique. The results indicate that the variables which the neural network models use to derive their output coincide to a great extent with those that are significant in parametric regression models. Out-of-sample estimates of prediction errors suggest that neural networks may have better predictive ability than parametric regression models, provided the number of observations is sufficiently large. Carbo valvered et al. (2011)

suggest that deposit interest rates should be higher in geographic areas characterized by greater in-migration and lower for depositors at banks with greater shares of existing (or so-called “locked-in”) depositors. These hypotheses are tested using a rich data set obtained for the Spanish banking industry. Results confirm that, all else equal, banks offer higher deposit rates in territories characterized by greater in-migration, and also that they tend to offer lower rates, the larger the number of their locked-in depositors. These findings confirm the existence of the trade-off between exploiting old customers and attracting new ones. Park (2013) discussed the challenges of quantifying Chinese development activities, introduces AidData’s Media-Based Data Collection (MBDC) methodology, provided an overview of Chinese development finance in Africa as tracked by this new database, and discussed the potential and limitations of MBDC as a resource for tracking development finance. Alexanian & Karim (2013) examined the premium/discount firm characteristic that fundamentally affects the value relevance of two key accounting line items, earnings and book values. They argue that from the perspective of both the residual income and option-style valuation models, the relative valuation roles of earnings and book values differ fundamentally between firms that trade at a premium vis-à-vis discount to book value. They find that book values play a significantly more important role in equity valuation than earnings when firms trade at a discount. They also find that other known influential conditions, such as the sign of earnings or the relative levels of earnings and book value become inconsequential when the premium/discount condition of the firm is controlled for. Radic (2014) represented a specifically tailored measure of the Economic Value Added approach, based on the shadow price of equity, and is developed in order to account for specific characteristics of the Japanese banking system. This new “shareholder value measure” is then used in a dynamic panel data model as a linear function of various bank-risk, bank-specific, and macroeconomic variables. This study finds that cost efficiency gains, credit risk and bank size are the most important factors in explaining the shareholder value creation in Japanese banking. Cost efficiency changes are also found to significantly influence cost of equity capital. Fiordelisi and Molyneux (2015) examined the determinants of shareholder value creation for a large sample of European banks between 1998 and 2005. They use a dynamic panel data model where the bank’s shareholder value is a linear function of various bank-specific, industry-specific and macroeconomic variables. They show that shareholder value has a positive relationship with cost efficiency changes, while economic profits are linked to revenue efficiency changes. Credit losses, market and liquidity risk and leverage are also found to substantially influence bank performance.

Research methodology

The main objective of this study is forecasting Stock prices of banks using artificial neural networks (GMDH). This research is applied research. Input variables are value of trading stock, trading volume and number of transactions and as well output is banks stock price. Data are gathered from library resources and also active banks in Tehran Stock Exchange. Monthly data were available for 25 banks during 2000-2014. Data were time series kind. In order to modelling data Excel, GMDH SHELL 3.5.8.0 are used.

Modelling using GMDH

GMDH neural network encompass some neurons that a second-degree polynomial gets produced by linking different pairs. The network by the combination of second degree polynomials of all neurons, describes a function approximate of $\hat{f}$ with output of $\hat{y}$, for a inputs collection of $X = (X_1, X_2, X_3, \dots, X_n)$ with the lowest error in comparison with real output of y . so for M data which includes n inputs and an output, real results are shown as bellow:

$$y_i = \hat{f}(X_{i1}, X_{i2}, X_{i3}, \dots, X_{in}) \quad (i = 1, 2, \dots, M) \quad (1)$$

The network can predict output of $\hat{y}$ for every input of X according to Eq. (2):

$$\hat{y}_i = \hat{f}(X_{i1}, X_{i2}, X_{i3}, \dots, X_{in}) \quad (i = 1, 2, 3, \dots, M) \quad (2)$$

The mean square error between real and predicted values must be the lowest:

$$MSE = \frac{\sum_{i=1}^M (\hat{y}_i - y_i)^2}{M} \rightarrow \min \quad (3)$$

General shape of input and output variables linking is represented in Eq. (15):

$$y = a_0 + \sum_{i=1}^n a_i X_i + \sum_{i=1}^n \sum_{j=1}^n a_{ij} X_i X_j + \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n a_{ijk} X_i X_j X_k + \dots \quad (4)$$

This polynomial is called Ivakhnenko. In most cases a second order and two variables in the polynomial is used as bellow:

$$\hat{y} = G(X_i, X_j) = a_0 + a_1 X_i + a_2 X_j + a_3 X_i^2 + a_4 X_j^2 + a_5 X_i X_j \quad (5)$$

Unknown coefficients of a_i in Eq. (5) by regression techniques as are obtained, that difference between real output of y and calculated output of $\hat{y}$ for each pair of X_i and X_j be the least. For each function of G_i (made by each neuron) coefficients of each neuron to minimize the total error in order to adapt the optimal input of all pairs, is:

$$E = \frac{\sum_{i=1}^M (y_i - G_i)^2}{M} \rightarrow \min \quad (6)$$

In basic methods of GMDH, all binary compounds (neurons) are made from n input variables and unknown coefficients of all neurons got calculated using least square. So $\binom{n}{2} = \frac{n(n-1)}{2}$ neurons are made in second layer which can be shown as Eq. (7).

$$\{(y_i, y_{ip}, y_{iq}) | (i = 1, 2, \dots, M) \& p, q \in (1, 2, \dots, M)\} \quad (7)$$

Of the second order function expressed in Eq. (5) for each M triple row is used, so these equations can be rewrite in matrix form:

$$Aa = Y \quad (8)$$

Where A is second order unknown coefficients vector represented in Eq. (5), means:

$$Y = \{a_0, a_1, \dots, a_5\} \quad (9)$$

And:

$$Y = \{y_1, y_2, y_3, \dots, y_M\}^T \quad (10)$$

From the input vectors and function form it is obvious that:

$$A = \begin{bmatrix} 1 & X_{1p} & X_{1q} & X_{1p}^2 & X_{1q}^2 & X_{1p}X_{1q} \\ 1 & X_{2p} & X_{2q} & X_{2p}^2 & X_{2q}^2 & X_{2p}X_{2q} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & X_{Mp} & X_{Mq} & X_{Mp}^2 & X_{Mq}^2 & X_{Mp}X_{Mq} \end{bmatrix} \quad (11)$$

The method of least squares multiple regression analysis, make solution of equations as bellow:

$$a = (A^T A)^{-1} A^T Y \quad (12)$$

This equation gives coefficient vectors f Eq. (5) for all triple rows of M (Narimanzadeh et al., 2002).

Research Findings

Descriptive statistics

In order to better analysis of the variables it is necessary that initially the data be analyzed in terms of statistical parameters. Descriptive statistics is the subject of regulation and classification of data, diagram, and calculate values such as facades, mean, median, and... that claims individual members of the community in question profile. In fact, descriptive statistics gives basic information about raw data that in normal mode are incomprehensible. In other word with descriptive statistics huge amount of raw data can be summarized in a way that with a glimpse on them general information be extracted.

Table2: descriptive statistics of variables

Index	Variable			
	Stock Price	Traded Stock Value	Transaction Volume	Trailers number
Mean	1948.514	4270000000	1945894.	142.3815
Middle	1607.292	1.97E+09	906613.6	66.63283
Max.	3539.105	3.39E+10	17224915	846.5714
Min.	908.5000	826653.7	573.6667	1.666667
Std.	740.8910	6.94E+09	3148776.	194.4922
n	1450	1450	1450	1450

As it is in table2 during 2000-2014 years average stock price was 1948.514, average of traded stock value per month was 4270000000 and average transaction volume was 1945894. These shows that during the time interval stock price market was not stagnant.

GMDH neural network modeling method

In this section predicting model is fitted. Forecasting is based on three independent variables of traded stock value, transaction volume and trailers number. Figure4 represents GMDH SHELL 3.5.8 software output with the aim of banks stock price forecasting. Grey lines shows actual values, blue lines shows model estimated values and red lines shows predicted values. After fitting the proper model and predicting the dependent variable evaluating whole model is available. As in the prediction models, target is estimation of one variable, so it is expected that a model be achieved which in that estimated values be as much as possible close to actual values. Therefore the aim is to minimizing error (the difference between actual and predicted values). It is obvious that the fitted model is very close to actual values so it can be said that this method is a proper model for estimating stock value.



Figure1: GMDH modelling of stock price

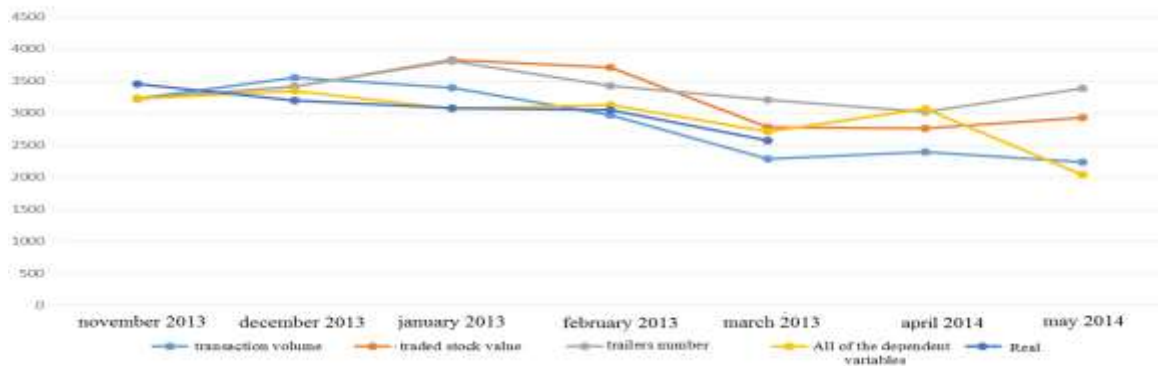


Figure2: figure1 red section magnification

Figure 2 magnifies prediction part of figure1. According to the visual guide the blue line represents actual value, light blue line represents forecasts based on the trading volume, the yellow line represents forecasts based on three components, gray line represents the prediction based on the trailers number and orange line represents forecasts based on transactions value. As it is obvious the yellow line at any 5 points comparable to the real values and in comparison with others has less error. This is normal because forecasting with three variables is better and more precise than one variable.

Conclusion

Main objective of this study was forecasting Stock prices of banks using artificial neural networks (GMDH). Modelling was based on three independent variables of traded stock value, transaction volume and trailers number and models output was banks stock price. Obtained results showed that both modelling methods have capability to predicting banks stock price using traded stock value, transaction volume and trailers number. As of limitations it must be noted that in this study stock price got modelled using tree variable of traded stock value, transaction volume and trailers number but it is obvious that there are many other variables that influence the stock price. It is suggested that potential variables get recognized and be ranked based on their importance beside related data must be very precise and as much as possible includes long run information and as many as possible more banks and financial institutions. It is recommended that in next studies more modelling methods including other neural networks, Fuzzy methods, FARIMA and radial basis functions get checked.

References

- Aleksanyan, M., & Karim, K. (2013). Searching for value relevance of book value and earnings: a case of premium versus discount firms. *Review of Quantitative Finance and Accounting*, 41(3), 489-511.
- Bastos, J. (2010). Predicting bank loan recovery rates with neural networks. CEMAPRE, ISEG, Technical University of Lisbon.
- Beyranvand, H. (2004). Integrated marketing communication process, monthly plan Year Issue 146.
- Carbo-Valverde, S., Hannan, T. H., & Rodriguez-Fernandez, F. (2011). Exploiting old customers and attracting new ones: The case of bank deposit pricing. *European Economic Review*, 55(7), 903-915.
- Daii, S., Sharifi, H., Ghasemiyan, L. (2013). The effect of macroeconomic variables on stock prices of banks "economic .mjlh, 13, 11, 12, 65-90.
- Fiordelisi, F., & Molyneux, P. (2010). The determinants of shareholder value in European banking. *Journal of Banking & Finance*, 34(6), 1189-1200.
- Etemadi, H., Saeedi, M. (2013). Analysis of factors affecting the profitability of banks using credit scoring approach: A Case Study of Sepah Bank branches in Zahedan, *Journal of Public Management*, Issue 18, Pages 135-152. The relationship between the current stock value of, profits and book value, with an emphasis on the role of investment in capital assets, securities analysis, financial literacy, 6, 20, 45-61.
- Fahimifard, S., Salarpoor, M., Saboohi, M. (2011). Compare neuro-fuzzy model ANFIS be predicted with ANN network model and ARIMA autoregressive price regiment egg case study "The Economics of agriculture and development, Vol XIX, No. 74.
- Jackson, C. (1996). Rescuing gender from the poverty trap. *World Development*, 24(3), 489-504.
- Jabbarzadeh, S., Mohammadreza, A. (2010). Identify factors affecting the initial release to the public at the time of abnormal stock returns of companies listed on Tehran Stock Exchange, the *Financial Accounting*.
- Khadem, A. (2013). The impact of capital markets on economic growth in Iran (2011-1991) using principal component analysis, *Economic Journal*, Volume 13, Issue 3.

- Khajavi, S., Allahyari, H. (2009). Check explanatory power and predictive power of accounting information of listed companies in Tehran Stock Exchange. Development and investment magazine, Issue 4, Page 50.
- Heydari, B., Hamidreza, K. (2010). Stock price forecasting using neural network researcher (Management) (Journal of Industrial Strategic Management);, Volume 7, Number 17; From page 49 to page 56.
- Monajjemi, A., Abzari, M., Raiiyati, A. (2009). Predict stock prices in the stock market using fuzzy neural networks and genetic algorithms and comparison with Artificial Neural Network, Journal of Quantitative Economics (former economic surveys) Volume 6, Number 3, pp. 26-1.
- Nariman-zadeh, N.; Darvizeh, A.; Darvizeh, M.; Gharababaei, H., (2002), "Modelling of Explosive Cutting Process of Plates Using GMDH-type Neural Network and Singular Value Decomposition". Journal of Materials Processing Technology, 128 (1-3): 80-87.
- Radić, N. (2015). Shareholder value creation in Japanese banking. Journal of Bankin & Finance, 52, 199-207
- Rostamiyan, F., Hajibabai, F., (2009). Bank liquidity risk measurement model using value at risk (Case Study: Bank), financial accounting and auditing research (Journal of Financial Accounting and Auditing), Volume 1, Number 3; From page 175 to page 198.
- Saeedi, H., Pageh, J. (2011). Ranking corporate bank customers' credit risk by the method of data analysis. Thesis, Bu-Ali Sina University, Faculty of Literature and Humanities Department of Economics.
- Sevai, M., Ma, A., & Wa, J. (2011), "The Reliability of Prospectus Based Profit Forecasts in Hong Kong, "Working paper, Hong Kong City University. Solutions, p. 3.
- Stewart, J (1991). Change in the management of public services. Public administration, 70(4), 499-518.
- Park, B., Tierney, M. J., Fuchs, A., Dreher, A., & Ramachandran, V. (2013). China's development finance to Africa: A media-based approach to data collection. Center for Global Development Working Paper, (323).
- Zariffard, A., Jahankhani, A. (1995). Are the directors and shareholders of the company are a good measure for measure? Financial Research, 7 and 8, pp. 41- 66.