



## Forecasting Stock Market Trends by Logistic Regression and Neural Networks: Evidence from KSA Stock Market

Dr. Makram ZAIDI<sup>1</sup>, Dr. Amina AMIRAT Grace Ofori-Abebrese<sup>2</sup>

1. Assistant professor at college of administrative sciences – Najran University (KSA), Email: [zaydi.makram@gmail.com](mailto:zaydi.makram@gmail.com)
2. Assistant professor at college of administrative sciences – Najran University (KSA), email: [amina.amirat@yahoo.fr](mailto:amina.amirat@yahoo.fr).

*Forecasting stock market trends is very vital for investors to take action for the next period for sustainable competition. It is especially important for policy makers to predict actions for development. KSA stock market is evolving rapidly. Due to increasing importance; the aim of this study is to forecast the stock market trends by using logistic model and artificial neural network. Logistic model is a type of probabilistic statistical classification model. It is also used to predict a binary response from a binary predictor, used for predicting the outcome of a categorical dependent variable (i.e., a class label) based on one or more predictor variables (features). Artificial neural networks are models which are used for forecasting because of their capabilities of pattern recognition and machine learning. Both methods are used to forecast the stock prices of upcoming period. The model has used the preprocessed data set of closing value of TASA Index. The data set encompassed the trading days from 5th April, 2007 to 1st January, 2015. Both methods give us estimation with up to 80% accuracy.*

**Keywords:** forecasting, logistic model, neural networks, stock index, trends

### INTRODUCTION

Recently predicting stock market trends is gaining more consideration, perhaps because of the fact that if the trend of the market is successfully forecasted the traders may be well directed. The profitability of trading in the stock market to a large extent rest on the predictability. More over the forecast trends of the market will support the regulators of the market in taking corrective measures.

Forecasting in stock market have been built on traditional statistical prediction methods for many decades. Linear models have been the center of such traditional statistical prediction. But, these methods have rarely proved successful owing to the presence of non-linearity and noise in the time series. The fruitful use of nonlinear methods in other fields of research has sparked the hopes of financials. Nonlinear methods suggests an

\*Corresponding author: Dr. Makram ZAIDI,  
Assistant professor at college of administrative sciences – Najran University (KSA)  
E-Mail: [zaydi.makram@gmail.com](mailto:zaydi.makram@gmail.com)

innovative way of observing stock prices, and it proposes new methods for empirically measuring their nature. This new technique proposes that past prices help determine future prices, but not in a straightforward way.

Due to increasing importance; the aim of this study is to forecast the stock market trends by using logistic model and artificial neural network. Logistic model is a type of probabilistic statistical classification model. It is also used to predict a binary response from a binary predictor, used for predicting the outcome of a categorical dependent variable (i.e., a class label) based on one or more predictor variables (features). Artificial neural networks are models which are used for forecasting because of their capabilities of pattern recognition and machine learning. Both methods are used to forecast the stock prices of upcoming period. The model has used the preprocessed data set of closing value of TASA Index. The data set encompassed the trading days from 5th April, 2008 to 1st January, 2015.

This paper contains four sections. The first one introduce the study, in the second we review literature linked to our paper, then we explain methodology of research and present results, the forth section concludes the paper.

## LITERATURE REVIEW

Logistic Regression, which is helpful for prediction of the presence or absence of a characteristic or outcome based on values of a set of predictor variables, is a multivariate analysis model (Lee, 2004). The applications of Logistic Regression have repeatedly been used in the area of corporate finance, banking and investments. Multivariate Discriminant Analysis has been used by many researchers for the default-prediction model; Altman, being the pioneer in this work in the year 1968 while Logistic Regression was used by the Ohlson to construct the default-prediction model in 1980. The early research on default prediction focuses on classifying firms as either defaulters or non-defaulters

In predicting financial distress and bankruptcy which have been widely applied as the evaluation models providing credit-risk information, Logistic Regression was used by Ohlson (1980) which was then followed by several authors such as Zavgren

(1985). Subsequently the same trend opted by Zmijewski (1984) for a Probit Analysis.

Logistic Regression technique yields coefficients for each independent variable based on a sample of data (Huang, Chai and Peng, 2007). Logistic regression models (LRM) with two or more explanatory variables are widely used in practice (Haines and Others, 2007). The parameters of the logistic regression model are commonly estimated by maximum Likelihood (Pardo, Pardo and Pardo, 2005). The advantage of logistic regression is that, through the addition of an appropriate link function to the usual linear regression model, the variables may be either continuous or discrete, or any combination of both types, and they do not necessarily have normal distributions (Lee, 2004).

In the past, neural networks have proved to be a promising area of research in the field of finance with many applications to its credit. Neural network applications in finance include assessing the risk of mortgage loans (Collins, Ghosh and Scofield, 1988), rating the quality of corporate bonds (Dutta & Shekhar, 1988), predicting financial distress (Altman, et.al., 1994), rating of credit card applicants (Suan & Chye, 1997) and predicting bond ratings (Hatfield & White, 1996). Neural networks are also applied in areas such as credit risk and bankruptcy prediction (Suan & Chye, 1997), pricing of derivatives (Hutchinson, 1994), asset allocation and portfolio management (Refenes & Zapranis, 1994), risk management, interest rate modeling, insurance, stock valuation, exchange rate forecasting, predicting stock market prices (Refenes, 1997, McCluskey, 1993), generate buy, design of trading system (Kusturelis, 1998), prediction of option and futures prices (Burgess), etc. Currently, apart from academic researchers many practitioners in the financial market have also realized the potential of neural networks.

Hadavandi, Shavandi et al. (Hadavandi, Shavandi et al. 2010) presents an integrated approach based on genetic fuzzy systems (GFS) and artificial neural networks (ANN) for constructing a stock price forecasting expert system, in order to fulfill the central purpose of achieving best results using minimum required input data and the least complex stock market model.

Kara, Acar Boyacioglu et al (Kara, Acar Boyacioglu et al. 2010) presented the use of artificial neural networks and support vector machines in Predicting direction of stock price index movement

for the sample of the Istanbul Stock Exchange. In this study, they attempted to develop two efficient models and compared their performances in predicting the direction of movement in the daily Istanbul Stock Exchange (ISE) National 100 Index. Shen, Guo et al. (Shen, Guo et al. 2010) made use of a radial basis function neural network (RBFNN) to train data and forecast the stock indices of the Shanghai Stock Exchange. They also made use of the artificial fish swarm algorithm (AFSA) to optimize RBF.

Furthermore, Chih-Fong and Yu-Chieh (Chih-Fong and Yu-Chieh 2010) combined multiple feature selection methods to identify more representative variables for better prediction and to effectively predict stock price for investors. In particular, three well-known feature selection methods, which are Principal Component Analysis (PCA), Genetic Algorithms (GA) and decision trees (CART), are Used.

One of the recent work in this direction is that presented by Esfahanipour and Aghamiri (Esfahanipour and Aghamiri 2010). In this work, the authors developed Neuro-Fuzzy Inference System adopted on a Takagi-Sugeno-Kang (TSK) type Fuzzy Rule Based System for stock price prediction. The TSK fuzzy model applies the technical index as the input variables and the consequent part is a linear combination of the input variables

Based on the fact that stock trend prediction is regarded as a challenging task and the recent confirmation by researches that a successful feature selection method can improve the prediction accuracy of stock market, Ni, Ni et al (Ni, Ni et al. 2011) hybridizes fractal feature selection method and support vector machine to predict the direction of daily stock price index.

Bajestani and Zare (Bajestani and Zare 2011) presented a new method to forecast TAIEX based on a high-order type 2 fuzzy time series. Extra observations are modeled as type 2 fuzzy sets and fourth-order fuzzy time series in this study, while using extra observations to improve forecasting performance. The reported results indicated that the proposed model outperforms previous studies. Finally, having presented a brief review of most of the recently published work in this area, we refer the reader to the work of Atsalakis and Valavanis (Atsalakis and Valavanis 2009) for further reviews on the earlier works in this direction. This paper

(Atsalakis and Valavanis 2009) presented detailed review of earlier stock market forecasting based on soft computing techniques. Other recent works in this direction can be found in (Lean Yu, Huanhuan Chen et al. 2009; Boyacioglu and Avci 2010; Shen, Guo et al. 2010; Ni, Ni et al. 2011).

## DATA, METHODOLOGY AND RESULTS

The data consists of daily index values of the Saudi stock market (TASA). The period under consideration is 05/04/2008 to 01/01/2015. The data set consists of 1342 data points. The data has been obtained from the official web site of Saudi stock market that provides daily stock market data. The entire analysis has been done basically on the daily returns of the stock market index TASA.

### Application of Logistic Regression

Logistic regression is used in our study because we assume that the relation between variables is non-linear. Also this type of regression is preferred when the response variable is binary which means that can take only two values 1 or 0.

Logistic regression could forecast the likelihood, or the odds ratio, of the outcome based on the predictor variables, or covariates. The significance of logistic regression can be evaluated by the log likelihood test, given as the model chi-square test, evaluated at the  $p < 0.05$  level, or the Wald statistic. Logistic regression has the advantage of being less affected than discriminant analysis when the normality of the variable cannot be assumed.

It has the capacity to analyze a mix of all types of predictors [Hair, 1995]. Logistic regression, which assumes the errors are drawn from a binomial distribution, is formulated to predict and explain a binary categorical variable instead of a metric measure. In logistic regression, the dependent variable is a log odd or logit, which is the natural log of the odds.

In the logistic regression model, the relationship between  $Z$  and the probability of the event of interest is described by this link function.

$$p_i = \frac{e^{z_i}}{1+e^{z_i}} = \frac{1}{1+e^{-z_i}}$$

$$z_i = \log(p_i/1 - p_i)$$

Where

$p_i$  is the probability the  $i^{\text{th}}$  case experiences the event of interest  
 $z_i$  is the value of the unobserved continuous variable for the  $i^{\text{th}}$  case

The z value is the odds ratio. It is expressed by

$$z_i = c + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}$$

Where

$x_{ij}$  is the  $j^{\text{th}}$  predictor for the  $i^{\text{th}}$  case

$\beta_j$  is the  $j^{\text{th}}$  coefficient

P is the number of predictors

$\beta$ s are the regression coefficients that are estimated through an iterative maximum likelihood method. However, because of the subjectivity of the choice of these misclassification costs in practice, most researchers minimize the total error rate and, hence, implicitly assume equal costs of type I and type II errors [Ohlson, 1980; Zavgren, 1985].

In order to carry out logistic regression analysis, first a method is needed for classifying returns as a "1" or "0" for a given day. In this study we use a method that is simple and objective, if the value of a return in day j is above the return in day j-1, it is noted as a "1"; otherwise, it is classified as a "0".

The return was calculated using the following formula:

$$return = \frac{p_j - p_{j-1}}{p_{j-1}} \times 100$$

Where:

$p_j$  is the closing price for day j

$p_{j-1}$  is the closing price for day j-1

As mentioned, the study contains 1678 data where 1342 are used for estimating and 336 used for validating the model. For variables we have the market return as dependent variable and five independent variables. One of them is fundamental variable oil, and the reminders are technical variables: open price, higher price, lower price and turnover.

We conduct the logistic regression in two steps. The first called general model where we introduced all our independent variables. The second is called the reduced model where we use only the significant variables of the general model. The estimation is done using the software Eviews 7 and the results are summarized table 1:

The final logistic regression equation is estimated in general model by using the maximum likelihood estimation:

$$Z = 0.178794 + 186.6705 * \text{Higher} + 191.2550 * \text{Lower} + 9.394324 * \text{Oil} - 189.6573 * \text{Open} + 0.854115 * \text{Turnover}$$

Where

$$Z = \log (p / 1 - p)$$

and 'p' is the probability that the variable is 1.

We note that the statistic LR is equal to 716.5421. This statistic suppose in the null hypothesis that all coefficients are equal to zero except the constant. Here we reject this hypothesis with zero probability to be wrong. This means that our model is globally significant. To enforce the results of the LR test we make the wald test which study the same hypothesis. We found F-statistics equal to 57.49 (prob =0) and chi-square equal to 287.48 (prob=0). So we reject the null hypothesis.

Also we have McFadden R-squared equal to 0.387 which is between 0.2 and 0.4 and that means the explicative power of the model is excellent.

Regarding the results of this model, the variable lower price has the higher influence on the market return and is more important in prediction among the technical variables chose in our estimation. while the variable turnover is significantly non-significant so it will be removed in the next step which is the reduced model (table 2).

The final logistic regression equation is estimated in a reduced model by using the maximum likelihood estimation:

$$Z = 0.173331 + 196.5624 * \text{Higher} + 187.9323 * \text{Lower} + 9.484048 * \text{Oil} - 191.7596 * \text{Open}$$

Where

$$Z = \log (p / 1 - p)$$

and 'p' is the probability that the variable is 1.

In this estimation the higher price has more influence on the market return and can be more significant in forecasting stock return.

We note that the statistic LR is equal to 713.5310. This statistic suppose in the null hypothesis that all coefficients are equal to zero except the constant. Here we reject this hypothesis with zero probability to be wrong. This means that our

model is globally significant. To enforce the results of the LR test we make the wald test which study the same hypothesis. We found F-statistics equal to 71.30506 (prob =0) and chi-square equal to 285.2202 (prob=0). So we reject the null hypothesis. To measure the quality of our model we calculate McFadden R square, Cox & Snell R square, max R square and Nagelkerke adjusted R square. The results are in table 3.

These indicators show at first that the explicative power still the same after moving to the reduced model. Secondly, the Nagelkerke adjusted R square is close to 1 which indicate that the significant power of our models is great.

We can also scan the information criteria that furnish a measure of the quality of information given by the model. We can find this criteria by a combination between the log likelihood, the number of observations and the number of variables. These criteria are: AKAIKE, Schwarz and Hannan-Quin (table 4)

We note from table (4) that the criteria is the same for Akaike while Schwarz and Hannan-Quinn decreased in the reduced model which means that the later give better information.

The last step to validate the reduced model is to check its predictive quality. For this reason, we will make the test of Goodness-of-Fit Evaluation for Binary Specification: Andrews and Hosmer-Lemeshow Tests. The results are in table 5.

This test reflect that the reduced model give a better predictive quality with zero probability of error. So, we will estimate the value of y using the reduced model for our 1342 observation. The test of Expectation-Prediction Evaluation for Binary Specification shows the results in table 6.

With the reduced model, 83.83% of estimation are correct. So we validate this model and we can now interpret the coefficient and make an estimation of the reminder data.

Here, we will speak about the marginal effect of variables. The marginal effect for jth explicative variable  $x_i^{[j]}$  is defined as :

$$\frac{\partial p_i}{\partial x_i^{[j]}} = f(x_i \beta) \cdot \beta_j = \frac{e^{x_i \beta}}{(1 + e^{x_i \beta})^2} \beta_j$$

Because of the sign of  $f(x_i \beta)$  is always positive, so the sign of this derivate is the same of the

coefficient. So the positive sign indicate an increase in the probability of y to be equal to 1 which is the case for variables: higher price, lower price, and oil. While a negative sign reflect a decrease in this probability which the case for the variable open price.

For more explanation we calculate the elasticity  $\varepsilon_{p_i/x_i^{[j]}}$  as the variation in percentage of the probability  $p_i$  that  $y_i=1$  occurs due to a variation of 1% of the jth explicative variable.  $x_i^{[j]}$ . Elasticity is defined as:

$$\forall i \in [1, N] \quad \varepsilon_{p_i/x_i^{[j]}} = \frac{x_i^{[j]} \beta_j}{1 + \exp^{x_i \beta}}$$

If these variables change by 1% the occurrence probability of  $y_i=1$  decrease for all variables except the open price (see table 7).

Using this result we will estimate the value of our dependent variable for the left 336 data in order to test the performance of our model. The result are shown in table 8.

From table 8 we can deduce that the accuracy of our model is of 81.55% which is very important and can help investor to implement the best investment strategy.

### 3-2- Application Of Artificial Neural Network

A neural network is a great data modeling device that is able to capture and represent complex input/output relations. The inspiration for the development of neural network technology stemmed from the wish to develop an artificial system that could make "intelligent" tasks similar to those done by the human brain. Neural networks resemble the human brain in the following two ways: At first, a neural network acquires knowledge through learning and secondly, a neural network's knowledge is stored within inter-neuron connection strengths known as synaptic weights.

From a very broad perspective, neural networks can be used for financial decision making in one of the three ways:

1. It can be provided with inputs, which enable it to find rules relating the current state of the system being predicted to future states.

2. It can have a window of inputs describing a fixed set of recent past states and relate those to future states.

3. It can be designed with an internal state to enable it to learn the relationship of an indefinitely large set of past inputs to future states, which can be accomplished via recurrent connections.

In this paper, one model of neural network is selected among the main network architectures used in finance. The basis of the model is neuron structure as shown in Fig. 1.

These neurons act like parallel processing units. An artificial neuron is a unit that performs a simple mathematical operation on its inputs and imitates the functions of biological neurons and their unique process of learning. From Fig. 1 will have:

$$v_k = \sum_{j=1}^m x_j w_{kj} + b_k$$

The neuron output will be:

$$y_k = f(v_k)$$

Neural systems are naturally organized in layers. Layers are made up of a number of artificial neurons, which hold an activation function. Patterns are offered to the network via the input layer, which transfers to one, or more hidden layers where the actual processing is done through a system of fully or partially weighted connections. The hidden layers are then connected to an output layer where the response is output. Multilayer perceptron (MLP) is the most commonly used neural network with the back-Propagation Algorithm networks. This kind of neural networks is excellent at both prediction and classification. Using this neural network algorithm, both input and corresponding desired output data are given to the calibrating phase (figure 2).

According to Wierenga and Kluytmans (1994) ANN can be proceeded in four steps:

**Step 1:** Set the number of input and output layers, the analyst must decide how many intermediate or hidden layers, the network offers only limited possibilities of adaptation, with one hidden layer is capable, with a sufficient number of neurons to approximate all continuous function (Hornik,

1991). A second hidden layer takes into account the discontinuities or detect relationships and interactions between variables. Indeed, in our study the use of hidden layers is very useful to detect all non-linear relationships between variables in the model. After several experimentation we used four (4) layers.

**Step 2:** Determine the number of neurons in layers. Each additional neuron allows taking into account specific profiles of input neurons. A more large number therefore allows better match the data presented but decreases capacity generalization of the network. Here again no general rule but empirical rules. The size of the hidden layer must be either equal to that of the input layer (Wierenga and Kluytmans, 1994) or is equal to 75% of the latter (and Venugopal Baet, 1994) or is equal to the square root of the number of neurons in the input and output layer (Shepard, 1990).

In our study we chose the approach of setting a maximum number of nodes in each hidden layer. Then, we eliminated those have no utility for the learning procedure. So the number of input nodes is equal to number of independent variables which is 5 (higher price, lower price, open price, oil and turnover). And the number of output nodes is equal to number of dependent variables which is 1 (index returns) (figure 3)

**Step 3:** Select the transfer function. because of the dependent variable is bounded [0,1], the logistic function is chosen for both hidden layers to the output layer (tansig) (figure 4)

**Step 4:** Choose learning. The retro-propagation algorithm requires determination of the adjustment parameter of the synaptic weights in each iteration. We have retained the learning method (trainrp) which is a network training function that updates weight and bias values according to the resilient backpropagation algorithm (RPROP).

The assessment of the prediction performance of the different soft computing models was done with matlab 2013 by quantifying the prediction obtained on an independent data set. The mean absolute error (MAE) was used to study the performance of the trained forecasting models for the testing years. MAE is defined as follows:

$$MAE = \frac{1}{N} \sum_{i=1}^N [P_{predicted\ i} - P_{actual\ i}]$$

Where:

$P_{predicted\ i}$  is the predicted return on day  $i$

$P_{actual\ i}$  is actual return on day  $i$

$N$  is the number of observations

We used also Mean square error is used for evaluating the prediction accuracy of the model:

$$MSE = \frac{1}{N} \sum_{i=1}^N (P_{predicted\ i} - P_{actual\ i})^2$$

The network is trained by repeatedly presenting it with both the training and test data sets. To identify when to stop training, two parameters, namely the MAE and MSE of both the training and test sets were used. After 70 iterations, the MAE and MSE of the training set was 0.2352 and 0.1183 respectively, while those of test set was 0.2660 and 0.1368. The training was stopped after 70 iterations as there was no significant decrease in both parameters. Thus, any further training was not going to be productive or cause any significant changes. The prediction accuracy of the training data is 84.12% and that of test data is 81.84%.

From the result shown in Table 9, it is observed that the predicted values are in good agreement with exact values and the predicted error is very less. Also the results obtained clearly demonstrate that MLP method is reliable and accurate and effective for forecasting stock returns. Therefore the proposed ANN model with the developed structure shown in Table 9 can perform good prediction in stock market.

## CONCLUSION

In this paper, an attempt is made to explore the use of logistic regression to determine the factors which significantly affect the evolution of the stock index. Logistic regression method helps the investor to form an opinion about the time to invest. It may be observed that four variables i.e. open price, higher price, lower price and oil can classify up to 81.55% into two categories up and down. This prediction rate is very good, so it can be used for prediction with higher accuracy. This study has also employed neural network to predict the direction of stock index return. Multi-

layer perceptron network is trained using Tan-sigmoidal algorithm. The prediction accuracy of the model is high both for the training data (84.12%) and test data (81.84%).

We can deduce from this study that the use of logistic regression and neural network give us very close result. But, both methods give us a good results of prediction with very high accuracy using a mixture of fundamental and technical variables. So this study must be extended to use these methods in the prediction and classification of stock returns.

## REFERENCES

- Altman, E., M.Giancarlo, and F.Varentto (1994), .Corporate Distress Diagnosis: Comparisons using Linear Discriminant Analysis and Neural Networks (the Italian Experience), *Journal of Banking and Finance*, Vol.18, pp.505-29.
- Bajestani, N. S. and A. Zare (2011). "Forecasting TAIEX using improved type 2 fuzzy time series." *Expert Systems with Applications* 38(5): 5816-5821.
- Boyacioglu, M. A. and D. Avci (2010). "An Adaptive Network-Based Fuzzy Inference System (ANFIS) for the prediction of stock market return: The case of the Istanbul Stock Exchange." *Expert Systems with Applications* 37(12): 7908-7912.
- Chih-Fong, T. and H. Yu-Chieh (2010). "Combining multiple feature selection methods for stock prediction: Union, intersection, and multi-intersection approaches." *Decis. Support Syst.* 50(1): 258-269.
- Collins, E., Ghosh, S. and Svofield. C. (1998), .An Application of a multiple Neural Network Learning System to Emulation of Mortgage Underwriting Judgements., *Proceedings of the IEEE*

- International Conference on Neural Networks, Vol.2, pp.459-466.
- Dutta, A. et al. (2008). "Classification and Prediction of Stock Performance using Logistic Regression: An Empirical Examination from Indian Stock Market: Redefining Business Horizons": McMillan Advanced Research Series. p.46-62.
- Dutta, S., and Shekhar, S. (1998), .Bond Rating: A Non-conservative Application of Neural Networks., Proceedings of the IEEE International Conference on Neural networks, Vol.2, pp.443-450.
- Esfahanipour, A. and W. Aghamiri (2010). "Adapted Neuro- Fuzzy Inference System on indirect approach TSK fuzzy rule base for stock market analysis." Expert Systems with Applications 37(7): 4742-4748.
- Hadavandi, E., H. Shavandi, et al. (2010). "Integration of genetic fuzzy systems and artificial neural networks for stock price forecasting." Knowledge-Based Systems 23(8): 800-808.
- Haines, L. M. et.al. (2007). "D-optimal Designs for Logistic Regression in Two Variables, mODa 8-Advances in Model-Oriented Designed and Anaysis" : Physica-Verlag HD. p.91-98.
- Hair, Black, Babin, Anderson, Tatham (2008) "Multivariate Data Analysis" Sixth edition Pearson Education, Inc.
- Hatfield, Gay B., and White, A. Jay (1996), .Bond Rating Changes: Neural Net Estimates for Bank Holding Companies., Working Paper.
- Huang, Q. Cai, Y., Peng, J. (2007) ."Modeling the Spatial Pattern of Farmland Using GIS and Multiple Logistic Regression: A Case Study of Maotiao River Basin," Guizhou Province, China.
- Environmental Modeling and Assessment, 12(1):55-61.
- Hutchinson, James M. Andrew W. Lo and Tomaso Poggio (1994), .A Non-Parametric Approach to Pricing and Hedging Derivatives Securities Via Learning Networks. . The Journal of Finance, Vol. XLIX, No.3, July.
- Kara, Y., M. Acar Boyacioglu, et al. (2010). "Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange." Expert Systems with Applications In Press, Corrected Proof.
- Kusturelis, Jason E. (1998) .Forecasting Financial Markets using Neural Networks: An Analysis of Methods and Accuracy., Master of Science in Management Thesis, Department of System Management, <http://www.nps.navy.mil/Code36/KUTSURE.html>
- Lean Yu, Huanhuan Chen, et al. (2009). "Evolving Least Squares Support Vector Machines for Stock Market Trend Mining." IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION VOL. 13(NO. 1).
- Lee, S. (2004). "Application of Likelihood Ratio and Logistic Regression Models to Landslide Susceptibility Mapping Using GIS". Environmental Management. 34(2):223-232.
- Lee, S., Ryu, J. Kim, L. (2007). "Landslide Susceptibility Analysis and Its Verification Using Likelihood Ratio, Logistic Regression, and Artificial Neural Network Models: Case Study of Youngin, Korea". Landslides. 4: 327-338.
- McCluskey, Peter C. (1993) .Feed-forward and Recurrent Neural Networks and Genetic Programs for Stock Market and



- Time Series Forecasting., Department of Computer Science, Brown University, September.
- Ni, L.-P., Z.-W. Ni, et al. (2011). "Stock trend prediction based on fractal feature selection and support vector machine." *Expert Systems with Applications* 38(5): 5569- 5576.
- Ohlson, J. (1980). "Financial ratios and the probabilistic prediction of bankruptcy". *Journal of Accounting Research*, 18:109-31.
- Pardo,J.A., Pardo L, Pardo, M.C (2005) "Minimum  $\theta$ -divergence Estimator in logistic regression Models", *Statistical Papers*, No 47, 91-108.
- Refenes , A.N., Bentz, Y. Bunn, D.W., Burgess, A.N. and Zapranis, A.D. (1997) .Financial Time Series Modeling with Discounted Least Squares back Propagation., *Neuro Computing* Vol. 14, pp.123-138.
- Refenes, A.N. and Zapranis, A.D. (1994) .Neural Networks in Tactical Asset Allocation: A Comparative Study with Regression Models., Department of Decision Science, London Business School, London.
- Shen, W., X. Guo, et al. (2010). "Forecasting stock indices using radial basis function neural networks optimized by artificial fish swarm algorithm
- Suan, Tan Sen and Koh Hian Chye (1997), .Neural Network Applications in Accounting and Business., *Accounting and Business Review*, Vol.4, No.2, July, pp.297-317.
- Tsalakis, G. S. and K. P. Valavanis (2009). "Surveying stock market forecasting techniques - Part II: Soft computing methods." *Expert Systems with Applications* 36(3, Part 2): 5932-5941.
- Zavgren, C. 1985. Assessing the vulnerability to failure of American industrial firms: A logistic analysis, *Journal of Business Finance and Accounting* 12(1):19-45.
- Zmijewski, M.E .1984. Methodological issues related to the estimation of financial distress prediction models, *Journal of Accounting Research* 22:59-82.