



STOCK MARKET FORECASTING TECHNIQUES: A SURVEY

¹G. PREETHI, ²B. SANTHI

¹Asst Prof., Department of Computer Application, SASTRA University, Thanjavur.

²Professors, School of Computing, SASTRA University, Thanjavur.

E-mail: ¹preethi@mca.sastra.edu, ²shanthi@cse.sastra.edu

ABSTRACT

This paper surveys recent literature in the area of Neural Network, Data Mining, Hidden Markov Model and Neuro-Fuzzy system used to predict the stock market fluctuation. Neural Networks and Neuro-Fuzzy systems are identified to be the leading machine learning techniques in stock market index prediction area. The Traditional techniques are not cover all the possible relation of the stock price fluctuations. There are new approaches to known in-depth of an analysis of stock price variations. NN and Markov Model can be used exclusively in the finance markets and forecasting of stock price. In this paper, we propose a forecasting method to provide better an accuracy rather traditional method.

Keywords: *Data Mining, Stock Market Prediction, Markov Model, Neuro-Fuzzy Systems, Forecasting Techniques, and Time Series Analysis.*

1. INTRODUCTION

Stock market forecasters focus on developing a successful approach for forecast/predict index values or stock prices. Ultimate aiming at earn high profit using well defined trading strategies. The vital idea to successful stock market prediction is achieving best results also minimize the inaccurate forecast the stock price. Indisputably, forecasting stock indices is very difficult because of the market volatility that needs accurate forecast model. The stock market indices are highly fluctuating that's fall the stock price or raising the stock price. Fluctuations are affecting the investor's belief. Determining more effective ways of stock market index prediction is important for stock market investor in order to make more informed and accurate investment decisions.

The survey of recent techniques such as Neural Network, Data Mining, Markov Model and Neuro-Fuzzy system offer useful tools for forecasting noisy environments like stock market. This research aims to provide intelligent techniques to forecast stock market indexes and stock market prices. A stock market index is represents the movement average of several individual stocks. Firm characteristics are not taken into consideration in the forecasting process. To overcome this limitation, the researchers could develop a model to forecast individual stock prices.

2. BRIEF HISTORY OF THE STOCK MARKET

In this section we are going to discuss about some of the basics of stock market i.e. stock market, market index, stock exchange and essential of stock markets. There are different kinds of customers with different requirements and preferences. The total market divided into number of parts are called market segmentation, from the segmentation to choose the best segment and design the different strategies for profitability serving rather than the company's competitors do.

2.1 Stock Market

A stock market is a public market for trading the company's stocks and derivative at an approved stock price. These are called securities, listed on a stock exchange as well as an investor traded privately.

In the stock market also known as secondary market is monitored by a regulatory body called SEBI (Security and Exchange Board of India). Stock market allows companies to buy and sell their shares. It depends upon the demand and supplies the prices are vary. The price will high when the demand is high, when the share is heavy to sell the decrease the price. This type of transaction is called trading and the companies, which are permitted to do the trading, are called "Listed companies".



2.2 Essential of Stock Market

Companies can raise their money with the help of stock market. This will allow businesses can trade publicly and it produces additional capital expansion to shelling shares of ownership of the company in public market. History has shown that the price of shares and other assets is important part of the economic activity. In fact, the stock market is one of the primary indicators of a country's economic strength and development. Rising share price will increase the business investments and growth of the company profitability. Share prices will affect the wealth of households and their consumption. Therefore, central banks tend to keep an eye on the control and behavior of the stock markets. Major exchanges also act as the clearinghouse for transaction, which means that they collect the orders and deliver the shares in proper time, and give guarantee payment to the seller of a security. This will reduce the risk to an individual buyer or seller on each transaction. Primary market deals with the new issues of securities and also securities are brought the share directly from the companies. But the secondary market, securities are bought and sold the shares among investors. Secondary market deals with excellent securities. This market is made of organized exchanges and it has trading floor, where orders are transmitted for exchange. All the trading of stocks are maintained and guided by the exchanges. The rules and regulations are set down by the exchanges.

3. LITERATURE REVIEW

R.K. and Pawar D.D. in [1] predicated the stock rate because it is a challenging and daunting task to find out which is more effective and accurate method so that a buy or sell signal can be generated for given stocks. Predicting stock index with traditional time series analysis has proven to be difficult an artificial neural network may be suitable for the task. Neural network has the ability to extract useful information from large set of data. In this paper the author also presented a literature review on application of artificial neural network in stock market Index prediction.

Halbert white in [2] reported some results of an on-going project using neural network modeling and learning techniques to search for and decode nonlinear regularities in asset price movements. Author, focus on case of IBM common stock daily returns. Having to deal with the salient features of economic data highlights the role to be played by

statistical inference and requires modifications to standard learning techniques that may prove useful in other contexts.

Jing Tao Yao and chew Lim tan in [3] used artificial neural networks for classification, prediction and recognition. Neural network training is an art. Trading based on neural network outputs, or trading strategy is also an art. Authors discuss a seven-step neural network prediction model building approach in this article. Pre and post data processing/analysis skills, data sampling, training criteria and model recommendation will also be covered in this article.

Tiffany Hui-Kuang and Kun-Huang Huarng in [4] used neural network because of their capabilities in handling nonlinear relationship and also implement a new fuzzy time series model to improve forecasting. The fuzzy relationship is used to forecast the Taiwan stock index. In the neural network fuzzy time series model where as in-sample observations are used for training and out-sample observations are used for forecasting. The drawback of taking all the degree of membership for training and forecasting may affect the performance of the neural network. To avoid this take the difference between observations. These reduce the range of the universe of discourse.

Akinwale adio T, Arogundade O.T and Adekoya Adebayo F in [5] examined the use of error back propagation and regression analysis to predict the untranslated and translated Nigeria Stock Market Price (NSMP). The author was used 5-j-1 network topology to adopt the five input variables. The number of hidden neurons determined the j variables during the network selection. Both the untranslated and translated statements were analyzed and compared. The Performance of translated NSMP using regression analysis or error propagation was more superior to untranslated NSMP. The result was showed on untranslated NSMP ranged for 11.3% while 2.7% for NSMP.

David Enke and Suraphan Thawornwong in [6] used machine learning for data mining to evaluate the predictive relationship of numerous financial and economic variables. Neural network model used for estimation and classification are then examined for their ability to provide an effective forecast of future values. A cross-validation technique was used to improve the generalization ability of several models. The trading strategies guided by classification models generate higher



risk-adjusted profits than the buy-and-hold strategy as well as guided by the level-estimation based forecast of the neural network and regression models. The author decides to deploy the forecast the stock dividends, transaction costs and individual-tax brackets to replicate the realistic investment practices.

K. Senthamarai Kannan, P. Sailapathi Sekar, M.Mohamed Sathik and P. Arumugam in [7] used data mining technology to discover the hidden patterns from the historic data that have probable predictive capability in their investment decisions. The prediction of stock market is challenging task of financial time series predictions. There are five Methods namely Typical price(TP), Bollinger bands, Relative strength index (RSI), CMI and MA used to analyzed the stock index. In this paper the author got the profitable signal is 84.24% using Bollinger Bands rather than MA, RSI and CMI.

Abdulsalam sulaiman olaniyi, adewole, kayoed s, Jimoh R.G in [8] used the moving average [MA] method to uncover the patterns, relationship and to extract values of variables from the database to predict the future values of other variables through the use of time series data. The advantage of the MA method is a device for reducing fluctuations and obtaining trends with a fair degree of accuracy. This techniques proven numeric forecasting method using regression analysis with the input of financial information obtained from the daily activity equities published by Nigerian stock exchange.

M. suresh Babu, N. Geethanjali and B. Sathyanarayana in [9] used the data mining techniques are able to uncover the hidden pattern, predict future trends and behaviors in financial market. Pattern matching techniques is found to be descriptive in time series analysis. In this paper, author applied ant algorithm to accommodate a flexible and dynamic pattern-matching task in time series analysis. Apart from segment size the ant to sub-time-series size affects the system performance. In this paper, the ratio was set to 1 and also the ratio reduced to obtain a better result.

Y.L. Hsieh, Don-Lin Yang and Jungpin Wu in [10] used data mining that the methods are association rule and sequential pattern mining. In association rule can be used to analyze the customer consumption behaviors and find the patterns of buying habits in the retailer business. The sequential pattern was used to help web viewers match their needs quickly but it won't

know when to buy or sell and also it does not include the time interval dimension. In this paper, the highest confidence was up to 70%. The author said that the usage of the association rule and sequential pattern mining methods are extend of causal relationship chain and improved the accuracy level.

Md. Rafiul Hassan and Baikunthu Nath in [11] used Hidden Markov Models (HMM) approach to forecasting stock price for interrelated markets. HMM was used for pattern recognition and classification problems because of its proven suitability for modeling dynamic system. The author summarized the advantage of the HMM was strong statistical foundation. It's able to handle new data robustly and computationally efficient to develop and evaluate similar patterns. The author decides to develop hybrid system using AI paradigms with HMM improve the accuracy and efficiency of forecast the stock market.

Ching-Hseue cheng, Tai-Liang chen, Liang-Ying Wei in [12] this paper proposed a hybrid forecasting model using multi-technical indicators to predict stock price trends. There are four procedures described such as select the essential technical indicators, the popular indicators based on a correlation matrix and use CDPA to minimize the entropy principle approach. Then use RST algorithm to extract linguistic rules and utilize genetic algorithm to refine the extracted rules to get better forecasting accuracy and stock return. The advantage was discovered that produce more reliable and understandable rules and forecasting rules based on objective stock data rather than subjective human judgments.

Kuang Yu Huang, Chuen-Jiuan Jane in [13] used the moving average autoregressive exogenous (ARX) prediction model is combined with grey system theory and rough set theory to create an automatic stock market forecasting and portfolio selection mechanism. Financial data were collected automatically every quarter and are input to an ARX prediction model for forecast the future trends. Clustered using a K means clustering algorithm and then supplied to a RS classification module which selects appropriate investment stocks by a decision-making rules. The advantages are combining different forecasting techniques to improve the efficiency and accuracy of automatic prediction. Efficacies of the fusion models are evaluated by comparing the forecasting accuracy of the ARX model with GM (1, 1) model. The hybrid



model provides a highly accurate forecasting performance.

Md. Rafiul Hassan, Baikunth Nath and Michael Kirley in [14] this paper author used a fusion model of HMM, ANN and GA to forecast financial market behavior. This tool was used for in depth analysis of the stock market. The daily stock price was transformed to independent sets of values that become input to HMM and GA to optimize the initial parameters of HMM. Then HMM used to identify and locate similar patterns. A weighted average of the price differences of similar patterns are obtained to prepare a forecast for the required next day. To simplify the implementation to chosen the number of states as the number attributes in the observation vectors may not suitable for some instance. The author decides to employ a new effective GA to find the best HMM architecture for the given data set.

Yi-Fan Wang, Shihmin Cheng and Mei-Hua Hsu in [15] used Markov chain concepts into fuzzy stochastic prediction of stock indexes to achieve better accuracy and confidence. In this paper examined the comparison of ANN and Markov model has major advantages. It generates high accurate result and requiring only one input of data. The first hour's stock index data was used as the input and it lead the prediction of the probable index at any given hour. This approach was not required the standard deviation of the prediction. This approach was provide not only improved profit performance also used to determine stop-losses with greater confidence.

Hsien-Lun Wong, Yi-Hsien Tu and Chi-chen Wang in [16] used ARIMA model and vector ARMA model with fuzzy time series method for forecasting. Fuzzy time series method especially heuristic model performs better forecasting ability in short-term period prediction. The ARIMA model creates small forecasting errors in longer experiment time period. In this paper, the author investigates whether the length of the interval will influence the forecasting ability of the models or not.

Samarth Agrawal, Manoj Jindal and G.N. Pillai in [17] used Adaptive neuro-fuzzy inference system for taking decisions based on the values of technical indicators. The system used weighted moving averages, divergence and RSI. RSI helps to signal overbought and oversold conditions in a security. The task of the training for ANFIS was

adjusting all the modifiable parameters to make the ANFIS output match the training data. A hybrid-learning algorithm contains both least-square method and back propagation and also used to identify the optimal values for the parameters. The author assign weights or percentage for signify the fraction of total investment that should be minimize risk and maximize profit in stock market.

George S. Atsalakis and Kimon P. Valavanis in [18] used Adaptive Neuro Fuzzy system controller used to control the stock market process model and also evaluate a variety of stocks. The Efficient Market Hypothesis was used to improve the prediction in short-term stock market trends. The result demonstrates clearly to use the proposed Rate of Return (ROR). The better returns was obtained the investor allocated assets to the risk-free government bonds once the predicted stock return turned negative. This is known as asymmetric outcomes of the stock markets. The second time the investor allocated assets to risk-free government bonds have some positive returns. This means gains from correct prediction and losses from incorrect prediction. The neuro-fuzzy system clearly demonstrates the potential for financial market prediction.

George S. Atsalakis, Emmanouil M. Dimitrakakis and Constantinos D. Zopounidis in [19] this paper suggest Wave Analysis Stock prediction based on the neuro-fuzzy. The techniques were used to forecast the trend of the stock prices and results derived. The Elliott wave principle was connected with the Fibonacci sequence, the Fibonacci sequence of numbers derived from the addition of the previous two numbers. Elliott's wave theory cannot constantly explain the market perfectly but the fuzzy estimates of the market behavior accurately to improve the stock market forecasting.

Fazel Zarandi M.H, Rezaee B, Turksen I.B and Neshat E in [20] used a type-2 fuzzy rule based expert system is developed for stock price analysis. The purposed type-2 fuzzy model applies the technical and fundamental indexes as the input variables. The model used for stock price prediction of an automotive manufactory in Asia. The output membership values were projected onto the input spaces to generate the next membership values of input variables and tuned by genetic algorithm. The type-1 method was used for inference and to increasing the robustness of the system. This method was used to robustness, flexibility and error

minimization. It is used to forecast more profitable trading in stock markets.

4. METHOD OF TIME SERIES ANALYSIS:

4.1 Random Walk

The stock market price changes have the same distribution and these are independent of each other. The stock prices are fluctuating and the financial status of a gambler can be modeled as random walk. Random walks can be used in many fields such as ecology, economics and psychology. The random walks explain the observed behavior of processes in these areas. This will serve as a fundamental model for the recorded of stochastic activity.

4.2 Moving Average

Moving average also called rolling average or rolling mean or running average is a type of finite impulse response filter used to analyze a set of data points by creating a series of averages of different subsets of the full data set in the stock market area. This is used to smooth out the short-term fluctuations with the help of time series analysis data and highlight longer-term stock market trends or cycles. This will use in technical analysis of financial data such as stock price, stock returns or trading volumes.

4.3 Regression Method

This method includes many techniques for modeling and analyzing several variables, which is used to focus on the relationship between a dependent variable and one or more independent variables. Regression analysis is widely used for prediction and forecasting, it has substantial overlap with the field of machine learning. This is used to understand which among the independent variables are related to the dependent variables and explore the relationship. The regression analysis carries out the methods are linear regression, ordinary least squares regression are parametric, which is defined in terms of finite number of unknown parameters that are estimated from the data set. This model is used for prediction even though the moderately violated data.

4.4 ARIMA Modal

This model is fitted to the time series analysis data for predict future points in the series. These models are applied in some cases where the data show evidence of nonstationarity also where an integrated part of the model can be applied to remove the nonstationarity. ARIMA models are

clearly identifiable trends such as a constant trend (i.e. zero average model), a linear trend (i.e. linear growth behavior), and a quadratic trend (i.e. quadratic growth behavior).

This work consolidated the available methods in fig (1). Limitations are highlighted in section 5.

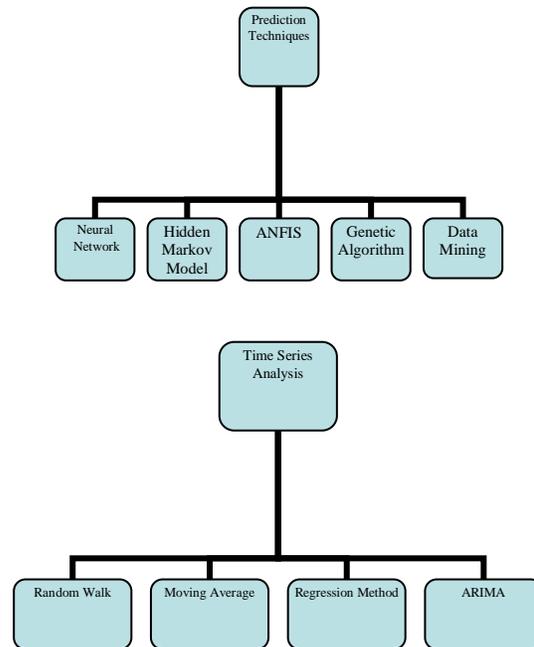


Fig. 1 Methods of Prediction Techniques and Time Series Analysis.

5. LIMITATIONS

- i) Neural Network results are unstable. The neural network functions are Block Box function. The rules of operations are completely unknown.
- ii) Back propagation networks can be take long time to train the large amount of data.
- iii) Data mining techniques are used to archive the highest confidence up to 70% only.
- iv) Unlike a regression model, ARIMA model is not support the stationary time series data.

The limitations listed above by the survey work could be overcome by the proposed methodology of combined neural network and Hidden Markov model to improve the accuracy and effectiveness of the stock market price movements. These techniques are used to predict the hours based stock price movements, which may be deployed to week and month based stock price movements.



6. PROPOSED ALGORITHM

Using the stock market database this algorithm predicts the next day stock price. Procedure as follows:

1. Open, High, Low, Close values of the daily shares are input.
2. Calculate the pivot point.
Pivot point = (High + Low + Close price)/3.
3. Require: trained HMM $\lambda = (\pi, A, B)$, likelihood update start date start_date, likelihood update range days.
4. Date \leftarrow start_date
5. O (Open, High, Low, Close) \leftarrow load (date)
6. p \leftarrow P(O/ λ) {forward Backward algorithm}, persist(date, p)
7. Date \leftarrow tomorrow (date), repeat step from 5 to 7.
8. Likelihood \leftarrow load_likelihood(yesterday)
9. Most_similar \leftarrow find(yesterday, likelihood)
10. Nextday_close_price \leftarrow load(Most_similar)
11. Calculate the share price changes.
Share price change = (today stock price – yesterday stock price)/yesterday stock price.
12. Calculate the percentage of closing price.
Percentage closing price = (Close price * percentage of share price changes)/100.
13. If current_stock_price \geq yesterday_stock_price then
SP= close price + percentage of closing price
Else
SP= close price – percentage of closing price
14. Predicted_close_price = yesterday_close_price + (nextday_close_price – most_similar.close)

Thus the above algorithms use the stock market database and predict the values.

7. CONCLUSION

This paper surveyed the Neural Network, Data mining, Neuro Fuzzy system and Markov Model in the area of stock market prediction. The NN and Markov model has ability to extract useful information from the data set so it is widely play very important role in stock market prediction. These approaches are used to control and monitor the entire the stock market price behavior and fluctuation. There are new approaches to known in-depth of an analysis of stock price variations. Markov Model can be used exclusively in the finance markets and forecasting of stock price. In this paper, we propose a forecasting method to

provide better accuracy rather traditional methods. In future work neural network and Markov modal can also explore for other applications and comparative study with other time series analysis and forecasting models.

REFERENCES:

- [1] Dase R.K. and Pawar D.D., “Application of Artificial Neural Network for stock market predictions: A review of literature” *International Journal of Machine Intelligence*, ISSN: 0975–2927, Volume 2, Issue 2, 2010, pp-14-17.
- [2] Halbert White,” Economic prediction using neural networks: the case of IBM daily stock returns” Department of Economics University of California, San Diego.
- [3] JingTao YAO and Chew Lim TAN , “Guidelines for Financial Prediction with Artificial neural networks“.
- [4] Tiffany Hui-Kuang yu and Kun-Huang Huarng, “A Neural network-based fuzzy time series model to improve forecasting”, *Elsevier*, 2010, pp: 3366-3372.
- [5] Akinwale Adio T, Arogundade O.T and Adekoya Adebayo F, “ Translated Nigeria stock market price using artificial neural network for effective prediction. *Journal of theoretical and Applied Information technology* , 2009.
- [6] David Enke and Suraphan Thawornwong, “The use of data mining and neural networks for forecasting stock market returns, 2005.
- [7] K. Senthamarai Kannan, P. Sailapathi Sekar, M.Mohamed Sathik and P. Arumugam, “Financial stock market forecast using data mining Techniques”, 2010, *Proceedings of the international multiconference of engineers and computer scientists*.
- [8] Abdulsalam sulaiman olaniyi, adewole, kayoed, Jimoh R.G, “Stock Trend Prediction using Regression Analysis – A Data Mining Approach”, *AJSS journal*, ISSN 2222-9833, 2010.
- [9] M. Suresh babu, N.Geethanjali and B. Sathyanarayana, “ Forecasting of Indian Stock Market Index Using Data Mining & Artificial Neural Network”, *International journal of advance engineering & application*, 2011.
- [10] Y.L.Hsieh, Don-Lin Yang and Jungpin Wu, “Using Data Mining to study Upstream and Downstream causal relationship in stock Market”.



- [11] Md. Rafiul Hassan and Baikunth Nath, "Stock Market forecasting using Hidden Markov Model: A New Approach", *Proceeding of the 2005 5th international conference on intelligent Systems Design and Application* 0-7695-2286-06/05, *IEEE* 2005.
- [12] Ching-Hsue cheng, Tai-Liang Chen, Liang-Ying Wei, "A hybrid model based on rough set theory and genetic algorithms for stock price forecasting", 2010, pp. 1610-1629.
- [13] Kuang Yu Huang, Chuen-Jiuan Jane, "A hybrid model stock market forecasting and portfolio selection based on ARX, grey system and RS theories", *Expert systems with Applications*, 2009, pp.5387-5392 .
- [14] Md. Rafiul Hassan, Baikunth Nath and Michael Kirley, "A fusion model of HMM, ANN and GA for stock market forecasting", *Expert systems with Applications*, 2007, pp. 171-180,
- [15] Yi-Fan Wang, Shihmin Cheng and Mei-Hua Hsu, "Incorporating the Markov chain concepts into fuzzy stochastic prediction of stock indexes", *Applied Soft Computing*, 2010, pp.613-617.
- [16] Hsien-Lun Wong, Yi-Hsien Tu and Chi-Chen Wang, "Application of fuzzy time series models for forecasting the amount of Taiwan export", *Experts Systems with Applications*, 2010, pp. 1456-1470.
- [17] Samarth Agrawal, Manoj Jindal and G.N. Pillai, "Preduction using Adaptive Neuro-Fuzzy Inference System (ANFIS)", *proceeding of the international Multiconference of engineers and computer scientists*, 2010, Vol I.
- [18] George S. Atsalakis and Kimon P.Valavanis, "Forecasting stock market short-term trends using a neuro-fuzzy based methodology", *Expert systems with Application*, 2009, pp. 10696-10707.
- [19] George S. Atsalakis, Emmanouil M.Dimitrakakis and Constantinos D. Zopounidis, "Elliot Wave Theory and neuro-fuzzy systems, in stock market predictions: The WASP system", *Expert systems with application*, 2011, pp. 9196-9206.
- [20] M.H. Fazel Zarandi, B. Rezaee, I.B. Turksen and E.Neshat, "A type-2 fuzzy rule-based experts system model for stock price analysis", *Expert systems with Applications*, 2009, pp. 139-154.