

# **An Efficient Approach to Forecast Indian Stock Market Price and their Performance Analysis**

**K.K.Sureshkumar**

Assistant Professor, Department of MCA  
Kongu Arts and Science College  
Bharathiar University  
Tamil Nadu, India

**Dr.N.M.Elango**

Professor and Head, Department of MCA  
RMK Engineering College  
Anna University of Technology  
Chennai, India

## **ABSTRACT**

Forecasting accuracy is the most important factor in selecting any forecasting methods. Research efforts in improving the accuracy of forecasting models are increasing since the last decade. The appropriate stock selections those are suitable for investment is a difficult task. The key factor for each investor is to earn maximum profits on their investments. Numerous techniques used to predict stocks in which fundamental and technical analysis are one among them. In this paper, prediction algorithms and functions are used to predict future share prices and their performance will be compared. The results from analysis shows that isotonic regression function offers the ability to predict the stock prices more accurately than the other existing techniques. The results will be used to analyze the stock prices and their prediction in depth in future research efforts.

## **General Terms**

Soft Computing and Artificial Neural Network

## **Keywords**

Artificial Neural Network, National Stock Exchange, Stock Prediction, Performance Measures.

## **1. INTRODUCTION**

Generally, stock market across the globe replicates the fluctuation of the market's economy, and attracts the attention of millions of investors. The stock market is characterized by high risk and high yield; hence investors are concerned about the analysis of the stock market and are trying to forecast the trend of the stock market. To accurately predict stock market, various prediction algorithms and models have been proposed in the literature. Forecasting is a process that produces a set of output with a set of variables [4]. The variables are normally past data of the market prices. Basically, forecasting assumes that future values are based, at least in part on these past data. Past relationships can be derived through the study and observation. Neural networks have been used to perform such task. The ideas of forecasting using neural network is to find an approximation of mapping between the input and output data through training. Later to predict the future stock value trained neural network can be used [1].

The rest of the paper is organized as follows. Section 2 reviews the literature in predicting the stock market price through Artificial Neural network. Section 3 focuses on the objectives of the research. Section 4 discusses about the National Stock Exchange (NSE), stock market, stock classification,

fundamental analysis and technical analysis. Section 5 explains the methodology of using Weka tool to predict the stock prices and calculating performance measures. Section 6 draws the conclusion of the paper.

## **2. LITERATURE REVIEW**

Prediction of stock price variation is a difficult task and the price movement behaves more like a random walk and varies with time. Since the last decade, stockbrokers and future traders have relied upon various types of intelligent systems to make trading decisions. Recently, artificial neural networks (ANNs) have been applied to predict and model stock prices [3, 12, 13]. These models, however, have their limitations owing to the tremendous noise and complex dimensionality of stock price data and besides, the quantity of data itself and the input variables may also interfere with each other [7].

Generally, there are two analytical models to predict the stock market. They are fundamental analysis and technical analysis. Fundamental analysis includes economic analysis, industry analysis and company analysis. Technical analysis is a method for predicting future price based on the past market data [8]. Prediction is made by exploiting implications hidden in past trading activities and by analyzing patterns and trends shown in price and volume charts (Smirlock and Starks, 1985; Brush 1986). Using neural networks to predict financial markets has been an active research area in both analysis, since the late 1980's (White, 1988; Fishman, Barr and Loick, 1991; Shih, 1991; Stein, 1991; Utans and Moody, 1991; Katz, 1992; Kean, 1992; Swales and Yoon, 1992; [21] Wong, 1992; Azoff, 1994; Rogers and Vemuri, 1994; Ruggerio, 1994; Baestaens, Van Den Breg and Vaudrey, 1995; Ward and Sherald, 1995; Gately, 1996; Refenes [18], Abu-Mostafa and Moody, 1996; Murphy, 1999; Qi, 1999; Virili and Reisleben, 2000; Yao and Tan, 2001; Pan, 2003a; Pan 2003b) [17]. Most of these published works are targeted at US stock markets and other international financial markets. Very little is known or done on predicting the Chinese or Asian stock markets.

## **3. OBJECTIVES OF THE STUDY**

The main objective of this paper is to use Weka 3.6.3 tool to obtain more accurate stock prediction price and to compare them with weka classifier functions such as Gaussian processes, isotonic regression, least mean square, linear regression, multilayer perceptron, pace regression, simple linear regression and SMO regression. This study can be used to reduce the error proportion in predicting the future stock prices. It increases the

chances for the investors to predict the prices more accurately by reduced error percentage and hence benefited in share markets.

#### **4. NATIONAL STOCK EXCHANGE OF INDIA (NSE)**

NSE was incorporated in November 1992, and received recognition as a stock exchange under the Securities Contracts (Regulation) Act, 1956 in April 1993. Since its inception in 1992, NSE of India has been at the vanguard of change in the Indian securities market. This period has seen remarkable changes in markets, from how capital is raised and traded, to how transactions are cleared and settled. The market has grown in scope and scale in a way that could not have been imagined at that time. Average daily trading volumes have jumped from Rs. 17 crore in 1994-95 when NSE started its Cash Market segment to Rs.16,959 crore in 2009-10. Similarly, market capitalization of listed companies went up from Rs.363,350 crore at the end of March 1995 to Rs.36,834,930 crore at end March 2011. Indian equity markets are today among the most deep and vibrant markets in the world. NSE offers a wide range of products for multiple markets, including equity shares, Exchange Traded Funds (ETF), Mutual Funds, Debt instruments, Index futures and options, Stock futures and options, Currency futures and Interest rate futures. Our Exchange has more than 1,400 companies listed in the Capital Market and more than 92% of these companies are actively traded. The debt market has 4,140 securities available for trading. Index futures and options trade on four different indices and on 223 stocks in stock futures and options as on 31<sup>st</sup> March, 2010. Currency futures contracts are traded in four currency pairs. Interest Rate Futures (IRF) contracts based on 10 year 7% Notional GOI Bond is also available for trading. The role of trading members at NSE is to the extent of providing only trading services to the investors; the Exchange involves trading members in the process of consultation and participation in vital inputs towards decision making [11].

##### **4.1 Stock Market**

A stock market index is a method of measuring a stock market as a whole. The most important type of market index is the broad-market index, consisting of the large, liquid stocks of the country. In most countries, a single major index dominates benchmarking, index funds, index derivatives and research applications. In addition, more specialized indices often find interesting applications. In India, we have seen situations where a dedicated industry fund uses an industry index as a benchmark. In India, where clear categories of ownership groups exist, it becomes interesting to examine the performance of classes of companies sorted by ownership group.

##### **4.2 Stock Classification**

Stocks are often classified based on the type of company it is, the company's value, or in some cases the level of return that is expected from the company. Below is a list of classifications which are generally known to us Growth Stocks, Value Stocks, Large Cap Stocks, Mid Cap Stocks, and Small Cap Stocks. Stocks are usually classified according to their characteristics. Some are classified according to their growth potential in the long run and the others as per their current valuations. Similarly, stocks can also be classified according to their market capitalization. S&P CNX NIFTY has NIFTY (50), Junior NIFTY (50), CNX IT (20), Bank NIFTY (12), NIFTY

Midcap50, CNX Realty (10) and CNX Infra (25). The sectoral distribution of NSE are Financial services or banks, Energy, Information Technology, Metals, Automobile, FMCG, Construction, Media & Entertainment, Pharma, Industrial Manufacturing, Cement, Fertilizers & Pesticides, Textiles, Power and Telecom [11]. Two ways of analyzing stock prices namely fundamental analysis and technical analysis are described in the next section.

##### **4.3 Fundamental Analysis**

Fundamental analysis involves analysis of a company's performance and profitability to determine its share price. By studying the overall economic conditions, the company's competition, and other factors, it is possible to determine expected returns and the intrinsic value of shares [14]. This type of analysis assumes that a share's current (and future) price depends on its intrinsic value and anticipated return on investment. As new information is released pertaining to the company's status, the expected return on the company's shares will change, which affects the stock price. So the advantages of fundamental analysis are its ability to predict changes before they show up on the charts. Growth prospects are related to the current economic environment.

##### **4.4 Technical Analysis**

Technical analysis is a method of evaluating securities by analyzing the statistics generated by market activity, such as past prices and volume. To predict the future movement and identify patterns of a stock, technical analysts use tools and charts ignoring the fundamental value [20]. Few analysts rely on chart patterns and while others use technical indicators like moving average (MA), relative strength index (RSI), on balance volume (OBV) and moving average convergence-divergence (MACD) as their benchmark. In any case, technical analyst's exclusive use of historical price and volume data is what separates them from their fundamental counterparts. Unlike fundamental analysts, technical analysts don't care whether a stock is undervalued - the only thing that matters is a security's past trading data and what information this data can provide about where the security might move in the future. Technical analysis really just studies supply and demand in a market in an attempt to determine what direction, or trend, will continue in the future.

Technical analysis studies the historical data relevant to price and volume movements of the stock by using charts as a primary tool to forecast possible price movements. According to early research, future and past stock prices were deemed as irrelevant. As a result, it was believed that using past data to predict the future stock price was impossible, and that it would only have abnormal profits. However, recent findings have proven that there was, indeed, a relationship between the past and future return rates. Furthermore, arguments have been made that by using past return rates, future return rates could also be forecasted. There are various kinds of technical indicators used in futures market as well [5]. There are 26 technical indicators which can be primarily used to analyze the stock prices. Based upon the analysis, stock trend either up or down can be predicted by the investor. For more detailed references, see [2, 9, 15, 16].

#### **5. RESEARCH METHODOLOGY**

Artificial Neural networks are potentially useful for studying the complex relationships between the input and output variables in the system. The stock exchange operations can greatly benefit

from efficient forecast techniques. In this paper, to predict the stock price of Infosys technologies a number of different prediction approaches have been applied such as Gaussian processes, isotonic regression, least mean square, linear regression, multilayer perceptron, pace regression, simple linear regression and SMO regression.

### 5.1 Prediction

The actual problem discussed in this paper is to forecast the stock price of National Stock Exchange in India. For this purpose we have used available daily stock data of Infosys Technologies (i.e., bhavcopy) from the National Stock Exchange beginning from 01-October-2007 to 12-October-2011 [10].

### 5.2 Data and Methodology

We choose consecutive 1000 day’s of NSE stock data of Infosys Technologies Company. The stock data of a company cannot be predicted only from the stock price itself. The data employed in this study consists of previous close, open price, high price and low price of Infosys Technologies. The data set encompassed the trading days from 01-October-2007 to 12-October-2011. The data is collected from the historical data available on the website of National Stock Exchange.

The task is to predict the stock price of Infosys Technologies will be up or down from today to tomorrow by using the past values of the company stock. In this research, Waikato Environment for Knowledge Analysis version 3.6.3 (Weka 3.6.3) tool is applied to preprocess the stock previous close price, open price, high price and low price data. As far as the computing equipment it is an Intel Core 2 Duo processor with 2 GB RAM and 500 GB HDD.

To derive effective input factors, Weka 3.6.3 tool is used throughout the process, this study choose 4 important attributes including previous close, open price, high price and low price. The performance of the neural network largely depends on the model of the neural network. Issues critical to the neural network modeling like selection of input variables, data preprocessing technique, network architecture design and performance measuring statistics should be considered carefully.

### 5.3 Data Preprocessing

In data preprocessing, we have consider 1000 days of Infosys technologies daily data. From this, minimum price, maximum price, mean, standard deviation, distinct and unique percentage has been found. The details are given in the table 1.

From the table 1, we found the Minimum, Maximum, Mean and Standard deviation for the four attributes. In view of the data, we find out in-between values for the stock prices. In each table we have ten classifications in the table and in the graphs each in four categories namely previous close, open price, high price and low price. If we add up the data it will have 1001 instances.

For example, consider the Table 2 prev close column containing the attribute data as 113 which means 113 days of data indicates stock price range lies between 1102.3 and 1340.21. Similarly for data 113, 155, 78, 61, 85, 127, 140, 97 and 31. The last data 31 which denotes 31 days of Infosys Technology stock price range lies between 3243.53 and 3481.45. Totally 1000 previous close data, open price, high price and low price data are represented in

the table 2 and table 3 lies within the minimum and maximum

Statistical Value of Infosys Technologies						
Attributes	Min.	Max.	Mean	Std Dev	Distinct	Unique %
Prev Close	1102.3	3481.45	2200.52	652.03	993.00	0.99
Open	1103.0	3480.00	2200.04	650.77	854.00	0.74
High	1119.6	3499.00	2232.04	650.27	894.00	0.80
Low	1040.0	3457.00	2168.27	653.36	940.00	0.89

value of Infosys technologies stock.

**Table 1. Statistical values of Infosys Technologies**

**Table 2. Data Classification of Previous Close, Open Price**

Prev Close (1102.3 - 3481.45)		Open Price (1103 -3480)	
Data	Lies Between the Price	Data	Lies Between the Price
113	1102.30 -1340.21	111	1103 -1340.70
113	1340.21-1578.13	117	1340.70 -1578.40
155	1578.13 - 1816.04	154	1578.40 -1816.10
78	1816.04 -2053.96	75	1816.10 -2053.80
61	2053.96 -2291.87	66	2053.80 -2291.50
85	2291.87 -2529.79	84	2291.50 -2529.20
127	2529.79 -2767.70	129	2529.20 -2766.90
140	2767.70 -3005.62	136	2766.90 - 3004.60
97	3005.62 -3243.53	100	3004.60 - 3242.30
31	3243.53 -3481.45	28	3242.30 - 3480

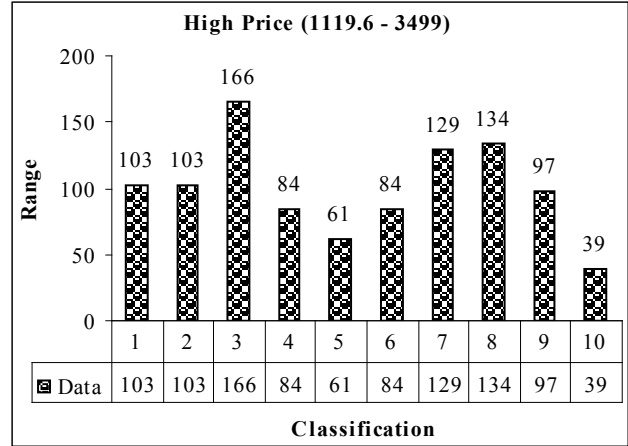
**Table 3. Data Classification of High Price, Low Price**

High Price (1119.6 - 3499)		Low Price (1040 - 3153.2)	
Data	Lies Between the Price	Data	Lies Between the Price
103	1119.60 -1357.54	86	1040 - 1259.72
103	1357.54 - 1595.48	95	1259.72 - 1479.45
166	1595.48 - 1833.42	142	1479.45 - 1699.18
84	1833.42 - 2071.36	112	1699.18 - 1918.90
61	2071.36 - 2309.30	39	1918.90 - 2138.63
84	2309.30 - 2547.24	81	2138.63 - 2358.36
129	2547.24 - 2785.18	76	2358.36 -2578.09
134	2785.18 - 3023.12	161	2578.09 -2797.81
97	3023.12 - 3261.06	108	2797.81 - 3017.54
39	3261.06 -3499	100	3017.51 -3457

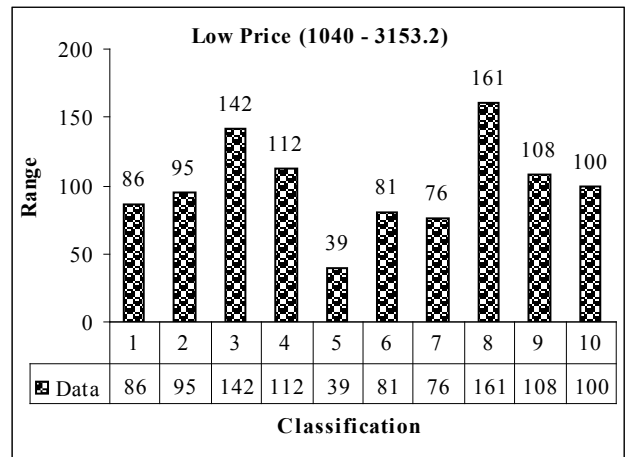
**Table 3. Data Classification of High Price, Low Price**

High Price (1119.6 - 3499)		Low Price (1040 - 3153.2)	
Data	Lies Between the Price	Data	Lies Between the Price
103	1119.60 -1357.54	86	1040 - 1259.72
103	1357.54 - 1595.48	95	1259.72 - 1479.45
166	1595.48 - 1833.42	142	1479.45 - 1699.18
84	1833.42 - 2071.36	112	1699.18 - 1918.90
61	2071.36 - 2309.30	39	1918.90 - 2138.63
84	2309.30 - 2547.24	81	2138.63 - 2358.36
129	2547.24 - 2785.18	76	2358.36 -2578.09
134	2785.18 - 3023.12	161	2578.09 -2797.81
97	3023.12 - 3261.06	108	2797.81 - 3017.54
39	3261.06 -3499	100	3017.51 -3457

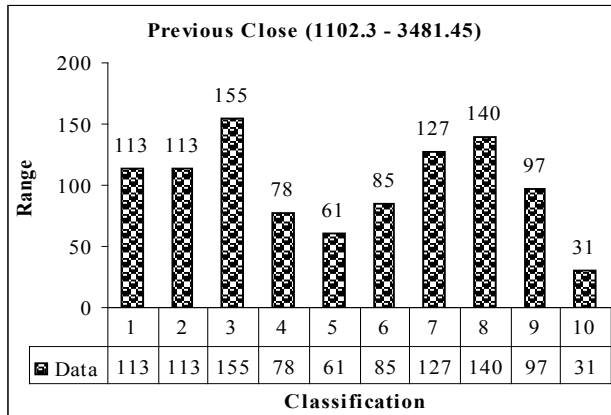
The classification of previous close, open price, high price and low price data are represented as a graph shown in Figure 1, Figure 2, Figure 3 and Figure 4.



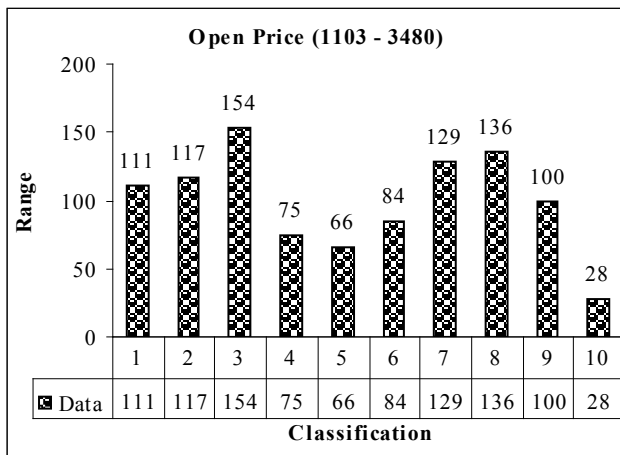
**Fig 3: Graph of High Price data classification**



**Fig 4: Graph of Low Price data classification**



**Fig 1: Graph of Previous close data classification**



**Fig 2: Graph of Open Price data classification**

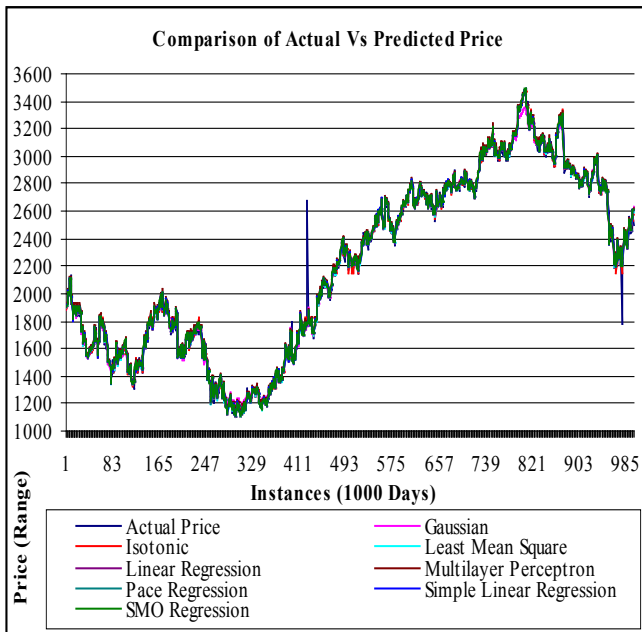
### 5.4 Results and Performance Analysis

The result of the predicted value has been shown in the Figure 5. In Figure 5, the actual stock price of Infosys technology have compared with the predicted value of the Gaussian, Isotonic regression, Least mean square, Linear regression, Multilayer Perceptron functions Pace regression, Simple linear regression and SMO regression values. Here, Instances refers to 1000 days of values taken into account for the input and the output has predicted for the same.

There are numerous methods to measure the performance of systems. In which, Root mean squared error and relative absolute error are very common in literature. In order to evaluate the net performance of the stock value four different indicators have been calculated. The indicators are mean absolute error (MAE), root mean squared error (RMSE), relative absolute error (RAE) and root relative squared error (RRSE). These indicators are used to evaluate the error rate of the stock price.

**Table 4. Comparison of Performance Measures**

Functions	Correlation coefficient	Mean absolute error	Root mean squared error	Relative absolute error	Root relative squared error	Time Taken to Build Model (S)	Total Number of Instances
Gaussian Processes	0.9976	24.5165	44.95	4.19%	6.90%	7.97	1000
Isotonic Regression	0.9985	15.1073	36.0041	2.58%	5.52%	0.52	1000
Least Mean Square	0.9982	16.6091	38.9525	2.84%	5.98%	1.84	1000
Linear Regression	0.9982	16.7118	38.8333	2.85%	5.96%	0.02	1000
Multilayer Perceptron	0.9982	24.535	42.5518	4.19%	6.53%	0.50	1000
Pace Regression	0.9982	16.7935	38.8104	2.87%	5.96%	0.03	1000
Simple Linear Regression	0.9982	16.7118	38.8333	2.85%	5.96%	0.02	1000
SMO Regression	0.9982	16.7087	38.8682	2.85%	5.96%	0.83	1000



**Fig 5: Comparison of Infosys Technology stock Actual Vs Predicted Price**

In this paper, the error percentage of the stock was evaluated by calculating MAE, RMSE, RAE and RRSE for each of the outputs for testing [22]. MAE is a quantity used to measure how close forecasts or predictions are to the eventual outcomes. Root mean squared error is widely used method to evaluate the performance of any system. RMSE is a frequently used measure of the differences between values predicted by a model or an estimator and the values actually observed from the thing being

modeled or estimated [19]. RAE is very similar to the relative squared error in the sense that it is also relative to a simple predictor, which is just the average of the actual values. In this case, though, the error is just the total absolute error instead of the total squared error. Thus, the relative absolute error takes the total absolute error and normalizes it by dividing by the total absolute error of the simple predictor. The relative squared error takes the total squared error and normalizes it by dividing by the total squared error of the simple predictor. By taking the square root of the relative squared error one reduces the error to the same dimensions as the quantity being predicted [6].

The sample data has been tested on Windows XP operating system. We have used Weka 3.6.3 tool for preprocessing and evaluation of the stock. From this analysis, it is found that the percentage of correct prediction results has shown in the above table 4. We acquire eight functions for this analysis to predict values and evaluate them. By comparing the results of the correlation coefficient values and error percentage the isotonic regression is the best suited method for predicting the stock prices.

## 6. CONCLUSIONS

In this paper, we examined and applied different neural classifier functions by using the Weka tool. By using correlation coefficient we compared various prediction functions, and found that Isotonic regression function offer the ability to predict the stock price of NSE more accurately than the other functions such as Gaussian processes, least mean square, linear regression, multilayer perceptron, pace regression, simple linear regression and SMO regression. This analysis can be used to reduce the error percentage in predicting the future stock prices. It increases the chances for the investors to predict the prices more accurately by reduced error percentage and hence increased profit in share markets.

In a highly volatile like Indian stock market, the performance levels of this various functions reported in the paper will be very useful. Especially, the prediction of the market direction with fairly high accuracy will guide the investors and the regulators. We believe that this tool gives a promising direction to the study of market predictions and their performance measures.

## 7. REFERENCES

- [1]. Abhyankar, A., Copeland, L. S., & Wong, W. (1997). "Uncovering nonlinear structure in real-time stock-market indexes: The S&P 500, the DAX, the Nikkei 225, and the FTSE-100". *Journal of Business & Economic Statistics*, 15, 1–14.
- [2]. Achelis, S.B., *Technical analysis from A to Z*, IL: Probus Publishing, Chicago, 1995.
- [3]. Aiken, M. and M. Bsar. "Forecasting Market Trends with Neural Networks." *Information Systems Management* 16 (4), 1999, 42-48.
- [4]. Atiya, A. F, El-Shoura, S. M, Shaheen, S. I and El-Sherif, M. S. "A comparison between neural network forecasting techniques case study: river flow forecasting," in *IEEE Transaction Neural Networks* pp. 402-409, 10-2, 1999.
- [5]. Brock, W, Lakonishok, J, and Lebaron, B, "Simple technical trading rules and the stochastic properties of stock returns", *Journal of Finance*, vol. 47, 1992, pp. 1731-1764.
- [6]. Carlson, W. L., Thorne, B., "Applied statistical methods", published by Prentice Hall, Inc., 1997.
- [7]. Chang, P.C., Wang, Y. W. and W. N. Yang. "An Investigation of the Hybrid Forecasting Models for Stock Price Variation in Taiwan." *Journal of the Chinese Institute of Industrial Engineering*, 21(4), 2004, pp.358-368.
- [8]. Chi, S. C., H. P. Chen, and C. H. Cheng. A Forecasting Approach for Stock Index Future Using Grey Theory and Neural Networks. *IEEE International Joint Conference on Neural Networks*, (1999), 3850-3855.
- [9]. Colby, R.W., (2003). *The Encyclopedia of Technical Market Indicators*. New York: McGraw-Hill.
- [10]. <http://www.nseindia.com/content/equities/scripvol/datafiles/01-01-2007-TO-12-10-2011INFYALLN.csv> accessed on 12/10/2011.
- [11]. National Stock Exchange of India, [http://www.nseindia.com/content/us/fact2011\\_sec1.pdf](http://www.nseindia.com/content/us/fact2011_sec1.pdf)
- [12]. Kimoto, T., and K. Asakawa Stock market prediction system with modular neural network. *IEEE International Joint Conference on Neural Network*, (1990), 1-6.
- [13]. Lee, J. W. Stock Price Prediction Using Reinforcement Learning. *IEEE International Joint Conference on Neural Networks*, (2001). 690-695.
- [14]. Mizuno, H., Kosaka, M., Yajima, H. and Komoda N., "Application of Neural Network to Technical Analysis of Stock Market Prediction", *Studies in Informatic and Control*, 1998, Vol.7, No.3, pp.111-120.
- [15]. Murphy, J.J. *Technical analysis of the Futures Markets: A Comprehensive Guide to Trading Methods and Applications*, Prentice-Hall, New York, 1986.
- [16]. Murphy, J.J. *Technical analysis of the financial markets*, Prentice-Hall, New York, 1999.
- [17]. Pan, H.P. "A joint review of technical and quantitative analysis of the financial markets towards a unified science of intelligent finance". *Proc. 2003 Hawaii International Conference on Statistics and Related Fields*, June 5-9, Hawaii, USA, 2003.
- [18]. Refenes, AN, Burgess, N and Bentz, Y. "Neural Networks in Financial Engineering: A study in methodology," in *IEEE Trans on Neural Networks*, pp. 1222-1267, 8(6), 1997.
- [19]. Refenes, Zapanis, and Francis, (1994) *Journal of Neural Networks*, "Stock Performance Modeling Using Neural Networks: A Comparative Study with Regression Models," Vol. 7, No. 2, 375-388.
- [20]. Steven B. Achelis "Technical Analysis from A-To-Z," 1st Ed. *Vision Books, NewDelhi, 2006*.
- [21]. Swales, G.S. and Yoon, Y.: "Applying artificial neural networks to investment analysis". *Financial Analysts Journal*, 1992, 48(5).
- [22]. Voelker, D, Orton, P, Adams, S, "Statistics", Published by Wiley, 2001.