A MODERN APPROACH TO PREDICT THE STOCK MARKET BEHAVIOUR FOR NAÏVE INVESTORS

[1]A. Kanimozhi, [2]Dr. B. Umadevi

[1]Research Scholar, [2]Assistant Professor & Head
[1]P.G. & Research Department of Computer Science
[1]Raja Doraisingam Govt. Arts College, Sivaganga, Tamil Nadu, India
[2]P.G. & Research Department of Computer Science
[2]Raja Doraisingam Govt. Arts College, Sivaganga, Tamil Nadu, India

---

Abstract

In investment plan, people new for the platform needs well defined guidelines as well as a concrete support for making any decisions. Without any understanding or knowledge, it will make disinterest for the market penetrator. Depth knowledge or regular market players are only capable to predict the trend. It limits the investment and controls with in a group of people. In general stock market is always vibrant and confused in nature. There are so many researches gave essential solutions for the investors in identifying the market behaviour. Even though a good guidelines must need for the naïve investors rather than regular marketer who deals with stock investments. This paper focuses on introducing a new mechanism using the Multi Layer Perceptron and Support Vector Machine. The stock indices TCS and HCL Infotech from *Nifty-Midcap50* are chosen for analysis for identifying the market behaviour.

Keywords: Naive, SVM, Classifier, MLP and Prediction.

---

1. Introduction

The stock market is essentially a non-linear, non-parametric system that is extremely hard to model with any reasonable accuracy [1]. Investors are trying to find a way to predict the stock market behaviour to decide the right stock to buy and sell. And also for making investment decision. The fundamental analysis considers the price of a stock depends on its intrinsic value and expected return on investment. Consequently, the stock price can be predicted reasonably well. Most people believe that fundamental analysis is a good method only on a long-term basis. However for short- and medium term speculations, fundamental analysis is generally not suitable.

In stock market price analysis people are following the traditional approach that based on the historical data of stock trading price and volume [2]. Technical analysis as illustrate and refers to the various methods that aims to predict future price movements using past stock prices and volume information. It is based on the hypothesis the future market directions can be determined by examining historical price data. Thus, it concludes that price trends and patterns are identified and utilized for profit. The technical analyses are highly subjective in nature and statistically invalid.

Recently, data mining techniques like decision trees, support vector machine, rough set approach, and artificial neural networks have been applied to this area. Data mining refers to extracting or mining knowledge from large [3] data stores or sets. Some of its functionalities are the discovery of concept or class descriptions, associations and correlations, classification, prediction, clustering, trend analysis, outlier and deviation analysis, and similarity analysis.

In this paper, the problem of stock market behaviour analysis and the mining algorithms that have been used to solve are discussed. The paper is structured as follows. The section one describes about introduction. Section 2 deals about background study and its related works. The methodology of the research work is explained in section three. In section four portraits the Experiment results. Finally the paper is concluded in last section.

2. Background and Related Works

Prediction of change in stock behaviour is a difficult task and the trend 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 [4].

Data miming has its own set of techniques that can be used to mine relevant and attractive knowledge from data. And mining also has several techniques such as association rule mining, classification and prediction, and clustering.

2.1 Literature Review

In this paper, diversified approaches using various data mining techniques are collected to analyse the stock behaviour and trend prediction at different levels. The study related to data mining for extracting and predicting the stock behaviour analysis used in various models and the comprehensive literature review of various researchers' works are stated below.

Pujana Paliyawan was used to predict the future market direction of the Stock Exchange of Thailand (SET). Time series forecasting was conducted and a suitable span of time for the stock market data is [5] examined. A novel approach to predict future market direction has been introduced based on chart patterns recognition by using data mining classification. Models are built through different methods including neural network, decision tree, naïve Bayes and k-nearest neighbours [6].

Soniya Dewangan, S R Tandan, Shagufta Farzana and Rohit Miri show the significant correlation between the changes in daily stock index values and the effect of financial news to the prediction of stock market prices.

Ruchi Desai and Snehal Gandhi presented a model that predicts the changes of stock trend by analyzing the influence of non- quantifiable information namely the news articles which are rich in information and superior to numeric data.

Abhinav Pathak pertained to develop a financial forecasting systems which can be used for performing an in-depth analysis of the stocks prices, downloading or importing data from the various locations and analyzing that data and producing charts to determine statistical trends. They proceed and perform a time series predictive analysis thereby predicting the h-days closing prices of a certain stock using the neural networks classification algorithm. The implementation is done using the open source software R & WEKA thereby aiming to reduce the analytics cost for any organization.

Qasem A. Al-Radaideh, Adel Abu Assaf and Eman Alnagi attempted to find to help the investors in the stock market to decide the better timing for buying or selling stocks based on the knowledge extracted from the historical prices of such stocks. The decision taken will be based on decision tree classifier which is one of the data mining techniques [7]. The CRISP-DM methodology was used over real historical data of three major companies listed in Amman Stock Exchange (ASE).

S.Prasanna and D.Ezhilmaran discussed several attempts made by researches for stock price prediction. These works show that data mining techniques can be applied for evaluation of past stock prices and acquire valuable information by estimating suitable [8] financial indicators.

3. Methodology
Data mining methodology is designed [9] to make sure that the mining effort leads to a stable model that successfully addresses the problem it is designed to solve [1]. Various data mining methodologies have been proposed to serve as blueprints for how to organize the process of gathering data, analyzing data, disseminating results, implementing results, and monitoring improvements. This methodology is proposed to analyse the nonproprietary standard process model for data mining. The following section describes the popular models which are used to predict the stock trend and behviour. The steps involved in Predicting the Stock behaviour is explained in Figure 1.

3.1 Multilayer Perceptron
Multi Layer Perceptron (MLPs) is feed forward neural networks trained with the standard back propagation algorithm. They are supervised networks so they require a desired response to be trained [10]. They learn how to transform input data into a desired response, so they are widely used for pattern classification. With one or two hidden layers, they can approximate virtually any [11] input-output map. They have been shown to approximate the performance of optimal statistical classifiers in difficult problem solving. Most neural network applications involve MLPs. The terms "Neural Network" (NN) and "Artificial Neural Network" (ANN) usually refer to a Multilayer Perceptron Network. It process the records one at a time, and "learn" by comparing their prediction of the record with the known actual record [11].

3.2 Support Vector Machine
Support Vector Machine (SVM) is a supervised [12] machine learning algorithm which can be used for both classification and regression challenges. However, it is mostly used in classification problems. Even though it's considered that [13] Neural Networks are easier to use than this, however, sometimes unsatisfactory results are obtained. A step in SVM classification [14] [15] involves identification as which are intimately connected to the known classes. This is called feature selection or feature extraction. Feature selection and SVM classification together have a use even when prediction [16] of unknown samples is not necessary [17]. They can be used to identify [18] key sets which are involved in whatever processes distinguish the classes [19].
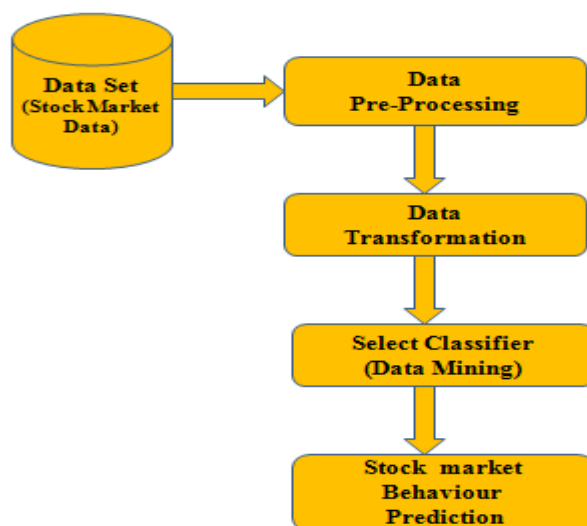
Figure 1: Steps involved in Predicting the Stock Market Behaviour

4. Experiment Results

The system evaluation on the stocks, the historical data has been collected for the period between January 2018 and June 2018 of NSE India which is given in Figure 2. The Nifty 50 companies such as TCS and HCL info technology are having the highest market capital values both have been selected for analysis.

The training can be made more efficient after certain preprocessing steps. The attributes Date and Actual Closing-Price are the two input parameters are taken for analysis. From the data set the Moving average is calculated for the actual closing price of each and every month. The date and average closing price of the index is supplied as input through the two classification algorithms such as Multi Layer Perceptron (MLP) and Support Vector Machine (SVM). The MLP Classifier that uses backpropagation to classify instances. The network can also be monitored and modified during training time. The nodes in this network are all sigmoid (except for when the class is numeric in which case the output nodes become unthresholded linear units).

The predicted Closing price is generated from the testing model and it is compared with the actual closing price. Since the closing price of the index is used to predict the market trend and behaviour in stock market analysis. The outcome of the algorithm results are given in Table 1 and Table 2 for the two index TCS and HCL InfoTech. The graphical output is shown in Figures 3 and 4.

The Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) are used to evaluate the performance of the two algorithms for the model generated. The comparison of the error accuracy is given in Figure 5.
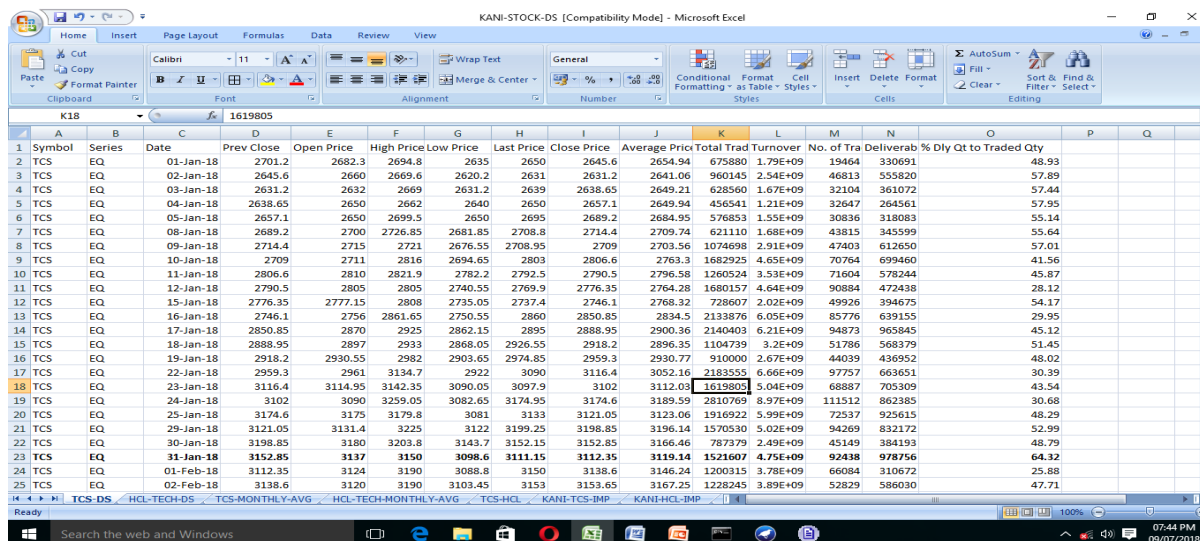
Figure 2 : Sample Data set for TCS.

Table 1 : TCS (Actual Closing-Price Vs Predicted)

Months	TCS (Actual)	TCS-MLP (Predicted)	TCS-SVM (Predicted)
Jan-18	2881.823	2881.823	2880.861
Feb-18	3009.337	3009.337	3007.059
Mar-18	2920.823	2920.823	2919.861
Apr-18	3176.336	3176.336	3174.983
May-18	3416.211	3416.211	3413.933
Jun-18	1795.9	1795.9	1797.787

Table 2 : HCL (Actual Closing-Price Vs Predicted)

Months	HCL (Actual)	HCL-MLP (Predicted)	HCL-SVM (Predicted)
Jan-18	942.1614	942.161	942.18
Feb-18	948.4658	948.466	948.447
Mar-18	950.7395	950.74	950.706
Apr-18	1008.195	1008.195	1008.145
May-18	918.4591	918.459	918.594
Jun-18	915.8619	915.862	916.013

Table 3 : Error Accuracy by Classifiers

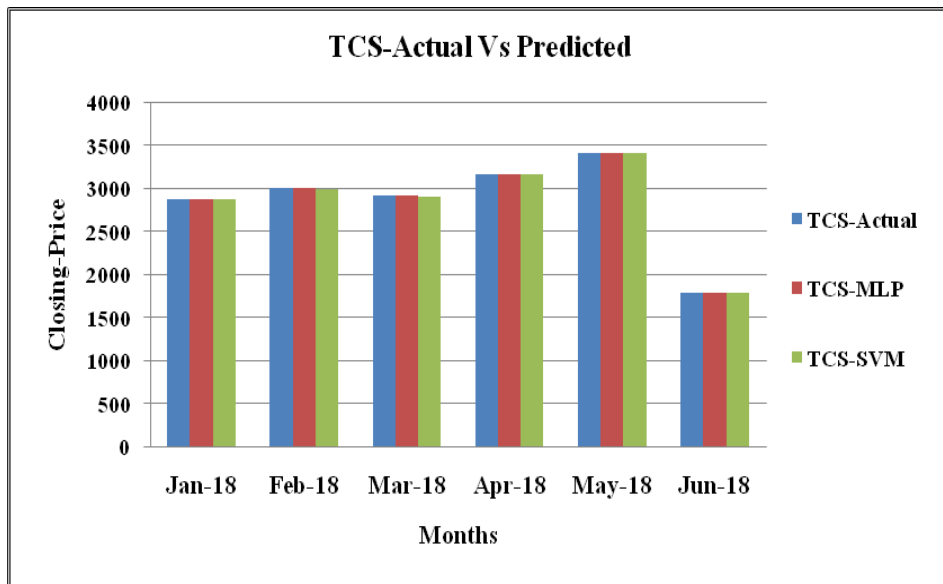| Stockindex -Classifier	MAE	RMSE
TCS-MLP	0.00%	0.00%
HCL-MLP	0.00%	0.00%
TCS-SVM	162.03%	171.40%
HCL-SVM	6.78%	8.70%

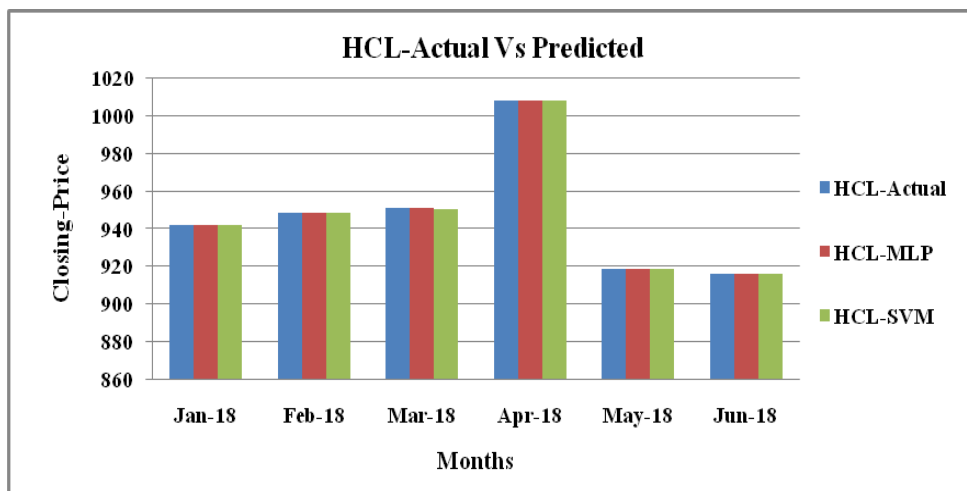**Figure 3 :** TCS (Actual Closing-Price Vs Predicted)

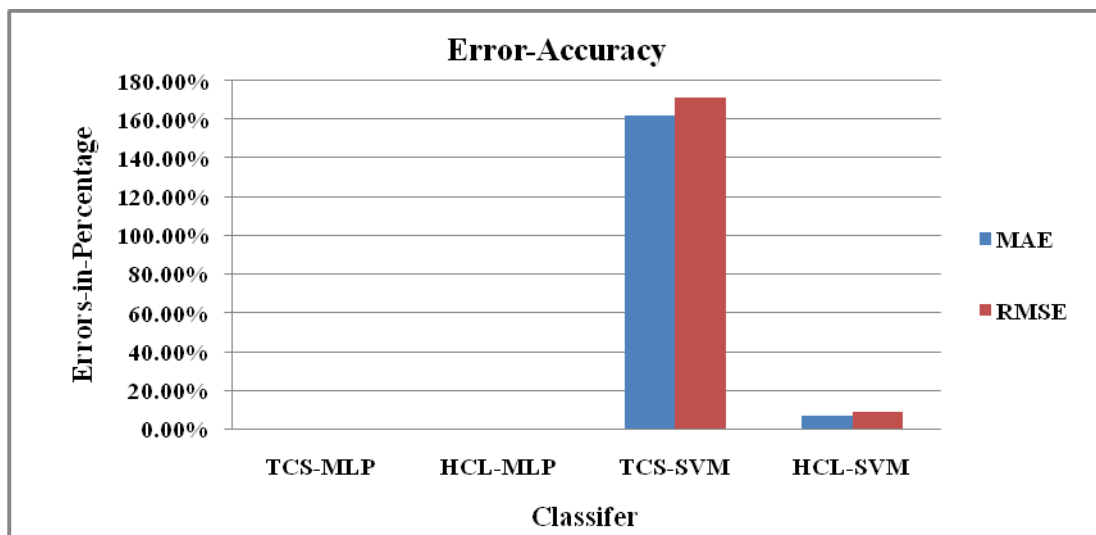**Figure 4 :** HCL (Actual Closing-Price Vs Predicted)

Figure 5 : Error Accuracy by Classifiers

5. Conclusion

The research highly supports for naïve people in making a sharp decision for their investments. The two indices such as TCS and HCL Infotech were taken for analysis. The both indices are compared through the algorithm MLP and SVM. MLP gives best result rather than SVM for TCS and HCL. The Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) are used to evaluate the performance of the two algorithms to show the error accuracy. The obtained accuracy shows that MLP has better accuracy than SVM. The TCS closing price value is much higher than HCL. So it makes a conclusion that any investor can do their investment in TCS.

References

[1] Qasem A. Al-Radaideh, Adel Abu Assaf and Eman Alnagi, 2013, "Predicting Stock Prices Using Data Mining Techniques", *The International Arab Conference on Information Technology*.

[2] Yuvraj,Wadghule .M and Sonawane V. R, 2017, "Stock Market Prediction And Forecasting Techniques: A Survey", *International Journal of Engineering Sciences & Research Technology*, ISSN: 2277-9655, Value: 3.00 CODEN: IJESS7.

[3] Roshani Gandhe, Manoj and Chaudhari.S, 2016 , "Performance Analysis Of Stock Market Prediction Techniques", *International Journal for Scientific Research & Development* , ISSN:2321-0613,Volume 4, Issue 05.

[4] Raj Kumar and Anil Balara, "Time Series Forecasting Of Nifty Stock Market Using Weka".

[5] Kamarul Faizal Hashim, Siti Hasanah Ishak and Mazida Ahmad, 2015, " A Study On Social Media Application As A Tool To Share Information During Flood Disaster", *ARPN Journal of Engineering and Applied Sciences*, Volume. 10 No. 3.

[6] Pujana Paliyawan, 2015, "Stock Market Direction Prediction Using Data Mining Classification", International Conference on Advancement in Computer Science and IT (ADVCIT2014) and *ARPN Journal of Engineering and Applied Sciences*, At Bandung, Indonesia, Volume: 10(3).

[7] Pooja Taori and Archana Nikose, "Prediction Of Future Trends In Capital Markets Using Mining Methodologies", *International Journal of Science and Research* (IJSR) ISSN : 2319-7064.

[8] Prasanna.S and Ezhilmaran Dr.D., 2013, "An Analysis On Stock Market Prediction Using Data Mining Techniques", ISSN:2229-3345, Volume 4, Issue.2.

[9] Mariam Adedoyin-Olowe , Mohamed Medhat Gaber and Frederic Stahl, "A Survey Of Data Mining Techniques For Social Media Analysis".

[10] Gulshan Kumar and Krishan Kumar, 2012, "The Use Of Multi Objective Genetic Algorithm Based Approach To Create Ensemble Of Ann For Intrusion Detection", *International Journal of Intelligence Science*, Volume.2, No.4.

[11] Gaurang Panchal , Amit Ganatra , Y P Kosta and Devyani Panchal, 2011, "Behaviour Analysis Of Multilayer Perceptrons With Multiple Hidden Neurons And Hidden Layers", *International Journal of Computer Theory and Engineering*, ISSN: 1793-8201, Volume. 3, No.2.

[12] Shivpvje Prakash.R and DeshmukhNilesh,Dr.K, 2017, "Accuracy Assessment & Analysis of Land Use & Land Cover Classification(LU/LC) using Machine Learning Algorithim SVM", *International Journal of Advance Engineering and Research Development*, e-ISSN(0):2348-4470, Volume 4, Issue 10.

[13] Sharayu Sadashiv Phule and Sharad Sawant.D, 2017 , "Abnormal Activities Detection For Security Purpose Unattainded Bag And Crowding Detection By Using Image Processing", *International Conference on Intelligent Computing and Control Systems*.

[14] Kanimozhi A. and Umadevi Dr. B., "A Survey On Trend And Behaviour Analysis Of The Stock Market", *International Journal of Creative Research Thoughts* (IJCRT), 2018 IJCR, Volume 6, Issue 2 April 2018 | ISSN: 2320-2882.

[15] Umadevi B. and Sundar D, Alli Dr P., January 2013, "An Effective Time Series Analysis for Stock Trend Prediction Using ARIMA Model for Nifty Midcap-50", *International Journal of Data Mining & Knowledge Management Process* (IJDKP),Vol.3.

[16] Umadevi B, Sundar D., Alli Dr.P., 10th June 2014, "Novel Framework For The Portfolio Determination Using PSO Adopted Clustering Technique", *Journal of Theoretical and Applied Information Technology* , Vol. 64- No.1.

[17] Umadevi B.and Sundar D., Alli Dr.P., 2013, "An Optimized Approach to Predict the Stock Market Behavior and Investment Decision Making using Benchmark Algorithms for Naive Investors",Computational Intelligence and Computing Research (ICCIC), *IEEE International Conference*, 26-28 Dec. 2013,Page(s):1 – 5978-1-4799-1594-1INSPEC Accession Number: 14061140.

[18] Harsh Singhal, Neelendra Badal, Amit Kumar Gupta, Devesh Singh Sisodia, Gautam Kumar Singh and Hemant Kumar Singh, 2013, "A Novel Approach for Load Balancing in Distributed System using FIFO-Support Vector Machine (FIFOSVM)", *International Journal of Science and Research* (IJSR) ISSN : 2319-7064 , Value : 6.14.

[19] Premalatha.M, Vijaya Lakshmi.C, 2013, " Svm Trade-Off Between Maximize The Marginand Minimize The Variables Used For Regression", *International Journal of Pure and Applied Mathematics*,ISSN: 1311-8080 , Volume: 87.

Authors' Biography

A. KANIMOZHI is an M.Phil Research Scholar in PG & Research Department Of Computer Science, Raja Doraisingam Government Arts College, Sivaganga, Tamilnadu, India. Her research concentration include in Data mining, Machine Learning and its applications. She also putlished papers in international journal.

Dr. B. UMADEVI has received her Doctoral degree in Computer Science from Manonmaniam Sundaranar University, Tirunelveli, India. Currently working as Assistant Professor & Head-P.G and Research Department of Computer Science, Raja Doraisingam Government Arts College, Sivagangai-Tamilnadu, India. She has over 23 years of Teaching Experience and published her research papers in various International, National Journals and Conferences. Her research interests include Data Mining, Soft Computing and Evolutionary Computing. She got the Best Paper Award for her publication in the IEEE International Conference on Computational Intelligence and Computing Research held on 27th Dec 2013 at VICKRAM College of Engineering and Technology.