

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. Forecasting stock return is an important financial subject that has attracted researchers' attention for many years. It involves an assumption that fundamental information publicly available in the past has some predictive relationships to the future stock returns.

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

INTRODUCTION

The stock market is essentially a non-linear, nonparametric system that is extremely hard to model with any reasonable accuracy [1]. Investors have been trying to find a way to predict stock prices and to find the right stocks and right timing to buy or sell. To achieve those objectives, and according to [2], [3-4] some research used the techniques of fundamental analysis, where trading rules are developed based on the information associated with macroeconomics, industry, and company. The authors of [5] and [6] said that fundamental analysis assumes that the price of a stock depends on its intrinsic value and expected return on investment. Analyzing the company's operations and the market in which the company is operating can do this. 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 mediumterm speculations, fundamental analysis is generally not suitable. Some other research used the techniques of technical analysis [2], in which trading rules were developed based on the historical data of stock trading price and volume. Technical analysis as illustrated in [5] and [7] refers to the various methods that aim to predict future price movements using past stock prices and volume information. It is based on the assumption that history repeats itself and that future market directions can be determined by examining historical price data. Thus, it is assumed that price trends and patterns exist that can be identified and utilized for profit. Most of the techniques used in technical analysis are highly subjective in nature and have been shown not to be statistically valid. Recently, data mining techniques and artificial intelligence techniques like decision trees, rough set approach, and artificial neural networks have been applied to this area [8]. Data mining [9] refers to extracting or mining knowledge from large 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. Data classification can be done in many different methods; one of those methods is the classification by using Decision Tree. It is a graphical representation of all possible outcomes and the paths by which they may be reached.

LETERATURE SURVEY

Over the past two decades many important changes have taken place in the environment of financial markets. The development of powerful communication and trading facilities has enlarged the scope of selection for investors. Forecasting stock return is an important financial subject that has attracted researchers' attention for many years.

It involves an assumption that fundamental information publicly available in the past has some predictive relationships to the future stock returns [10]. In order to be able to extract such relationships from the available data, data mining techniques are new techniques that can be used to extract the knowledge from this data.

For that reason, several researchers have focused on technical analysis and using advanced math and science. Extensive attention has been dedicated to the field of artificial intelligence and data mining techniques [11]. Some models have been proposed and implemented using the above mentioned techniques, the authors of [5] made an empirical study on building a stock buying/selling alert system using back propagation neural networks (BPNN), their NN was codenamed NN5. The system was trained and tested with past price data from Hong Kong and Shanghai Banking Corporation Holdings over the period from January 2004 to December 2005. The empirical results showed that the implemented system was able to predict short-term price movement directions with accuracy about 74%.

The research by [2] used decision tree technique to build on the work of Lin [12] where Lin tried to modify the filter rule that is to buy when the stock price rises k% above its past local low and sell when it falls k% from its past local high. The proposed modification to the filter rule in [12] was by combining three decision variables associated with fundamental analysis. An empirical test, using the stocks of electronics companies in Taiwan, showed Lin's method outperformed the filter rule. According to [2], in Lin's work, the criteria for clustering trading points involved only the past information; the future information was not considered at all. The research by [2] aimed to improve the filter rule and Lin's study by considering both the past and the future information in clustering the trading points. The researchers used the data of Taiwan stock market and that of NASDAQ to carry out empirical tests. Test results showed that the proposed method outperformed both Lin's method and the filter rule in the two stock markets.

The model of [11] applied the concept of serial topology and designed a new decision system, namely the twolayer bias decision tree, for stock price prediction. The methodology developed by the authors differs from other studies in two respects; first, to reduce the classification error, the decision model was modified into a bias decision model. Second, a two-layer bias decision tree is used to improve purchasing accuracy. The empirical results indicated that the presented decision model produced excellent purchasing accuracy, and it significantly outperformed than random purchase.

The authors of [10] presented an approach that used data mining methods and neural networks for forecasting stock market returns. An attempt has been made in this study to investigate the predictive power of financial and economic variables by adopting the variable relevance analysis technique in machine learning for data mining. 3 The authors examined the effectiveness of the neural network models used for level estimation and classification. The results showed that the trading strategies guided by the neural network classification models generate higher profits under the same risk exposure than those suggested by other strategies.

The research by [13] was basically a comparison between the work of Fama and French's model [14-15] and the artificial neural networks in order to try to predict the stock prices in the Chinese market. The purpose of this study is to demonstrate the accuracy of ANN in predicting stock price movement for firms traded on the Shanghai Stock Exchange. In order to demonstrate the accuracy of ANN, the authors made a comparative analysis between Fama and French's model and the predictive power of the univariate and multivariate neural network models. The results from this study indicated that artificial neural networks offer an opportunity for investors to improve their predictive power in selecting stocks, and more importantly, a simple univariate model appears to be more successful at predicting returns than a multivariate model.

Al-Haddad et al., [16] presented a study that aimed to provide evidence of whether or not the corporate governance & performance indicators of the Jordanian industrial companies listed at Amman Stock Exchange (ASE) are affected by variables that were proposed and to provide the important indicators of the relationship of corporate governance & firms' performance that can be used by the Jordanian industrial firms to solve the agency problem. The study random sample consists of (44) Jordanian industrial firms. The study finds a positive direct relationship between corporate governance and corporate performance.

Hajizadeh et al. [17] provided an overview of application of data mining techniques such as decision tree, neural network, association rules, and factor analysis and in stock markets. Prediction stock price or financial markets has been one of the biggest challenges to the AI community. Various technical, fundamental, and statistical

indicators have been proposed and used with varying results. Soni [18] surveyed some recent literature in the domain of machine learning techniques and artificial intelligence used to predict stock market movements. Artificial Neural Networks (ANNs) are identified to be the dominant machine learning technique in stock market prediction area.

El-Baky et al., [19], proposed a new approach for fast forecasting of stock market prices. The proposed approach uses new high speed time delay neural networks (HSTDNNs). The authors used the MATLAB tool to simulate results to confirm the theoretical computations of the approach.

METHODOLOGY STUDY

Data mining methodology is designed to ensure that the data mining effort leads to a stable model that successfully addresses the problem it is designed to solve. 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 [9]. To build the model that analyses the stock trends using the decision tree technique, the CRISP-DM (CrossIndustry Standard Process for data mining) [20] is used. This methodology was proposed in the mid-1990s by an European consortium of companies to serve as a nonproprietary standard process model for data mining. This model consists of the following six steps:

- i. Understanding the reason and objective of mining the stock prices.
- ii. Understanding the collected data and how it is structured.
- iii. Preparing the data that is used in the classification model.
- iv. Selecting the technique to build the model.
- v. Evaluating the model by using one of the well known evaluation methods.
- vi. Deploying the model in the stock market to predict the best action to be taken, either selling or buying the stocks.
- vii. Understanding the reason and objective of building the model

The main reason and objective of building the model is to try to help the investors in the stock market to decide the best 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 one of the data mining techniques; the decision tree classifiers.

PROPOSED SYSTEM.

As there are many disadvantages of exiting system the proposed system works by combining two methods. First one is the Back Propagation using ANN and second is Regression. Regression is one of the techniques for predicting values of any entity. But it is a statistical method and less accurate than Artificial Neural Network. Another method is using "Backpropagation Algorithm" using Artificial Intelligence. This method is much more accurate than statistical method. So we are combining above two methods using historical data. The obtained by combining above two methods will be definitely accurate and correct.

CONCLUSION

We have studied different methodologies for Stock Market Prediction which will help the investor for making the correct decision for buy or sell the stocks. Each method is having some limitation and some disadvantage. The limitations can be overcome by selecting suitable prediction techniques for specific domains. In future one can combine the two method and get proper result and output. The proposed system can get accurate result.

ACKNOWLEDGMENT

I would like to show my gratitude to the Prof. Sonwane V.R project guide for sharing their pearls of wisdom with me during the course of this research. I am also immensely grateful to Prof. Shaikh I.R for their comments on an earlier version of the manuscript, although any errors are my own and should not tarnish the reputations of these esteemed persons

REFERENCES

- [1] Wang, Y.F., (2003) "Mining stock price using fuzzy rough set system", Expert Systems with Applications, 24, pp. 13-23.
- [2] Wu, M.C., Lin, S.Y., and Lin, C.H., (2006) "An effective application of decision tree to stock trading", Expert Systems with Applications, 31, pp. 270-274.

- [3] Al-Debie, M., Walker, M. (1999). "Fundamental information analysis: An extension and UK evidence", *Journal of Accounting Research*, 31(3), pp. 261–280.
- [4] Lev, B., Thiagarajan, R. (1993). "Fundamental information analysis", *Journal of Accounting Research*, 31(2), 190–215.
- [5] Tsang, P.M., Kwok, P., Choy, S.O., Kwan, R., Ng, S.C., Mak, J., Tsang, J., Koong, K., and Wong, T.L. (2007) "Design and implementation of NN5 for Hong Kong stock price forecasting", *Engineering Applications of Artificial Intelligence*, 20, pp. 453-461.
- [6] Ritchie, J.C., (1996) *Fundamental Analysis: a Backto-the-Basics Investment Guide to Selecting Quality Stocks*. Irwin Professional Publishing.
- [7] Murphy, J.J., (1999) *Technical Analysis of the Financial Markets: a Comprehensive Guide to Trading Methods and Applications*. New York Institute of Finance.
- [8] Wang, Y.F., (2002) "Predicting stock price using fuzzy grey prediction system", *Expert Systems with Applications*, 22, pp. 33-39.
- [9] Han, J., Kamber, M., Jian P. (2011). "Data Mining Concepts and Techniques". San Francisco, CA: Morgan Kaufmann Publishers.
- [10] Enke, D., Thawornwong, S. (2005) "The use of data mining and neural networks for forecasting stock market returns", *Expert Systems with Applications*, 29, pp. 927- 940
- [11] Wang, J.L., Chan, S.H. (2006) "Stock market trading rule discovery using two-layer bias decision tree", *Expert Systems with Applications*, 30(4), pp. 605-611.
- [12] Lin, C. H. (2004) Profitability of a filter trading rule on the Taiwan stock exchange market. Master thesis, Department of Industrial Engineering and Management, National Chiao Tung University.
- [13] Cao, Q., Leggio, K.B., and Schniederjans, M.J., (2005) "A comparison between Fama and French's model and artificial neural networks in predicting the Chinese stock market", *Computers & Operations Research*, 32, pp. 2499-2512.
- [14] Fama, E.F., French, K.R., (1993) "Common risk factors in the returns on stocks and bonds", *The Journal of Finance*, 33, pp. 3-56.