

A Comparative Study on Financial Stock Market Prediction Models

¹ Mr. Amit B. Suthar, ² Ms. Hiral R. Patel, ³ Dr. Satyen M. Parikh

^{1,2} Department of Computer Science, Ganpat University

³ Dean, Faculty of Computer Applications, Ganpat University

-----Abstract-----

Now a day's investors invest money in financial products. Financial operations are not local but wide-ranging to all the countries in the world. Stock Market is the market for security where organized issuance and trading of Stocks take place either through trading or over the counter in electronic or physical form. In stock markets, many guiding principle such as price limits are made to get involved between the financial operations so that a volatility of stock prices is more uncertain than one without disturbance. Investment in the financial market is one of the best ways to obtain high rewards, but it is also a great risk among many investments. The financial market is considered as a high complex and dynamic system. So the price prediction is one of the most important issues to be investigated by researchers. The objective of the proposed paper is to do study, improvement in the machine learning approaches to predict the financial products. For financial market prediction different approaches like Artificial Intelligence, Machine Learning Techniques and various data mining techniques are used.

Keywords: - Data Mining, Genetic Algorithms, Neural Network, Prediction, Volatility.

Date of Submission: 03, December, 2012



Date of Publication: 25, December 2012

I. INTRODUCTION

Ashore market is a place of high attention to the investors as it presents them with an opportunity to help financially by investing their resources on shares and offshoot of various companies. According to chaos theorem, the behavioral characteristics of share prices are unpredictable and uncertain. So researchers were enforced to discover a technique which can estimate the effect of this vagueness to the stream of share prices. They analyses various statistical models, Artificial Neural Networks and data mining techniques which are analogous to nonparametric, nonlinear, regression models. They find the potentiality to discriminate mysterious and buried patterns in data which can be very effectual for share market prediction. There are several approaches that have been applied in order to predict financial product value [4].

For steady financial markets, many governments around the world legalize some policies to control such a dynamic environment [1]. The stock price limits are artificial boundaries set by market governors to control daily movements of security prices. Price limits are currently utilized in many stock exchanges worldwide including: Austria, Belgium, France, Italy, Japan, Korea, Malaysia, Mexico, Netherlands, Spain, Switzerland, Taiwan, and Thailand [6]. Hence, price limit mechanisms indeed affect a substantial part of capital markets around the world. This paper presents a work of imaginary tale concepts which will useful for how to make a proper decision for investing money in financial markets by combining best prediction approaches. By this paper different models which already have been given to predict the financial product value are compared. This type of comparative analysis is useful for defining the combination of best approaches for price prediction and also overcome the limitation of existing one.

The financial market value prediction has been done using different approaches. From last past decade, Neural Networks have been used in financial market value prediction. One of the first such projects was by Kimoto et al. [7] who had used neural network for the forecast of Tokyo stock exchange index. Minzuno et al. Some researchers proposed Back propagation method for training the Neural Network and Multilayer Feed forward network in order to forecast the share values. [1] Ruiling Liu, Hengjin Cai^{1*}, Cheng Luo use clustering with Manifold learning approaches on closing price of stocks according to their trends. Manifold learning is an important field of machine learning which gains very good realization in exploring inner law of nonlinear data. It assumes data is even sampled from a low dimensional manifold embedded in a high dimensional Euclidean space. The purpose of manifold learning is to learn this low dimensional manifold from high dimensional data, to get the relevant embedded mapping and finally to reduce the dimension of the data and realize visualization. [3] The characteristic that all Stock Markets have in common is the uncertainty, which is related with their short and long term future state. This feature is disagreeable for the investor but it is also necessary whenever the financial Market is selected as the investment tool. [2] So as per Random Walk, the best prediction you can have about tomorrow's value is today's value. So for the best prediction, four things are required to data-mine effectively: high-quality data, the "right" data, an adequate sample size and the right tool. [5].

II. MACHINE LEARNING APPROACHES

As the trading grew, people tried to find methods and tools which can accurately predict the share prices increasing their gains and minimizing their risk. Fundamental analysis is the physical study of a company in terms of its product sales, manpower, quality, infrastructure etc. to understand its standing in the market and thereby its profitability as an investment. The technical analysis foretells the fitting time to buy or sell a share. Technical analysts use charts which contain technical data like price, volume, uppermost and lowly prices per dealing to envisage future share whereabouts. Price charts are used to be acquainted with inclination. These trends are understood by supply and demand issues that often have recurring or some sort of perceptible patterns. [8]. The following are the different approaches used to predict financial product value.

A. Statistical Analysis

Statistical analysis will be done using regression, Co linearity and ARIMA... etc. Statistical approach is deal with the linear data. Statistical model is linear non stationary model but it will fails to model non linearity. [9] The Traditional Time Series forecasting scrutinize historical data and endeavor to estimate upcoming values of a time series as a linear amalgamation of these historical data. In econometrics there are two basic types of time series forecasting: univariate (simple regression) and multivariate (multivariate regression). These types of regression models are the most common tools used in econometrics to foretell time series. The way they are applied in practice is that first a set of factors that influence (or more specific is presumed that influence) the series under prediction is formed [2].

B. Data Mining

Data mining is the process of extracting knowledge from a database. A data-warehouse is a place where information is stored. Four things are required to mine the data effectively high-quality data, the "right" data, an adequate size of data and the right tool to perform data mining techniques. Data mining models can be categorized according to the tasks they perform. Data mining techniques are predictive or descriptive. Classification Prediction, Clustering, Association Rules are the data mining techniques from which Classification and prediction is a predictive model, but clustering and association rules are descriptive models. Classification recognizes patterns that describe the group to which an item belongs. Prediction is the construction and use of a model to assess the class of an unlabeled object or to assess the value or value ranges of a given object is likely to have. Forecasting is different from predictions because it estimates the future value of continuous variables based on patterns within the data. [5] Data mining provides various techniques to analyze the data. It also provides the methods for data preprocessing which have different processes such as data selection, Attribute reduction, dimensionality reduction, etc... Data Mining used in various area. It's useful for finding unknown patterns.

C. Neural Network

An artificial neural network (ANN), often just called a "neural network" (NN), is a mathematical model or computational model based on biological neural networks, in other words, is an emulation of biological neural system. It consists of an interconnected group of artificial neurons and processes information using a connectionist approach to computation. In most cases an ANN is an adaptive system that changes its structure based on external or internal information that flows through the network during the learning phase. [5] The artificial neural network is basic models of the genetic neuron system, is a extraordinarily parallel distributed processing system made up of highly interrelated neural computing essentials that can learn and thereby acquire knowledge and make it available for use. [2] Neural network provides various benefits such as High Accuracy, Noise tolerance, independent from prior assumption, ease of maintenance, and problem tolerance. The neural network is very most useful for dynamic behavioral application domain such as Fraud detection, telecommunications, medicines, marketing, bankruptcy prediction, insurance, the list goes on. Now days it will mostly used for financial, marketing and accounting like application domains. The only problem with neural network is to optimize number of neurons essential for explaining any problem and selection of data set which fully describes problem to be solved. [5]

D. Genetic Algorithm

Genetic Algorithms are generally used to improve the performance of neural network. Genetic algorithms provide a method of problem-solving which is based upon implementations of evolutionary processes. According to David J. Montana, Genetic algorithm is way of encoding solutions to the problem on chromosomes. It is used to assign real-valued weight and bias associated with each connection of neurons. A genetic algorithm repeatedly modifies a population of individual solutions. A genetic algorithm repeatedly modifies a population of individual solutions based on evolutionary process. It works using evolutionary solution to the problems for that evaluation function requires which returns rating for each chromosome and optimized algorithms for providing optimized solution.

III. COMPARATIVE ANALYSIS ON PREDICTION MODELS

Chun-Min Hung and Chun-Wu Yeh, propose novel concept to design TTCR and then searches the cumulative distance of the volatility is consistent. They input time series financial data and produce transaction type that is the investor has to buy or sell the financial products or wait for future market stability. In this model, the closing price of every five days used to calculate an average of upper and lower limits. They also calculate the distance and sequence them as per time horizon. to determine the trend from the pattern of waveforms in different time periods, and to evaluate the coverage between the range accumulated. At last the exception distance (δ) calculated between the actual price and expected price in a parabolic track. This will help investors to develop better strategies in stock markets. But they conclude that the future study will seek for estimating a more precise time span for early warning of stock prices. [1]

Yusuf Perwej and Asif Perwej, Propose model using ANN with Genetic Algorithm to predict BSE Market returns. In this model first they refer the investment theory and as per the theory they define the investment affection parameters which will reflect the financial product value. Then they will predict the market. Regression models have been used to forecast stock market time series. Machine learning is a procedure that begins with the detection of the learning area and trimmings with testing and using the results of the knowledge. These methods have been applied to market prediction for the most part to the forecast of the market on a daily basis on non linear type of data. They prepared the model for BSE Data appropriately trained; the network can be put to effective use in solving unidentified or untrained instances of the problem. They defining the relations among the different parameters by hiding neurons layer and assign the weight. Then they applied the trained genetic algorithm with back propagation. They applied randomness tests on the excess return time series the run and the test and we rejected randomness so model provides the management of inaccuracy and non linearity but their performance is influenced by the way that their weights are initialized. [2]

Ruiling Liu, Hengjin Cai*, Cheng Luo, propose the model using clustering with Manifold learning on Stock CSI data. Clustering analysis of stocks is necessary when investigating in stocks. By a synthetic evaluation index system to measure the correspondence of stocks, investors can use the estimated scope and assumption possible variation of stock price. They uses Manifold learning which acquires very good achievements in exploring inner law of nonlinear data. First they apply the preprocessing techniques on stock data. They also apply the dimensionality reduction by Isomap algorithm can reduce dimensionality of these high dimensional data. Reduced data points distribute in a two-dimensional coordinate system. By comparing the effects of clustering by Isomap and by LLE, they uses global geometric manifold learning algorithm which is based on Multidimensional Scaling to fits the linearity. They are provides the basic research concept and conclude that the complex linearity requires more complex environment. [3]

Zabir Haider Khan and his co author, propose the prediction model for share market using Artificial Neural Network. Trading shares and commodities were mostly based on perception. There are many variables that could affect the share market directly or indirectly. There are no significant relations between the variables and the price so to derive the any mathematical relation among the variables is too complex. So they uses the neural network by providing proper analyzed input and the main advantage of neural network is that it will ignore the input which do not contribute to output. The model also uses back propagation for the data pre processing and to remove erroneous data. The model assigns weights by each hidden layer neuron calculation and also calculate error rate. After this the prediction phase is start by trained network as per the input data. [4]

DR. YASHPAL SINGH, ALOK SINGH CHAUHAN, proves the usage of neural network with data mining. Neural networks are non-linear statistical data modeling tools which is used to define complex relationships among input and output. To acquiring knowledge from data set the data mining is useful. Neural network encompass of architecture or model, learning algorithm and activation function. Data mining task is dividing in two categories: predictive and descriptive. Descriptive data mining endow with information to understand what is occurrence inside the data without a determined idea. Predictive data mining works on unknown field values, and the system will estimate the unknown values based on previous patterns exposed form the database. They conclude that only data mining is not right tool for best result. The use of neural networks in data mining is a talented field of research especially given the ready availability of large mass of data sets and the reported ability of neural networks to detect and assimilate relationships between a large numbers of variables. [5].

IV. CONCLUSION

The researchers and investors attempt to outperform the market. The use of neural networks to forecast financial market value will be a continuing area of research. Neural network provides the unsurpassed accuracy to define the relationship among complex variables but it's depends on the data input to the neural network. So the inputted data must be preprocessed and also the selective measures are provided. It has been proven through the research that the estimation of the return on investment in financial markets through any of the traditional techniques is tedious, complex, costly and a time consuming process. The conclusion is that train the model with more input data set by combining neural network with data mining to generate more error free prediction of financial market product value.

REFERENCES

Journal Papers:

- [1] Chun-Min Hung and Chun-Wu Yeh*, “The Reliable Integrated Decision for Stock Price by Multilayer Integration Time-series of Coverage Reasonability” in Proceedings of the International MultiConference of Engineers and Computer Scientists 2009 Vol I IMECS 2009, March 18 - 20, 2009, Hong Kong.
- [2] Yusuf Perwej, Asif Perwej, “ Prediction of the Bombay Stock Exchange (BSE) Market Returns Using Artificial Neural Network and Genetic Algorithm”, Journal of Intelligent Learning Systems and Applications, 2012, 4, 108-119, doi:10.4236/jilsa.2012.42010 Published Online May 2012 (<http://www.SciRP.org/journal/jilsa>)
- [3] Ruiling Liu, Hengjin Cai*, Cheng Luo, “ Clustering Analysis of Stocks of CSI 300 Index Based on Manifold Learning” in Journal of Intelligent Learning Systems and Applications, 2012, 4, 120-126, doi:10.4236/jilsa.2012.42011 Published Online May 2012 (<http://www.SciRP.org/journal/jilsa>)
- [4] Zabir Haider Khan, Tasnim Sharmin Alin and Md. Akter Hussain, “Price Prediction of Share Market using Artificial Neural Network (ANN)”, International Journal of Computer Applications (0975 – 8887), Volume 22– No.2, May 2011
- [5] DR. YASHPAL SINGH, ALOK SINGH CHAUHAN, “NEURAL NETWORKS IN DATA MINING” in Journal of Theoretical and Applied Information Technology, © 2005 - 2009 JATIT, Available: <http://www.jatit.org>
- [6] Roll, R., 1989. Price volatility, international market links, and their implications for regulatory policies. Journal of Financial Services Research 3, 211-246.
- [7] Kimoto,T., asakawa, K., Yoda , M., and Takeoka, M. (1990),Stock market prediction system with modular neural network, in proceedings of the International Joint Conference on Neural Network, 1-6.
- [8] Samarth Agarwal,Manol Jindal,G.N.Pillai “Momentum Analysis based Stock Market Prediction using ANFIS”. In Proceeding of the International Multiconference of Engineering and Computer Scientists 2010 Vol.1, IMECS 2010, March 2010, Hong Kong
- [9] Jung Hua Wang, “Stock market trend prediction using ARIMA based Neural Networks” 0-7803-3210-5, IEEE 1996.