

Prediction of Stock Market using Data Mining and Artificial Intelligence

G. S. Navale
Savitribai Phule Pune
University
SITS Narhe, Pune-411041

Nishant Dudhwala
Savitribai Phule Pune
University
SITS Narhe, Pune-411041

Kunal Jadhav
Savitribai Phule Pune
University
SITS Narhe, Pune-411041

Pawan Gabda
Savitribai Phule Pune University
SITS Narhe, Pune-411041

Brij Kishor Vihangam
Savitribai Phule Pune University
SITS Narhe, Pune-411041

ABSTRACT

Predicting anything is very hard where the relationship between inputs and outputs are non-linear in nature. The prediction of stock market values is one of a challenging task of financial time series. Online application for buying and selling the shares is used in high amounts these days. The next step of this web application will be not just registering, buying and selling the shares but it will also be predicting the values for particular shares in the market. We are proposing the system which will study the database of shares and will give predictions according to it. With the help of study of neural networks the system will be designed and based on. For prediction particularly ARMA (autoregressive-moving-average) algorithm is used. Hence the system will be able to give highest probability predictions for particular shares.

General Terms

Data Mining, Prediction, Stock Market

Keywords

Artificial Neural Network, ARMA Algorithm, News articles, Text mining

1. INTRODUCTION

Data mining is analytic process design to explore data (usually large amount of data-typically business or market related- also known as “Big Data”) in search of consistent patterns and/or systematic relationships between variables, and then to validate the findings by applying the detected patterns to new.

Stock market is very volatile in nature. Prices of stocks changes almost instantly. Financial analysts who purchases stocks are not aware of all factors like inflation, economic growth affecting stocks prices. They do not have idea in which stocks to invest and sell. They can be easily manipulated by the stock brokers. Stock prices depend on news appearing in news articles. It is not possible for an average buyer to analyze such large amount of information .To deal with this problem Data Mining technique can be used. Data mining can automatically extract important information from large amount of data that is affecting the stock prices [9].

Predicting the stocks prices accurately can be done by Artificial Neural Network (ANN). The advantage of using ANN is that it can deal with both linear and non-linear data for forecasting the stock prices. Network is set of interconnected nodes and a node is a computational unit

which produces an output on receiving an input. The nodes can be both unidirectional and bidirectional. In unidirectional nodes, information can flow in one way while in bidirectional nodes, information can flow in both ways. So, ANN is neural network consisting of artificial neurons. ANN is inspired by the way our brain functions [10].

2. RELATED WORK

Table 1

| Paper ID | Methodology | Advantage | Disadvantage |
|----------|--|---|--|
| [1] | Genetic Algorithm, Support vector machines. | SVM transform the inputs into decision classes. There is correlation between prices of certain stocks. Considering closing, opening, mean, standard deviation and number of days for which correlation is found is considered | Various political, economic factors, company policy decide trends of markets are not considered while calculation. |
| [2] | Sentiment Analysis, Trading model. | They collected aggregating information from multiple online sources. They performed sentiment analysis on given data and filtered out dataset as a | It is necessary to analyze effects of applying different sentiments analysis methodology. |

| | | | |
|-----|--|--|--|
| | | <p>result of sentiment analysis and they found the ratio of sentimental signals.</p> <p>Based on this, they created on trading model to predict stock prices and trend of market.</p> | |
| [3] | <p>ANN(artificial neural network),</p> <p>Back-propagation algorithm</p> | <p>It can be used in field where accurate mathematical model cannot be produced, for example stock market</p> <p>It can deal with noisy data.</p> | <p>Designing is challenging as it requires tedious trial and error process.</p> <p>Selection of data set is complex.</p> |
| [4] | <p>Linear regression,</p> <p>Data mining.</p> | <p>Linear regression is used to perform operation data set where target values.</p> <p>It establishes relation between target values and predicted values.</p> <p>Data mining technique have more successful performance in predicting various fields as it uses hidden knowledge of data.</p> | <p>Calculations using linear regression are very complex.</p> <p>In linear regression, Accuracy is low.</p> |
| [5] | <p>Linear regression,</p> <p>Neural networks,</p> <p>Genetic</p> | <p>It is used to find accurate results among them.</p> <p>Helpful for gathering financial data.</p> | <p>Depend on sentiments and opinion over news content and</p> |

| | | | |
|-----|---|--|--|
| | <p>Algorithm,</p> <p>Support Vector Machine,</p> <p>Case based reasoning.</p> | <p>They are helpful to map the relations among financial product and financial news.</p> | <p>global events.</p> |
| [6] | <p>Typical price,</p> <p>Chwkin money flow indicator,</p> <p>Relative Strength Index.</p> | <p>It calculates the high, low and close value of the market.</p> <p>Also, it tends to give mid value so that customer can buy and sell share according to the values given</p> | <p>Problem is determining the probability that the relationships are not random at all market condition</p> |
| [7] | <p>Data collection,</p> <p>Feed-forward neural network.</p> | <p>Several machine learning techniques are used in parallel to predict most optimal stock market price.</p> <p>The main advantage is that it provide a very systematic approach and its ability to predict changes before they show up on the chart.</p> | <p>Requires large amount of historical data.</p> <p>It has very high time consuming factor depends on the accuracy of the data provided.</p> |
| [8] | <p>NewsCAT,</p> <p>Text preprocessing,</p> <p>Automatic text categorization.</p> | <p>It automatically analyzes and categorizes press releases derive stock trading recommendatio n from them.</p> <p>It can significantly outperform old trading strategies like buying and</p> | <p>Selection of categorization is poor.</p> <p>NewsCAT engine needs to be enhanced.</p> |

| | | | |
|------|--------------------------------|--|---|
| | | shorting stocks randomly immediately after press release. | |
| [9] | Data Mining. | It helps to find hidden pattern in from historic data that have probable predictive capability. It uses real time news to predict its effect on stocks. | It requires large amount of historic data. Large amount of data processing is required. |
| [10] | Artificial neural network(ANN) | It helps to build relation between non-linear input and output. It is very intelligent system and works like human brain. | ANN have not been fully explored. Prediction is satisfactory but still lot of improvement is needed. |

3. CONCLUSION

By using data mining and artificial intelligence, we can get accurate prediction result. Most of the authors have used methodologies in artificial intelligence to achieve accuracy and performance as shown Table 1. Still there is a need to improve the parameters accuracy and performance. This can be achieved with the help of Data Mining. Data Mining and Artificial Intelligence when put together will result in nearly accurate accuracy. The drawbacks of the other methodologies can be addressed by collaborating Artificial Intelligence and Data Mining. Investors gets an idea where to invest their valuable money.

4. REFERENCES

[1] Rohit Choudhry, Kumkum Garg, “A Hybrid Machine Learning System for Stock Market Forecasting” in International Journal of Computer, Electrical,

Automation, Control and Information Engineering Vol:2,No:3,2008.

- [2] Nann,Stefan , Krauss, Jonas, Schoder, Detlef , “Predictive Analytics on Public Data-The Case of Stock Markets” in the 21st European conference on Information Systems.
- [3] Manjul Saini, A.K.Singh, “ Forecasting Stock Exchange Market and Weather Using Soft Computing” in International Journal of Advanced Research in Computer Science and Software Engineering.ISSN:2277 128x,Volume 4,Issue 5,May 2014.
- [4] Farhad Soleimanian Gharehchopogh, Tahmineh Haddadi Bonab, Seyyed Reza Khaze, “ A LINEAR REGRESSION APPROACH TO PREDICTION OF STOCK MARKET TRADING VOLUME: A CASE STUDY” in International Journal of Managing Value and Supply Chains (IJMVSC) Vol.4, No. 3, September 2013.
- [5] Ms. Hiral R. Patel, Mr. Amit B. Suthar, Dr. Satyen M. Parikh, “ A Proposed Prediction Model for Forecasting the Financial Market Value According to Diversity in Factor” in Int. J. Computer Technology & Applications, Vol 5 (1),131-135.
- [6] K. Senthamarai Kannan, P. Sailapathi Sekar, M.Mohamed Sathik and P. Arumugam, “Financial Stock Market Forecast using Data Mining Techniques” in the International MultiConference of Engineers and Computer Scientists 2010 Vol I,IMECS 2010, March 17-19,2010, Hong Kong.ISSN:2078-0966(Online).
- [7] Zahid Iqbal, R. Ilyas, W. Shahzad, Z. Mahmood and J. Anjum, “ Efficient Machine Learning Techniques for Stock Market Prediction” in Int. Journal of Engineering Research and Applications, ISSN : 2248-9622, Vol. 3, Issue 6, Nov-Dec 2013, pp.855-867.
- [8] Marc-André Mittermaye, “Forecasting Intraday Stock Price Trends with Text Mining Techniques” in the 37th Hawaii International Conference on System Sciences – 2004.
- [9] Ruchi Desai, Prof. Snehal Gandhi, “Stock Market Prediction Using Data Mining” in International Journal of Engineering Development and Research, 2014 IJEDR | Volume 2, Issue 2 | ISSN: 2321-9939.
- [10] Prakash Ramani, Dr. P. D. Murarka, “Stock Market Prediction Using Artificial Neural Network” in International Journal of Advanced Research in Computer Science and Software Engineering. ISSN: 2277-128x, Volume 3, Issue 4, April 2013