



Data Mining Techniques in Stock Market

KEYWORDS

Stock Market, Data Mining, Applications of data mining techniques in stock market

Kainaz Bomi Sherdiwala

Dept. of Computer Application, Z.F. Wadia Women's & N.K.Jhota College of Commerce, Surat

ABSTRACT Data mining is being actively applied to stock market since 1980s.. This research paper has mainly concentrated on applications of data mining algorithms in stock markets. An overview of data mining techniques such as decision tree, neural network, association rules, factor analysis and etc in stock markets is provided.

Introduction

Stock market is a place where buying and selling of stocks/shares takes place. When an investor buys stocks of a certain company he becomes a part owner of that company according to the number of shares held by him. To make maximum profit, right investment should be made at the right time. The main function of a stock market is the dealings of stocks between investors. Stocks are grouped into industry groups according to their primary business focus. A transaction is the willing of an investor to sell some stocks and the request of another to buy them. Each stock is not only characterized by its price but also by many others variables.

The main variables are shown in the table below [7][8].

Table 1. Stock Variables

Variable Description

Price	Current price of a stock
Opening Price	Opening price of a stock for a specific trading day
Closing Price	Closing price of a stock for a specific trading day
Volume	Stock transactions volume (buy/sell)
Change	Opening and Closing stock value difference
Change (%)	Percentile Opening and Closing stock value difference

What is Data Mining?

Data are any facts, numbers, or text that can be processed by a computer. This includes: Operational or transactional data such as sales, cost, inventory, payroll & accounting; Non - operational data such as industry sales, forecast data & macro economic data and meta data such as logical data base design or data dictionary definitions. Data Mining is an analytic process designed to explore data and in search of consistent patterns and /or systematic relationships between variables, and then to validate the findings by applying the detected patterns to new subsets of data. In other words, Data mining is the extraction of hidden predictive information from large data bases. It is used to increase revenues and reduce costs. It finds patterns and correlations or relationships in data by using sophisticated techniques.

Application of Data Mining Techniques in Stock Markets

Data mining is an analytic process designed to explore 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 subsets of data. The ultimate goal of data mining is prediction and predictive data mining is the most common type of data mining and one that has the most direct business applications.

There are various data mining techniques which are applicable in stock market:

1) Application of decision tree in stock markets

Decision trees are excellent tools for making financial or number based decisions where a lot of complex information needs to be taken into account. They provide an effective structure in which alternative decisions and the implications of taking those decisions can be laid down and evaluated. They also help you to form an accurate, balanced picture of the risks and rewards that can result from a particular choice. In this section, we present some of the application of decision trees in stock markets. Decision trees are excellent tools for making financial or number based decisions where a lot of complex information needs to be taken into account. They provide an effective structure in which alternative decisions and the implications of taking those decisions can be laid down and evaluated. They also help you to form an accurate, balanced picture of the risks and rewards that can result from a particular choice. In a stock market, how to find right stocks and right timing to buy has been of great interest to investors. To achieve this objective, Muh-Cherng et al. (2006) present a stock trading method by combining the filter rule and the decision tree technique Listed companies' financial distress prediction is important to both listed companies and investors. Jie and Hui (2008) present a data mining method combining attribute-oriented induction, information gain, and decision tree, which is suitable for preprocessing financial data and constructing decision tree model for financial distress prediction. Accurately, forecasting stock prices has been extensively studied. Jar-Long and Shu-Hui (2006) provided a proposal to use a two-layer bias decision tree with technical indicators to create a decision rule that makes buy or not buy recommendations in the stock market.

2) Application of neural network in stock markets

Neural networks have been successfully applied in a wide range of supervised and unsupervised learning applications. Neural network methods are commonly used for data mining tasks, because they often produce compre-

hensible models. A neural network is a computational technique that benefits from techniques similar to ones employed in the human brain. It is nowadays a common notion that vast amounts of capital are traded through the stock markets all around the world. National economies are strongly linked and heavily influenced by the performance of their stock markets. Moreover, recently the markets have become a more accessible investment tool, not only for strategic investors but for common people as well. Consequently they are not only related to macroeconomic parameters, but they influence everyday life in a more direct way. Therefore they constitute a mechanism which has important and direct social impacts. The characteristic that all stock markets have in common is the uncertainty, which is related to their short and long term future state. This feature is undesirable for the investor but it is also unavoidable whenever the stock market is selected as the investment tool. The best that one can do is to try to reduce this uncertainty. Stock market prediction is one of the instruments in this process. The main advantage of neural networks is that they can approximate any nonlinear function to an arbitrary degree of accuracy with a suitable number of hidden units. Neural networks can forecast the buying and selling signs according to the prediction of future trends to stock market, and provide decision-making for stock investors so that the different investors could benefit from it. Neural network and time series models are used for forecasting the volatility of stock price index in two view points: deviation and direction.

3) Application of Clustering in Stock Markets

Clustering is a tool for data analysis, which solves classification problems. Its objective is to distribute cases (people, objects, events etc.) into groups, so that the degree of association can be strong between members of the same cluster and weak between members of different clusters. In clustering, there is no pre classified data and no distinction between independent and dependent variables. Instead, clustering algorithms search for groups of records. The algorithms discover these similarities. This way each cluster describes, in terms of data collected, the class to which its members belong. Clustering is a discovery tool. It may reveal associations and structure in data which, though not previously evident, nevertheless are sensible and useful once found. As part of a stock market analysis and prediction system consisting of an expert system and clustering of stock prices, data is needed. Stock markets are recently triggering a growing interest in the physicists' community. Basaltoa et al. (2005) apply a pair wise clustering approach to the analysis of the Dow Jones index companies, in order to identify similar temporal behavior of the traded stock prices. The objective of this attention is to understand the underlying dynamics which rules the companies' stock prices. In particular, it would be useful to find, inside a given stock market index, groups of companies sharing a similar temporal behavior. To this purpose, a clustering approach to the problem may represent a good strategy.

4) Application of Association Rules in Stock Markets

The associations' rules algorithm is used mainly to determine the relationships between items or features that occur synchronously in the database. For instance, if people who buy item X also buy item Y, there is a relationship between item X and item Y, and this information is useful for decision makers. Therefore, the main purpose of implementing the association rules algorithm is to find synchronous relationships by analyzing the random data and to

use these relationships as a reference during decision making. Association rule mining finds interesting associations and/or correlation relationships among large set of data items. Association rules shows attributed value conditions that occur frequently together in a given dataset. Mining association rules on large data sets has received considerable attention in recent years. Association rules are useful for determining correlations between attributes of a relation and have applications in marketing, financial, and retail sectors. Furthermore, optimized association rules are an effective way to focus on the most interesting characteristics involving certain attributes.

5) Application of Factor analysis in stock market

Factor analysis is particularly useful in situations where a large number of variables are believed to be determined by a relatively few common causes of variation. Also, it should be particularly useful for analyzing financial markets because if financial markets are efficient, nominal returns will be affected by default and market risk and by expected inflation and inflation uncertainty. Factor analysis models are used to examine hidden patterns of relationships for a set of stocks. Recent research on dynamic factor models finds that the information in a large number of economic time series can be effectively summarized by a relatively small number of estimated factors, affording the opportunity to exploit a rich base of information more likely to span the information sets of financial market participants than in previous analyses. In doing so, their study contributes to the empirical literature by evaluating both the potential role of omitted information in the estimated risk-return relation as well as the robustness of previous results to conditioning on richer information sets.

Conclusion

With the increase of economic globalization and evolution of information technology, financial data are being generated and accumulated at an unprecedented pace. As a result, there has been a critical need for automated approaches to effective and efficient utilization of massive amount of financial data to support companies and individuals in strategic planning and investment decision making. Data mining techniques have been used to uncover hidden patterns and predict future trends and behaviors in financial markets. The competitive advantages achieved by data mining include increased revenue, reduced cost, and much improved marketplace responsiveness and awareness. There has been a large body of research and practice focusing on exploring data mining techniques to solve financial problems.

REFERENCE

- (1) Jo Ting, Tak-chung Fu, and Fu-lai Chung. Mining of Stock Data: Intra- and Inter-Stock Pattern Associative Classification | (2) S.Karthik, K.K.Sureshkumar. Analysis of Stock Market Trend using Integrated Clustering and Weighted Rule Mining Technique. International Journal of Computer Science and Management Research Vol 1 Issue 5 December 2012. | (3) Rajesh V. Argiddi, Sulabha S. Apte. A Study Of Association Rule Mining In Fragmented Item-Sets For Prediction Of Transactions Outcome In Stock Trading Systems. International Journal of Computer Engineering and Technology (IJCET), Vol 3 Issue 2 September 2012. | (4) Rajesh V. Argiddi, Sulabha S. Apte. Future Trend Prediction of Indian IT Stock Market using Association Rule Mining of Transaction data. International Journal of Computer Applications (0975 – 8887) Volume 39– No.10, February 2012. | (5) Hongjun Lu, Jiawei Han, Ling Feng. Stock Movement Prediction And N-Dimensional Inter-Transaction Association Rules. | (6) Debashish Das and Mohammad Shorif Uddin. Data Mining And Neural Network Techniques In Stock Market Prediction: A Methodological Review. International Journal of Artificial Intelligence & Applications (IJAAA), Vol.4, No.1, January 2013. | (7) Gerasimos Marketos, Konstantinos Padiaditakis, Yannis Theodoridis, and Babis Theodoulidis. Intelligent Stock Market Assistant Using Temporal Data Mining. | (8) Mrs. Keerti. S. Mahajan & R. V. Kulkarni. Application Of Datamining Tools For Stock Market. | (9) Dattatray P. Gandhma, Ranjeetsingh B. Parihar, Rajesh V. Argiddi. An Optimized Approach to Analyze Stock market using Data Mining Technique | (10) Sachin Kambey, S. Thakur & Shailesh Jalori. Application of Data Mining Technique in Stock Market : An Analysis. | (11) Savinderjit Kaur And Veenu Mangat. Applications Of Data Mining In Stock Market. Journal of Information and Operations Management. | (12) D. Venugopal Setty, T.M.Rangaswamy, K.N.Subramanya. A Review on Data Mining Applications to the Performance of Stock Marketing. International Journal of Computer Applications (0975 – 8887) Volume 1 – No. 3. |