

# Inter-stock Trend Prediction of Stock Market using Outlier Mining and Association Rule Mining

R. V. Argiddi

Professor

Department of Computer Science and Engineering  
Walchand Institute of Technology, Solapur

S. T. Patel

Student M. E.

Department of Computer Science and Engineering  
Walchand Institute of Technology, Solapur

## ABSTRACT

With the advancement of storage techniques and Digitization of work in every field, the amount of stored data is tremendously increasing. Influence in Information Technology has caused a sizeable change in every sector of the digitized world. One of such sectors is the stock market where data changes constantly. The economy of the country is indicative of the stock market; this sector needs more support for its development in developing countries, which now rely to a great extent on Investments. Stock market generates a large amount of data on daily basis. Using Data Mining techniques like Clustering, Outlier Mining, Association Rule various operations will be performed to analyze the data and retrieve information. This information will serve us to predict the trend of the stock. Ups and downs in stocks of different companies may be related and so may be their trends. The historical data of such companies will be used to derive the relation to determine the collateral effect on the related stocks and the trend, if any.

## Keywords

Data Mining, Clustering, Outlier Mining, Clustering, Association Rule, Anomaly, Data Science.

## 1. INTRODUCTION

In today's technical era, all the sectors have been computerized. Stock market, from early days have a large amount of data; which makes it a candidate of data mining. If proper data mining techniques are applied on this data, a number of patterns could be retrieved. Data in stock market is a financial time series, which change dynamically and selectively. Difficulty in prediction arises in such time series because the problem is non-linear, non-stationary and have many noises [1, 5].

Contradicting to efficient market hypothesis [2], some trading price proves that market is not efficient in the real world [1], which makes the stock trend prediction possible. It is a common procedure for stockbroker to find relationship between different stocks by carefully analyzing the trade chart history. Many fields such as economic factors, political factor, social psychology of investment, etc., are some multidimensional element results in the price of one stock [2].

A various methods such as K-means Clustering, Artificial Neural Networks, Hypergraph Modeling, Association Rule, Fuzzy Sets, and Outlier Data Mining Algorithm can be applied in data mining. However, trends of stocks are inconsistent and unpredictability creates trouble to the investors and researchers. Outlier detection or anomaly detection is a problem to find patterns that do not comply with rules. Non-complying patterns are referred by various names according to the domains such as anomalies, outliers,

discordant observation, exceptions, aberration, surprises, peculiarities, contaminants [9].

The use of various data mining techniques like Data Preprocessing, Clustering, Outlier Detection, Association Rule, etc., in combination with each other, it may help to provide better prediction of stock market and can even examine the relation between the companies in more precise way.

## 2. LITERATURE REVIEW

In the previous research, various data mining techniques for stock price prediction and stock trend prediction are proposed.

**Zhao, Lei and Wang, Lin [1]** have focused on a novel outlier mining algorithm to detect anomalies on volume sequence of tick-by-tick data for stock market. This paper concludes with successfully locating of anomaly from the high frequency tick-by-tick data. Comparing with the traditional clustering algorithms, proposed algorithms are more effective and more practical.

**Yongen Luo, Jicheng Hu, Xiaofeng Wei, Dongjian Fang, Heng Shao [2]** have proposed the hypergraph model of clustering algorithm. In this paper for prediction, they have considered a stocks trends in recent days and other stocks trends that lie in the same cluster. The accuracy rate under this model is claimed to be more than 50%.

**Suraiya Jabin [3]** have determined a computational approach for stock market prediction. With this approach, the technique used is Feed-forward Artificial Neural network (ANN) which concludes that ANN capable of extracting vital information from a huge set of data and hence plays a vital role in stock market prediction.

**Yu-Feng Jiang, Chun-Ping Li, Jun-Zhou Han [4]** have considered an algorithm that discovers the correlation of stock price with respect to the notion of time series having recurring patterns. The characteristics of their approach is to make predictions based upon the occurrence of recurring pattern.

**Priti Saxena, Bhaskar Pant, R. H. Goudar [5]** have a perspective on pattern based data mining approach which is used to convert the numeric stock data into symbolic notations. Further analysis is carried out through apriori and proposed reverse apriori concept.

## 3. PROPOSED METHODOLOGY

The proposed method uses number of different data mining techniques for recognizing the ups and downs in the patterns of the stock. The research work will consist mainly of Clustering, Outlier Mining and Association Rule Mining; forming the basis of this methodology.

Following are the steps that will form the functional flow.

1. Prepare Tick-by-Tick Data of various stocks.
2. Form Clusters of tick-by-tick data.
3. Find anomalies using outlier Mining Algorithm.
4. Find Relations between stocks using Association Rule Mining.

The research work for stock trend prediction will require Tick-by-Tick data of various stocks. Tick-by-Tick data proves more helpful as it records every trade for every stock in the market.

Data Preprocessing would derive a ratio matrix which will provide us the price in price sequence and volume number in volume sequence.

In the further part, clustering is done on the data. Applying an Outlier Mining Algorithm on the clusters, anomalies can be obtained which will help in predicting the future trend of the stocks.

Association Rule Mining will help to determine inter stock relation among the stocks. Apply Association Rule Mining on the results of trend prediction obtained in the previous step.

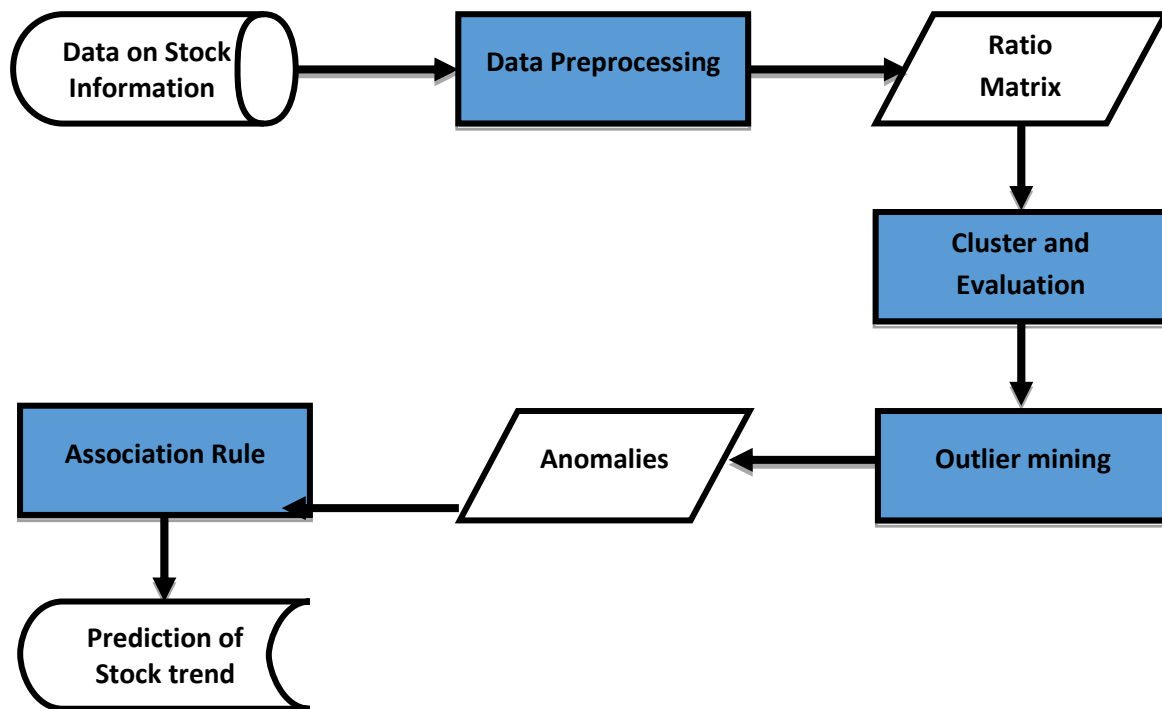


Fig1: System Architecture

#### 4. CONCLUSION

In developed and developing countries, stock market is accounted as a major economic factor. Not only for the countries progress but also for the individuals as it gain good profits. Outlier data Mining Algorithm would help obtaining the future trend of the stock. Association Rule Mining and Pattern Detection would be helpful to find the whether stock of other companies are related.

#### 5. REFERENCES

- [1] Zhao L., Wang L. Price Trend Prediction of Stock Market Using Outlier Data Mining Algorithm., IEEE, 2015, DOI 10.1109/BDCLOUD.2015.19
- [2] Yongen Luo, Jicheng Hu, Xiaofeng Wei, Dongjian Fang, Heng Shao, Stock Trend Prediction Based on Hypergraph Modeling Clustering Algorithm, IEEE, 2014, 978-1-4799-2030-3
- [3] Suraiya Jabin, "Stock Market Prediction using Feed-forward Artificial Neural Network", International Journal of Computer Application, August 2014, Volume 99 – No. 9
- [4] Yu-Feng Jiang, Chun-Ping Li, Jun-Zhou Han, Stock Temporal Prediction Based on the Time Series Motifs,

Proceedings of the English International Conference on Machine Learning and Cybernetics, Baoding, July 2009.

- [5] Priti Saxena, Bhaskar Pant, R. H. Goudar, Inter-Transactional Pattern Discovery Applying Comparative Apriori and Modified Reverse Apriori Approach., IEEE 8th Proceedings International Conference on Intelligent System and Control (ISCO), 2014.
- [6] P. K. Padhiary and A. P. Mishra, Development of improved artificial neural network model for stock market prediction, International Journal of Engineering Science and Technology, Vol. 3, 2011, pp. 1576-1581.
- [7] Stein E 1, Stein J 2. Stock Price Distributions with Stochastic Volatility: An Analytic Approach [J]. Review of Financial Studies, 1991, volume 4(4):727-752(26).
- [8] F. Allen and G. Gorton. Stock price manipulation, market microstructure and asymmetric information. European Economic Review, pages 624C630, 1992.
- [9] Varun Chandola, Arindam Banerjee, "Anomaly Detection: A Survey", ACM computing surveys, 2009, Volume 41 Issue 3 Article No. 15.