

“A Literature Review on Data Mining of Stock Data By Modifying Clustering Algorithm for Effective Decision Making for Future Stock”

Sonia Sawale, Dinesh Sahu, Amitesh Paul

Abstract— Classification and patterns extraction from customer data is very important for business support and decision making. Timely identification of newly emerging trends is needed in business process. The main Aim and the objective of this is to get better decision making for improving sale, services and quality as to identify the reasons of dead stock, slow-moving, and fast-moving products, which helps for business support, investment and surveillance. In this paper we proposed an algorithm for mining patterns of large stock data to predict Factors affecting the sale of products. In the first phase, we divide the stock data in three different clusters on the basis of product categories and sold quantities i.e. Dead-Stock (DS), Slow-Moving (SM) and Fast-Moving (FM) using K-means algorithm. In the second phase we have proposed Most Frequent Pattern (MFP) algorithm.

Index Terms— K-Means ,Dead Stock ,Slow Moving ,Fast Moving

I. INTRODUCTION

In every organization Data is very important for their business. Data that was measured in gigabytes until recently, is now being measured in terabytes, and will soon approach the petabyte range. In order to achieve our goals, we need to fully exploit this data by extracting all the useful information from it. Unfortunately, the size and complexity of the data is such that it is impractical to manually analyze, explore, and understand the data. As a result, useful information is often overlooked, and the potential benefits of increased computational and data gathering capabilities are only partially realized. Sale data classification has different market trends. Some clusters or segments of sale may be growing, while others are declining. The information produced is very useful for business decision making. Decision can take place on the basis classification of Dead-Stock (DS), Slow-Moving(SM) and Fast-Moving (FM) of the sale.

Sonia Sawale, PG Student ME Computer Science & Engg, Rkdf University, Agnoss college of Technology, Bhopal, 462033, Madhya Pradesh.

Dinesh Sahu, Professor Department of Computer Science&Engg, Rkdf University, Agnoss college of Technology, Bhopal, 462033, Madhya Pradesh.

Amitesh Paul, Professor Department of Computer Science&Engg, Rkdf University, Agnoss college of Technology, Bhopal, 462033, Madhya Pradesh.

Segment by- segment sales forecasting can produce very useful information.

The forecasting can be short term, midterm and long term. Long term forecasting may not produce accurate predictions.

II. LITERATURE REVIEW:

Data mining researchers often try to find most feasible and efficient methods for extraction of useful patterns from stock data. Most of the research regarding stock data mining uses the history of transactions as it likely that may persist in future. These can help to predict the customer behavior and future trend.

M. C. Lo [26] considered a model for inventory decision support system in which ordering quantity, ordering cost, safety factor, lead time and backorder discounts are decision variables, the algorithm is applied to find the optimal solution for the case where the lead time demands follows a general distribution.

J. ting et al [22] proposed a pattern based stock data mining approach for intra-stock mining which perform focuses on finding frequently appearing pattern for the stock time series data and inter stock mining which discover the strong relationship among the several stocks.

L. K. Soon et al [27] generate a list of stocks which are influential to Kuala Lumpur Composite Index (KLCI), and then produce classification rules, which denotes the inter-relationships among the stocks in terms of their trading performance with respect to KLCI. The DCX case study [28] a survey on the classification on the data mining technique in car manufacturing domain, which help in accurate prediction of future demand for car.

III. PROPOSED METHOD

Our proposed approach is a two phased model. First we generate clusters using K-Mean algorithm, and then MFP is designed for counting frequencies of items under their specified attributes. The block diagram of the whole process is given in figure 1. In phase-1 the first step is to collect sample data from real store inventory data. We have process the data to remove the noise first, so the incomplete, missing and irrelevant data are removed and formatted according to the required format. The Most Frequent Patter is shown below

“A Literature Review on Data Mining of Stock Data By Modifying Clustering Algorithm for Effective Decision Making for Future Stock”

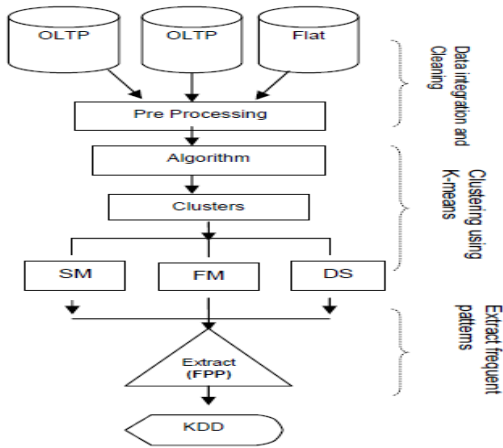


Figure 1: Block Diagram of Proposed Architecture

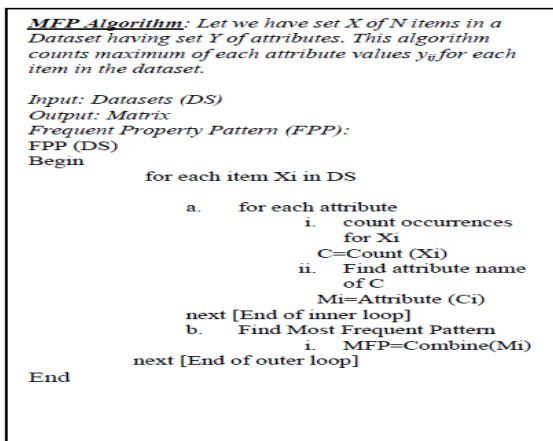


Figure 2: Proposed Algorithm for Frequent Pattern extraction

IV. CONCLUSION

In this project the problem of pattern discovery from stock data mining is addressed. Hybrid clustering association mining approach is proposed to classify stock data and find compact form of associated patterns of sale

REFERENCES:

- [1]. Abubakar, Felix “Customer satisfaction with supermarket retail shopping”, 2002.
- [2]. <http://www.roselladb.com/sales-trendforecast.htm>, visited January, 2010.
- [3]. M. Braglia, A. Grassi, R. Montanari “Multiattribute classification method for spare parts inventory management” 2004.
- [4]. L. Cao, C. Luo, J. Ni, DanLuo, C. Zhang “Stock Data Mining through Fuzzy Genetic Algorithm” Proceeding of JCIS s , 2008.
- [5]. Jo Ting, “Mining of stock data: inter- and inter-stock patternassociative classification” proceedings of 2006 international conference on data mining Las Vegas,USA, June 2006.
- [6]. Darken, C. Moody, J. Yale Comput. Sci., New Haven, “ Fast adaptive k-means clustering” IEEE- 2002
- [7]. <http://maya.cs.depual.edu/~mobasher/webminer/survey/node25.html> visited 2008.
- [8]. L.K.Soon, Sang Ho Lee, “An Emperical Study of Similarity Search in Stock Data” Australian Computer Society, Second International Workshop on Integrating AI and Data Mining, AIDM-2007. pp 31-38